Proceedings of International Academia
ISSN: 3083-2330
Volume 1, Issue 1, 2026
pp. 1-18

The Impact of Shadowing Technique Using eJOY on Improving the Final Sound Pronunciation

Duyen Le*
Nong Lam University Ho Chi Minh City, Vietnam
*Corresponding author email: duyen.lethimy@hcmuaf.edu.vn
https://doi.org/10.65956/procia.2026.66

Abstract

This study examines the effectiveness of eJOY-supported shadowing in improving English word-final consonant production among Vietnamese adult EFL learners. Using a classroom-based action research design, the intervention was implemented over two months in a communication course with ten learners at A2-B1 proficiency levels. The eJOY platform, which incorporates automatic speech recognition and real-time feedback, was used to facilitate structured shadowing practice. Data were collected through pre- and post-tests to assess pronunciation accuracy, weekly recordings to monitor changes in students' pronunciation and provide timely instructional support, and a post-intervention questionnaire to gather learners' perceptions and experiences. Findings revealed clear pre-post gains in overall pronunciation scores and increased consistency in learners' attempts to articulate common final consonants, alongside uniformly positive evaluations of feasibility and perceived benefits. The findings suggest that eJOY-supported shadowing is a practical, technology-enhanced option for targeting persistent final-sound difficulties in adult EFL contexts.

Keywords shadowing technique, final sound pronunciation, eJOY

Article history Received: 15 Apr 2025 | Accepted: 05 May 2026 | Available: 05 May 2026

INTRODUCTION

Understandable pronunciation is a core component of communicative competence and a long-standing priority in language instruction (Pourhosein Gilakjani, 2016a). In spoken interaction, pronunciation influences how efficiently listeners decode a speaker's message, shaping intelligibility and communicative success (Evers & Chen, 2022). Because meaning is interpreted in real time, even "small" segmental deviations may disrupt comprehension when they affect contrasts or grammatical marking (Derwing & Munro, 2015). Accordingly, pronunciation instruction remains essential for helping learners participate effectively in social and professional communication through the target language (Martinsen et al., 2017).

In the Vietnamese EFL context, a persistent and pedagogically consequential difficulty concerns the production of English word-final consonants. Due to differences in Vietnamese and English phonotactics, Vietnamese learners often omit, weaken, substitute, or modify codas, particularly in connected speech (Ha, 2005; Nguyen, 2007; Tang, 2007). These patterns are not merely accent-related: final sounds frequently carry lexical contrasts and grammatical information (e.g., plural -s/-es, third-person -s, past tense -ed), and their underproduction can reduce clarity and grammatical precision in spoken communication (Crystal, 2011; Kobilova, 2022). For adult learners, such

patterns may become stable over time, making it necessary to implement focused and sustained practice that targets final sounds explicitly rather than treating them as incidental aspects of general speaking practice (Nguyen, 2007).

One instructional approach that may be particularly relevant for this goal is shadowing, in which learners repeat speech immediately after hearing it with minimal delay, aiming to track and imitate a model in real time (Lambert, 1992). Shadowing has been discussed as a potentially useful technique for helping adult learners adjust entrenched pronunciation habits (Ricard, 1986), and it has been used in EFL pedagogy to develop listening and oral performance through intensive perception-production coupling (Hamada, 2018). From a pronunciation-learning standpoint, shadowing is promising because it repeatedly exposes learners to a stable spoken model while requiring immediate production, thereby creating frequent opportunities for learners to attend to timing and segmental detail (Hamada, 2014). However, implementing sufficiently frequent and consistent pronunciation practice can be challenging in adult EFL classes with limited time (Derwing & Munro, 2015). Technology-supported practice can increase access to input and enable repeated rehearsal beyond classroom hours (Lord, 2005; Pourhosein Gilakjani & Ahmadi, 2011). In this study, the eJOY platform was selected because it integrates interactive subtitles and immediate speech-recognition feedback within the same interface, unlike many video-based platforms that primarily support passive viewing or vocabulary learning. By enabling repeated auditory exposure followed by immediate oral reproduction, the platform supports structured shadowing practice, an approach shown to enhance learners' attentional focus on segmental detail (Hamada, 2014).

Although previous research has documented Vietnamese learners' persistent difficulty with English word-final consonants and highlighted the potential of scaffolded shadowing for pronunciation development (Foote & McDonough, 2017; Ha, 2005; Hamada, 2014, 2015, 2016a, 2018; Nguyen, 2007; Tang, 2007) evidence remains limited on whether a structured, technology-supported shadowing intervention yields feature-specific gains in word-final consonant production among Vietnamese adult learners, and on how such gains are reflected in shifts in error patterns over time. Addressing this gap, the present proceedings study examines the effects of eJOY-supported shadowing on learners' word-final consonant production using pre-/post-test evidence, supplemented by qualitative analysis of recurrent coda-related error patterns in recordings.

Research Questions

Building on the concerns and opportunities above, this action research investigates whether eJOY-supported shadowing can improve Vietnamese adult EFL learners' production of English final sounds and how learners evaluate the practicality and perceived benefits of this intervention. The study addresses the following research questions:

(1) To what extent does eJOY-supported shadowing improve Vietnamese adult EFL learners' production of English word-final consonants?

(2) How do learners evaluate the feasibility and perceived benefits of eJOY-supported shadowing for improving word-final consonants?

LITERATURE REVIEW

Pronunciation and Vietnamese Learners' Word-Final Consonant Difficulty

Gilakjani (2016b) defines pronunciation as the action that produces specific sounds. Pronunciation, while a component of phonetics, encompasses more than just sound production; it is intrinsically linked to both articulation and auditory perception (McLeod et al., 2022; Younes & Mueller, 2022). Sewell (2016) and Szyszka (2016) concur, framing pronunciation as the nuanced orchestration and interpretation of a language's segmental and suprasegmental elements. In dissecting these elements, segmental sounds encompass vowel sounds, consonant sounds, and sound combinations, while suprasegmental sounds encompass stress, rhythm, pitch, and intonation. Learning a new language frequently requires acquiring an entirely distinctive way of using speech systems; this is particularly the case for Vietnamese speakers learning English (Ha, 2005).

The final sound, also known as the word-final consonant, refers to the phoneme occurring at the end of a word or syllable (Carr, 2019). It is an integral component of English pronunciation that contributes to the clarity and intelligibility of spoken language. Achieving accurate final consonant sound pronunciation is essential for preventing confusion, ensuring effective communication, and enhancing overall language proficiency (Crystal, 2011; Derwing & Munro, 2015). Final sound omission refers to the tendency of learners to neglect or omit the consonant or vowel sounds occurring at the end of words or syllables in English speech. Tang (2007) discovered that Vietnamese language learners frequently simplify or omit the consonants in English words' final syllables. When pronouncing English word-final consonants, Vietnamese speakers will try to add schwa, remove the sound, or replace it with a sound that is more similar to their own tongue (Nguyen, 2007). According to Ha (2005), the phenomenon of deleting or reducing final syllable clusters among native Vietnamese speakers has the potential to cause a breakdown in communication. This tendency is not only habitual but is founded in the Vietnamese language's restricted phonotactics, which differ from English phonology (Nguyen, 2007). The distinct phonological patterns between Vietnamese and English pose a particular challenge for learners of English in Vietnam (Schuberg et al., 2013).

While most Vietnamese speakers' omission of English word-final sounds has piqued the interest of researchers (Ha, 2005; Ngo, 2005; Nguyen, 2007; Osburne, 1996; Tang, 2007), there have been few studies on structured classroom interventions that explicitly target word-final consonant accuracy and track feature-level change over time. This gap points to the need for pedagogical approaches that provide repeated, focused practice on codas and support learners in noticing and repairing recurring errors, particularly for adult learners whose pronunciation routines may be stable and resistant to change.

Shadowing as a Pronunciation-Focused Practice Technique

Shadowing, an instructional technique commonly utilized to enhance the listening skills of EFL learners, has its origins as a teaching method for simultaneous interpretation (Hamada, 2018; Lambert, 1992). Tamai (1992) characterized it as an engaged and cognitively intensive process wherein learners mirror the spoken words they hear by articulating them distinctly and concurrently absorbing the auditory input. Lambert (1992) describes shadowing as a paced, auditory tracking exercise that involves immediately vocalizing presented auditory stimuli. This distinguishes shadowing from delayed repetition tasks where learners have more time to plan and reflect before

producing language (Hamada, 2016a). Because shadowing requires rapid perception–production alignment, it is often characterized as a cognitively demanding activity that promotes bottom-up processing and sustained phonological attention (Hamada, 2018; Shiki et al., 2010). It has also been discussed as a potential means to help adult learners adjust entrenched pronunciation patterns through intensive imitation (Ricard, 1986).

The utility of shadowing in pronunciation instruction remains a hot topic in the domain of language acquisition. Certain critics contend that shadowing, due to its controversial audiolingual approach, could potentially result in pointless repetition resembling that of a parrot, which runs counter to the current emphasis on creative output in language learning environments (Bovee & Stewart, 2009). Nonetheless, numerous studies on shadowing have reported improvements in several dimensions of oral performance, including fluency and aspects of pronunciation and comprehensibility, although changes in accentedness tend to be limited (Foote & McDonough, 2017; Hamada, 2016b, 2018). Studies have also documented benefits for prosodic features such as intonation and rhythm (Hsieh et al., 2013; Hamada, 2014; Harmon, 2014). In Southeast Asian contexts, learners' perceptions of shadowing are frequently positive, and perceived gains include increased confidence and speaking performance (Luu & Do, 2023). Comparative findings further suggest that shadowing may yield gains that are as strong as - or stronger than - some alternative practice types, with improvements extending beyond pronunciation to broader spoken performance (Salim et al., 2020).

Nevertheless, the literature also highlights boundary conditions. Hamada (2016a) notes that shadowing effects may be more pronounced among higher proficiency learners because accurate tracking requires robust perception and rapid production. Implementation features also matter: learners may benefit from scaffolding such as previewing content before shadowing (Hamada, 2014) and recording/reviewing their performance to support self-monitoring, since learners can overlook their own output while concentrating on following the model (Hamada, 2015). Importantly, while many studies report broad pronunciation gains, fewer isolate a specific segmental target. Evidence remains limited regarding whether shadowing can systematically improve word-final consonant accuracy.

Technology Support and the Opportunity of eJOY

Technology-enhanced pronunciation practice is often argued to increase access to input and provide flexible opportunities for repeated practice, which can be difficult to sustain solely through classroom instruction (Lord, 2005; Gilakjani & Ahmadi, 2011). In shadowing contexts, mobile tools may also make repeated practice more feasible outside class, supporting intensity and consistency of exposure and production (Kessler, 2018). Foote and McDonough (2017), for example, reported gains in comprehensibility and fluency in an eight-week mobile-supported shadowing program. Empirical studies on mobile-assisted pronunciation learning generally report increased learner engagement and measurable gains in fluency or comprehensibility, particularly when practice is sustained over time (Foote & McDonough, 2017). However, many of these tools prioritize general speaking fluency or vocabulary development rather than targeting specific segmental features. These considerations underscore the need to examine how a platform such as eJOY may function when applied to focused segmental practice.

eJOY is a Vietnamese-developed English learning platform that provides video-based input and practice functions relevant to shadowing, including access to model audio/video, recording, repeated rehearsal, and automated feedback (refer to Appendix A). Action research on eJOY, however, is still limited. A recent study by Saskiana et al. (2024) offers preliminary evidence that eJOY use may improve learners' overall speaking performance in an Indonesian high school context. However, pronunciation in that study was embedded as one component of a broader speaking rubric (with vocabulary, grammar, fluency, and content), making it difficult to determine which pronunciation features improved and to what extent. In addition, the intervention was framed as general learning with eJOY rather than as a shadowing-centered program explicitly designed to address persistent segmental difficulties such as English word-final consonants. The learner population and instructional context also differ substantially from the Vietnamese adult EFL setting, where word-final consonant omission has been repeatedly documented as a widespread issue with implications for intelligibility and communicative effectiveness (Ha, 2005; Nguyen, 2007). Addressing this gap, the present study investigates the effects of eJOY-supported shadowing on Vietnamese adult learners' word-final consonant production using pre-/post-test evidence supplemented by qualitative analysis of error patterns across recordings.

METHODOLOGY

Action Research

The present study adopted a classroom-based Action Research (AR) design. Action Research differs from quasi-experimental research in that its primary purpose is not to establish generalizable causal relationships between controlled variables, but to address a context-specific pedagogical issue through systematic cycles of planning, action, observation, and reflection (Dick, 1993; Kemmis et al., 2014). Although pre- and post-test measures were incorporated to document learning outcomes, these functioned as evaluative components within a practitioner-led intervention rather than as elements of a controlled experimental comparison. This design aligns with the study's dual objective of examining pronunciation development while also exploring the practical effectiveness and learner perceptions of a technology-enhanced approach in an authentic classroom setting (Baumfield et al., 2013). Within this framework, the teacher-researcher actively implemented and monitored the instructional response to persistent word-final consonant difficulties, drawing on multiple sources of classroom evidence to inform ongoing pedagogical decisions (Cohen et al., 2017). Such an approach is particularly appropriate for pronunciation development in adult learners, whose entrenched speech patterns often require sustained, feedback-rich practice (Mertler, 2019; Omar & Umehara, 2010).

Participants

The study was conducted at an English center in Vietnam, where participants were enrolled in a communication course meeting twice weekly for two hours per session. Participants were ten Vietnamese adult learners (aged 23-35) with pre-intermediate to intermediate proficiency (approximately A2-B1), based on course placement and classroom performance. Participants were selected through convenience sampling, as the intact class taught by the teacher-researcher was identified as experiencing persistent difficulty with English word-final consonants and agreed to participate in the intervention. All participants were L1 Vietnamese speakers born and educated in

Vietnam and were employed in diverse professional domains (e.g., architecture/engineering, business/administration, science/technology). None of the participants had prior experience using eJOY for structured shadowing practice, although some reported occasional exposure to mobile language-learning applications for vocabulary review or listening practice. Therefore, brief orientation and procedural training were provided at the beginning of the intervention to ensure that all learners understood how to use the app's shadowing and feedback features consistently. Participation in the study was voluntary. Learners were informed of the research purposes, data collection procedures, and their right to withdraw at any time without academic penalty. These students are capable of engaging in everyday conversations on familiar topics. In the classroom, they demonstrate enthusiasm and active participation, with the aim of improving their communication skills for job promotion or personal purposes. However, they often struggle with pronouncing final sounds in English, which hinders their ability to effectively communicate in the language. All participants had reliable Internet access and personal digital devices (smartphones and/or computers), enabling participation in the technology-supported components of the intervention.

Materials

The intervention integrated eJOY as the primary platform to support shadowing practice. eJOY was selected because it offers features relevant to pronunciation practice, including: (a) access to short video-based input, (b) support for repeated listening and recording, (c) phonetic information for vocabulary items, and (d) automated feedback intended to guide learners' attention to pronunciation accuracy. In this study, eJOY was used primarily as a practice environment to increase the amount and consistency of shadowing practice. Authentic short videos available on eJOY served as shadowing materials. To ensure feasibility for A2-B1 learners and to reduce cognitive overload, videos were selected using the following criteria: (1) relevance to course topics, (2) manageable length (each clip under approximately 1 minute 30 seconds, maximum 150 words), (3) clear speech and moderate pace, and (4) availability of subtitles and translation support within the platform. Selecting course-relevant clips also ensured that at least part of the vocabulary and topic content was familiar to learners, which helped learners allocate more attention to pronunciation rather than comprehension. Google Drive was used to organize data collection and feedback exchanges. Each participant was assigned a private folder to upload weekly shadowing recordings and screenshots of practice results. The teacher-researcher used the same folder to return feedback files or comments.

Research Instruments

This study employed two primary research instruments: (1) a pronunciation test administered as pre-test and post-test, and (2) a post-intervention questionnaire designed to capture learners' perceptions of the instructional approach. The pronunciation test required participants to produce controlled speech samples containing target word-final consonants. Recordings were evaluated using a standardized pronunciation rubric (see Appendix B). To enhance scoring reliability and minimize researcher bias, all recordings were anonymized and independently rated by two experienced EFL instructors. Discrepancies were resolved through discussion to ensure consistency of scoring. The second instrument was a post-intervention questionnaire developed for this study (see Appendix C). The items targeted key perception domains relevant to the intervention, including enjoyment,

perceived usefulness for final-sound improvement, increased awareness of word-final consonants, motivation, confidence, ease of use, and intention to continue using the method.

Data Collection and Data Analysis

This study used a mixed-methods approach to evaluate the effectiveness of eJOY-supported shadowing for improving pronunciation. A pre-test was conducted prior to the intervention to establish baseline pronunciation performance, followed by a post-test after the intervention to examine changes in production. Weekly recordings were also collected throughout the instructional period to monitor changes in students' pronunciation and provide timely instructional support within the action research framework. To address Research Question 1, pre-test and post-test recordings were compared to examine changes in pronunciation accuracy. The resulting scores were analyzed using descriptive statistics, including comparisons of means and standard deviations to determine the extent of improvement. Given the small sample size, the analysis also considered individual developmental trajectories. In addition to score-based comparisons, recordings were subjected to a structured error analysis of word-final consonants. Tokens were categorized using a predefined set of coda-related error types (e.g., deletion, substitution, epenthesis, cluster simplification), and patterns were compared across pre- and post-test recordings to characterize developmental shifts. A subset of the data was double-coded by the two raters to confirm the applicability of categories and to support the trustworthiness of the qualitative coding. This qualitative analysis was used to contextualize and interpret the quantitative gains. To address Research Question 2, questionnaire responses were analyzed quantitatively using descriptive statistics to summarize overall response trends and variability across participants.

Procedure

The investigation followed Lewin's (1946) action research cycle of planning, acting, observing, and reflecting across an approximately two-month period. In the planning phase, the teacher-researcher identified learners' persistent omission of English word-final consonants and designed a structured shadowing protocol delivered through eJOY. At the outset of implementation, participants were trained in core shadowing procedures with explicit attention to word-final consonants, and they were oriented to using eJOY and submitting weekly recordings via individual Google Drive folders to standardize data collection. The acting phase involved systematic eJOY-supported shadowing practice outside class over eight weeks using curated short video clips aligned with course themes and appropriate for A2-B1 learners; participants submitted at least one recording per week. In the observing phase, a recorded pronunciation pre-test and a parallel post-test were administered to document change in word-final consonant production, weekly recordings were reviewed to track developmental patterns, and a post-intervention 7-item Likert questionnaire was collected to capture learners' evaluations of enjoyment, perceived benefits, and feasibility. Finally, in the reflecting phase, the researcher synthesized test outcomes, weekly performance evidence, and questionnaire trends to evaluate intervention impact and to inform minor adjustments to feedback and practice guidance, as well as to interpret whether learners were moving toward more stable, proceduralized production of final consonants.

FINDINGS

RQ1: Effects of eJOY-Supported Shadowing on Final Sound Pronunciation

Pre-Post test evidence of improvement

To address Research Question 1, which examined the extent to which eJOY-supported shadowing improved learners' production of word-final consonants, a paired-samples t-test (Table 1) was conducted to determine whether the observed improvement was statistically significant. The analysis revealed a highly significant difference between pre-test and post-test scores, $t(9) = 28.25$, $p < .001$.

Table 1. *Pretest & Posttest Mean Scores and Paired t-test Results*

Test	Mean Score	Standard Deviation	T-value	p-value
Pretest	53.0	5.715		
Posttest	69.2	4.779		
Difference	+16.2		-28.248	.000

*n = 10

This finding suggests that the improvement in final sound pronunciation scores was unlikely to be due to chance and is consistent with a meaningful intervention effect. The detailed results from Table 2 indicate a clear upward shift from pre-test to post-test performance across the entire group. As shown in the individual score profile, all 10 participants improved from pre-test to post-test, with gains ranging from +12 to +18 points. The group mean increased from 53.0 on the pre-test (SD = 5.72) to 69.2 on the post-test (SD = 4.78), yielding an average gain of +16.2 points. Notably, the post-test scores also showed slightly reduced variability, suggesting that learners' performance became not only higher but also more consistent after the intervention.

Table 2. *Learners' Pretest and Posttest Output*

Respondent	Pretest	Posttest
1	52	68
2	48	65
3	60	72
4	55	70
5	45	62
6	58	75
7	50	68
8	62	77
9	47	64
10	53	71
Overall	53	69

These results provide quantitative support that implementing shadowing practice through eJOY was associated with substantial gains in learners' final sound pronunciation performance. The uniform improvement across all participants is particularly noteworthy, as word-final consonant production

is often resistant to change for adult EFL learners. The consistent gains align with the theoretical rationale underpinning the intervention: repeated, focused, feedback-informed shadowing can strengthen perception–production mapping and support the proceduralization of segmental features - here, word-final consonants - under time pressure. At the same time, because this study employed a one-group pretest-posttest action research design, the findings should be interpreted as evidence of improvement within this instructional context rather than definitive causal proof under controlled experimental conditions.

Development and error patterns in final sound production

In addition to overall pre-post score gains, learners' recordings were qualitatively coded across both tests to trace changes in word-final consonant production, providing a feature-level account that contextualizes the quantitative results. A consistent developmental tendency was increased stability in the production of high-frequency English codas, particularly alveolar stops (/t/, /d/) and alveolar fricatives (/s/, /z/). In the pre-test, these segments were frequently reduced or omitted, especially when learners produced longer stretches of speech where maintaining fluency appeared to take priority over segmental completion. In the post-test, learners more often attempted to realize these word endings, with fewer instances of complete deletion and more instances of audible coda production. Importantly, the change was not merely "adding a sound"; rather, learners appeared to treat word endings as more salient and communicatively consequential, which is pedagogically important given the grammatical and lexical functions carried by codas.

A second noticeable change involved increased self-monitoring and repair behaviors. In the post-test recordings, learners more frequently paused briefly, restarted a word, or adjusted articulatory timing after producing a weak or incomplete ending. Such behaviors are developmentally meaningful because they indicate emerging capacity to notice mismatches between the intended target and the produced form and to attempt repair. This pattern is compatible with a skill-learning interpretation of shadowing: repeated perception-production cycles, particularly when learners receive platform-based feedback and teacher cues, can strengthen attention to low-salience segments and support more controlled execution of coda articulation under time pressure.

Despite these improvements, several error types persisted. Codas that were articulatorily demanding or phonotactically marked remained less stable, particularly in items involving affricate codas (e.g., /tʃ/, /dʒ/) and, more consistently, complex final consonant clusters (e.g., /kt/, /kst/). Learners frequently simplified these clusters via partial deletion (retaining only one consonant) or complete cluster reduction, resulting in reduced phonological contrast. Qualitatively, these patterns appeared linked to a speed-accuracy trade-off: when learners prioritized keeping pace with the model (a core requirement of shadowing), cluster completion was especially vulnerable, suggesting that real-time processing demands can exceed learners' current articulatory automatization for multi-consonant codas.

RQ2: Learners' Perceptions of eJOY-Supported Shadowing

To address Research Question 2, which explored learners' evaluations of the feasibility and perceived benefits of eJOY-supported shadowing, descriptive statistics were calculated for the post-intervention questionnaire. Table 3 summarizes learners' perceptions of the eJOY-supported

shadowing intervention on a five-point Likert scale. Given the small sample size (N = 10) and the short instrument (7 items), internal consistency should be interpreted cautiously. Nonetheless, the scale demonstrated good exploratory reliability (Cronbach's α = .865), suggesting that the items showed a coherent pattern as a brief indicator of learners' overall perceptions in this action-research context.

Table 3. *Descriptive Statistics for Learners' Perceptions of eJOY-Supported Shadowing*

Item Descriptor	Mean	SD
Enjoyment (Q1)	4.50	.527
Motivation (Q4)	4.10	.568
Awareness of final consonants (Q2)	4.20	.422
Perceived improvement in accuracy (Q3)	4.00	.471
Increased confidence in pronunciation (Q5)	4.00	.667
Ease & convenience of shadowing with eJOY (Q6)	4.50	.527
Intention to continue using eJOY & shadowing (Q7)	4.50	.527

*n = 10

Across all seven items, responses were consistently positive, with mean ratings ranging from 4.00 to 4.50. This narrow, high range indicates broad endorsement of the intervention rather than a polarized or mixed reaction, and it also implies a possible ceiling tendency typical of small-sample classroom evaluations. To provide a more informative interpretation, the results are discussed thematically in relation to feasibility/acceptance, perceived learning impact, and motivational sustainability.

Feasibility and acceptability of the intervention

Learners rated the intervention as both enjoyable and practically manageable. Enjoyment received one of the highest mean scores (Q1: $\bar{x}$ = 4.50, SD = 0.53), suggesting that participants generally experienced shadowing on eJOY as engaging rather than tedious. Similarly, learners strongly agreed that eJOY made shadowing easy and convenient (Q6: $\bar{x}$ = 4.50, SD = 0.53). These parallel peaks are important because they jointly reflect two conditions that are often difficult to achieve in pronunciation work: learners not only liked the practice, but also perceived it as feasible to sustain, which is essential when practice occurs largely outside the classroom.

Perceived impact on final sounds

With respect to the study's focal feature, learners reported heightened phonological awareness of word-final consonants (Q2: $\bar{x}$ = 4.20, SD = 0.42). This relatively high mean coupled with a smaller SD suggests a fairly shared perception that the intervention drew attention to word endings-an instructional prerequisite for improving low-salience segments such as codas. Learners also agreed that their final consonant accuracy improved (Q3: $\bar{x}$ = 4.00, SD = 0.47). While this mean is slightly lower than enjoyment and convenience, it still indicates clear perceived progress, and the moderate dispersion suggests that perceived improvement may differ across individuals-plausibly reflecting differences in baseline error severity, practice time, or the difficulty of specific target sounds.

A related perception concerned confidence in pronunciation (Q5: $\bar{x} = 4.00$, SD = 0.67). Notably, this item showed the largest variability among the seven statements. This pattern is meaningful: affective outcomes such as confidence often vary more than cognitive outcomes such as "awareness," particularly among adult learners whose prior experiences may include long-standing pronunciation anxiety or self-efficacy differences. In other words, while most participants agreed that the intervention supported pronunciation confidence, the strength of this effect was less uniform than the perceived increase in awareness.

Motivation and intention for continued practice

Motivation was also rated positively (Q4: $\bar{x} = 4.10$, SD = 0.57), indicating that learners generally felt the platform helped sustain their effort to improve pronunciation. The strongest behavioral-orientation signal came from learners' intention to continue using eJOY and shadowing (Q7: $\bar{x} = 4.50$, SD = 0.53). In small-scale interventions, continuance intention is particularly informative because it reflects perceived value beyond a one-time classroom requirement. Here, high intention scores-together with high convenience and enjoyment-suggest that learners viewed the practice as realistic to maintain in their own routines

DISCUSSION

Interpretation of Findings

This action research provides preliminary evidence that an eJOY-supported shadowing routine can support Vietnamese adult EFL learners' production of English word-final consonants. The observed pre-post improvement aligns with previous studies showing that sustained shadowing enhances aspects of oral performance, such as fluency and comprehensibility (Foote & McDonough, 2017; Hamada, 2014). However, unlike earlier research that focused primarily on global speaking gains, the present study examined segmental development in word-final consonants, offering a more fine-grained account of pronunciation change.

The qualitative error analysis further clarifies how improvement occurred. Learners demonstrated more consistent attempts at producing codas, suggesting heightened perceptual salience and strengthened perception-production links. This finding is particularly meaningful in the Vietnamese EFL context, where final consonant omission is widely documented (Ha, 2005). Nonetheless, gains were selective: articulatorily complex or typologically marked codas remained unstable, indicating that shadowing alone may not fully stabilize difficult segments without additional explicit guidance.

The perception results complement these findings. Consistent with prior mobile-assisted shadowing research (Foote & McDonough, 2017), learners reported positive evaluations of feasibility and usefulness. By integrating segment-specific performance analysis with learner perception data within a classroom-based AR framework, the study contributes context-sensitive evidence regarding both the effectiveness and practical sustainability of app-supported shadowing.

Pedagogical Implications

Several practical implications emerge for instructors who want to address word-final consonant difficulty in adult EFL settings. First, shadowing can be positioned as a targeted pronunciation routine rather than a general speaking activity. Previous research suggests that structured shadowing enhances attention to phonological detail when learners are explicitly guided toward specific features (Foote & McDonough, 2017; Hamada, 2014). To maximize benefits for final sounds, teachers can explicitly frame shadowing tasks around "word-ending accountability," such as reminding learners to maintain codas carrying grammatical meaning and to attend to final consonant clarity during practice. Second, shadowing should complement - not replace - explicit pronunciation instruction. The persistence of difficult codas and cluster simplification indicates that some targets require brief articulatory explanation, controlled practice, and corrective feedback before learners can produce them reliably in time-pressured shadowing. A short "micro-teaching" step (e.g., articulator placement, minimal pairs, or cluster breakdown) can therefore be used to prepare learners for more fluent imitation.

Moreover, out-of-class practice is feasible when students are trained in a stable routine. Mobile-assisted language learning research highlights the importance of manageable task design and perceived usability for sustaining engagement beyond class time (Kessler, 2018). After learners understand the procedure, shadowing can be assigned as home practice using short, level-appropriate clips. To protect sustainability, teachers can set realistic weekly targets and encourage learners to use features such as replay/loop or slowed playback when needed, particularly for final clusters. At the same time, instructors should anticipate implementation constraints. Technical issues and possible changes to platform access may affect continuity. Teachers can mitigate this risk by preparing alternative sources of short authentic clips while keeping the same core routine and assessment criteria.

Limitations and Directions for Future Research

This study has several limitations that should guide interpretation. Most importantly, the sample size was small and drawn from a single instructional context, which limits generalizability. In addition, the intervention period was relatively short; while measurable gains were observed, longer follow-up work is needed to determine whether improvements in final sound production are maintained over time and transfer to spontaneous speech. Future research can extend these preliminary findings by (a) using larger samples across varied contexts, (b) incorporating longitudinal follow-up to examine durability and transfer, and (c) comparing structured shadowing against alternative pronunciation practice types. Given the limited empirical work on eJOY for feature-specific pronunciation development, further studies should also clarify how platform-supported practice interacts with targeted instruction for difficult codas and final clusters.

CONCLUSION

This action research investigated the extent to which eJOY-supported shadowing could improve Vietnamese adult EFL learners' production of English word-final consonants and how learners evaluated the feasibility and perceived benefits of this practice. Overall, the pre-post results suggest that a structured shadowing routine implemented through eJOY can contribute to measurable improvement in final sound production within a short intervention period. Complementing the score gains, the feature-level inspection of recordings indicated meaningful shifts in coda-related

performance: learners more consistently attempted to realize word endings and showed greater salience of final consonants in their speech, which is pedagogically important because these segments often carry grammatical and lexical information. At the same time, improvement was not uniform across all targets. Persistent difficulties remained for articulatorily demanding sounds and complex word-final clusters, suggesting that shadowing alone may not stabilize all final consonants for adult learners. For classroom application, the findings support positioning eJOY-supported shadowing as a practical out-of-class pronunciation routine that can increase practice intensity, while still being complemented by brief, explicit phonetic instruction and targeted feedback for difficult codas and clusters. Learners' questionnaire responses were consistently positive across items related to enjoyment, convenience, perceived improvement, awareness of final sounds, motivation, confidence, and intention to continue, indicating that the intervention was viewed as both acceptable and feasible for sustained practice. Because the sample size was small, reliability indices should be interpreted descriptively; nevertheless, the overall response trend provides supportive learner-voice evidence aligned with the observed performance gains. In sum, the present study offers preliminary, context-specific evidence that integrating eJOY with shadowing can support adult Vietnamese learners in developing clearer word endings. Future research with larger samples, longer follow-up periods, and comparative designs is needed to examine durability, transfer to spontaneous speech, and the conditions under which technology-supported shadowing is most effective for complex final consonant targets.

ACKNOWLEDGEMENTS

The author sincerely thanks all the participants who voluntarily took part in the survey. Their time, honesty, and cooperation were essential to the completion of this study. This research received no specific external funding.

REFERENCES

Baumfield, V., Hall, E. & Wall, K. (2013). *Action research in education* (2nd ed.). Thousand Oaks, CA: SAGE. https://doi.org/10.33524/cjar.v17i2.263.

Bovee, N., & Stewart, J. (2009). The utility of shadowing. In A.M. Stoke (Ed.), *JALT 2008 Conference Proceedings*. Tokyo: JALT.

Carr, P. (2019). *English phonetics and phonology: An introduction*. John Wiley & Sons.

Cohen, L., Manion, L., & Morrison, K. (2017). Action research. In *Research methods in education* (8th ed., pp. 440-456). Routledge. https://doi.org/10.4324/9781315456539.

Crystal, D. (2011). *A dictionary of linguistics and phonetics*. John Wiley & Sons.

Derwing, T. M., & Munro, M. J. (2015). *Pronunciation fundamentals: Evidence-based perspectives for L2 teaching and research* (Vol. 42). John Benjamins Publishing Company. https://doi.org/10.1075/lllt.42.

Dick, B. (1993). A beginner's guide to action research. *Arcs Newsletter, 1*(1), 5-9.

Evers, K., & Chen, S. (2022). Effects of an automatic speech recognition system with peer feedback on pronunciation instruction for adults. *Computer Assisted Language Learning, 35*(8), 1869-1889. https://doi.org/10.1080/09588221.2020.1839504.

Foote, J. A., & McDonough, K. (2017). Using shadowing with mobile technology to improve L2 pronunciation. *Journal of Second Language Pronunciation*, *3*(1), 34-56. https://doi.org/10.1075/jslp.3.1.02foo.

Gilakjani, A. P. (2016a). English pronunciation instruction: A literature review. *International Journal of Research in English Education, 1*(1), 1-6.

Gilakjani, A. P. (2016b). What factors influence the English pronunciation of EFL learners? *Modern Journal of Language Teaching Methods (MJLTM), 6*(2), 314-326.

Gilakjani, A., P. & Ahmadi, M. R. (2011). Why is pronunciation so difficult to learn. *English Language Teaching, 4*(3), 74-83.

Ha, C. T. (2005). Common pronunciation problems of Vietnamese learners of English. *VNU Journal of Foreign Studies*, *21*(1).

Hamada, Y. (2014). The effectiveness of pre-and post-shadowing in improving listening comprehension skills. *The Language Teacher, 38*(1), 3-10.

Hamada, Y. (2015). Monitoring strategy in shadowing: self-monitoring and pair-monitoring. *The Asian EFL Journal, Professional Teaching Articles, 81*, 4-25.

Hamada, Y. (2016a). Shadowing: Who benefits and how? Uncovering a booming EFL teaching technique for listening comprehension. *Language Teaching Research, 20*(1), 35-52.

Hamada, Y. (2016b). *Teaching EFL learners shadowing for listening: Developing learners' bottom-up skills*. Routledge.

Hamada, Y. (2018). Shadowing: What is it? How to use it. Where will it go? *RELC Journal*, *50*(3), 386–393. https://doi.org/10.1177/0033688218771380.

Harmon, D. I. (2014). Teaching fluency and rhythm to learners using reading aloud and shadowing. *Teaching the sound system of English: Temple University Japan Studies in Applied Linguistics, 90*, 33-43.

Hsieh, K. T., Dong, D. H., & Wang, L. Y. (2013). A preliminary study of applying shadowing technique to English intonation instruction. *Taiwan Journal of Linguistics*, 11(2).

Kemmis, S., McTaggart, R., & Nixon, R. (2014). *The action research planner: Doing critical participatory action research*. Springer. http://dx.doi.org/10.1007/978-981-4560-67-2.

Kessler, G. (2018). Technology and the future of language teaching. *Foreign Language Annals, 51*(1), 205-218. https://doi.org/10.1111/flan.12318.

Kobilova, N. R. (2022). Importance of pronunciation in English language communication. *Academic Research in Educational Sciences, 3*(6), 592-597.

Lambert, S. (1992). Shadowing. *Meta, 37*(2), 263-273.

Lewin, K. (1946). Action research and minority problems. *Journal of Social Issues, 2*(4), 34-46. https://doi.org/10.1111/j.1540-4560.1946.tb02295.x.

Lord, G. (2005). (How) can we teach foreign language pronunciation? On the effects of a Spanish phonetics course. *Hispania, 88*(3), 557. https://doi.org/10.2307/20063159.

Luu, T. H., & Do, T. H. (2023). The effects of shadowing technique on non-English majors' speaking skills at HPU2: An action research project. *Scientific Journal of Hanoi Metropolitan University*, *2*(69), 102-111.

Martinsen, R., Montgomery, C., & Willardson, V. (2017). The effectiveness of video-based shadowing and tracking pronunciation exercises for foreign language learners. *Foreign Language Annals*, *50*(4), 661-680. https://doi.org/10.1111/flan.12306.

McLeod, S., Margetson, K., Wang, C., Tran, V. H., Verdon, S., White, K., & Phạm, B. (2022). Speech acquisition within a 3-generation Vietnamese-English family: The influence of maturation and ambient phonology. *Clinical Linguistics & Phonetics, 36*(4-5), 470-493. https://doi.org/10.1080/02699206.2020.1862915.

Mertler, C. A. (2019). *Action research: Improving schools and empowering educators.* Sage Publications.

Ngo, N. B. (2005). The Vietnamese learning framework–Part One: Linguistics. *COTSEAL Journal*.

Nguyen, T. T. T. (2007). Difficulties for Vietnamese when pronouncing English final consonants. Retrieved from: https://www.diva-portal.org/smash/get/diva2:518290/FULLTEXT01.pdf.

Nguyen, T. H., Hwang, W. Y., Pham, X. L., & Ma, Z. H. (2018). User-oriented EFL speaking through application and exercise: Instant speech translation and shadowing in authentic context. *Journal of Educational Technology & Society*, *21*(4), 129–142. https://www.jstor.org/stable/10.2307/26511544.

Omar, H. M., & Umehara, M. (2010). Using a shadowing technique to improve English pronunciation deficient adult Japanese learners: An action research on expatriate Japanese adult learners. *The Journal of ASIA TEFL, 7*(2), 199-230.

Osburne, A.G. (1996). Final cluster reduction in English L2 speech: A case study of a Vietnamese speaker. *Applied Linguistics, 17*(2), 164-181. https://doi.org/10.1093/applin/17.2.164.

Ricard, E. (1986). Beyond fossilization: A course in strategies and techniques in pronunciation for advanced adult learners. *TESL Canada Journal, 3*, 243 -253. https://doi.org/10.18806/tesl.v3i0.1009.

Salim, A., Terasne, T., & Narasima, L. (2020). Enhancing the students' pronunciation using shadowing technique at senior high school students. *Journal of Languages and Language Teaching*, *8*(1), 20-28. https://doi.org/10.33394/jollt.v8i1.2212.

Saskiana, W. O., Fadhilawati, D., & Aini, M. R. (2024). Evaluating the Efficacy of the EJOY English Application to Escalate Eleventh-Grade Students' Speaking Proficiency. *JOSAR (Journal of Students Academic Research)*, *9*(2), 107-119. https://doi.org/10.35457/josar.v9i2.3879.

Schuberg, B., Masciarelli, K., & Nguyen, N. (2013). Influence of Vietnamese pronunciation on the production of English final consonant sounds. *Linguistics*, 51(1), 137-162.

Sewell, A. (2016). *English pronunciation models in a globalized world: Accent, acceptability and Hong Kong English*. Routledge.

Shiki, O., Mori, Y., Kadota, S., & Yoshida, S. (2010). Exploring differences between shadowing and repeating practices: An analysis of reproduction rate and types of reproduced words. *ARELE: Annual Review of English Language Education in Japan, 21*, 81-90. https://doi.org/10.20581/arele.21.0_81.

Szyszka, M. (2016). *Pronunciation learning strategies and language anxiety*. Springer International Publishing.

Tamai, K. (1992). The effect of "follow-up" on development of listening ability and the relationship between "follow-up" ability and listening ability. *STEP Bulletin*, *4*, 48- 62.

Tang, G. M. (2007). Cross-linguistic analysis of Vietnamese and English with implications for Vietnamese language acquisition and maintenance in the United States. *Journal of Southeast Asian American Education & Advancement, 2*(1), 1-33. https://doi.org/10.7771/2153-8999.1085.

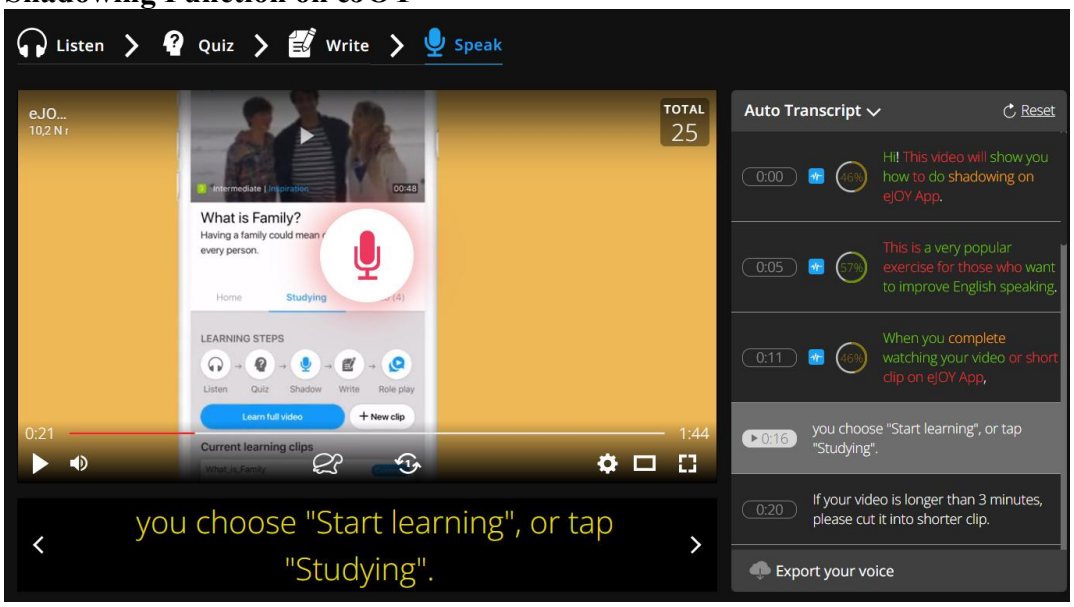
Younes, S. A., & Mueller Gathercole, V. C. (2022). Phonological interaction in Spanish-English bilinguals: effects of cognate usage on voice onset time of voiced stops. *International Journal of Bilingual Education and Bilingualism, 25*(2), 709-727. https://doi.org/10.1080/13670050.2020.1715914.

APPENDICES

Appendix A. eJOY Instruction

Platform (eJOY Go): https://ejoy-english.com/go/
Introduction Video: https://ejoy-english.com/go/video/how-to-do-shadowing-on-ejoy-app/61207

Shadowing Function on eJOY



How to do shadowing on eJOY (Speak mode):

When you complete watching the video on eJOY Go website, you choose "Speak".

- Step 1. You listen to the speaker. You can listen again or listen at a slower speed. If there are words you don't know how to pronounce, just tap on it to see the transcription.
- Step 2. Record your voice. Tap on the micro icon to record. Say it aloud. Make sure to allow eJOY to have access to your microphone. When done, tap on it again.
- Step 3. Check your result. Green words are those pronounced correctly. Red are those the app couldn't catch your words. Yellow are mostly correct. The white text is what the app can hear from you. (You can try it again if you think you can do it better.)

Appendix B. Pronunciation Assessment Rubric

Evaluation Criteria (0–20 each)	Poor (0–8)	Fair (9–12)	Good (13–16)	Excellent (17–20)

1. Accuracy of Final Sound Production	Final consonants are frequently omitted or substituted; many endings are incorrect.	Produces some final consonants correctly, but recurrent errors remain (omission/substitution /epenthesis), especially in connected speech.	Produces most final consonants correctly; occasional errors occur, typically with less familiar sounds or clusters.	Consistently accurate production of final consonants (incl. common grammatical endings -s/-es, -ed) with minimal errors.
2. Intelligibility of Word Endings (Clarity for listener)	Word endings are often unclear; meaning/grammar is frequently hard to infer.	Word endings are sometimes understandable but reductions cause occasional misunderstanding or ambiguity.	Word endings are mostly clear; minor reductions rarely affect meaning.	Word endings are clear and reliably support meaning/grammar; minimal listener effort required.
3. Self-correction / Monitoring of Final Sounds	Rarely notices or repairs final-sound errors; little evidence of self-monitoring.	Occasional self-correction, but inconsistent or only partly effective.	Frequently notices and repairs errors (repeats/adjusts endings) with generally successful outcomes.	Consistently monitors and repairs final-sound errors promptly and effectively; strong self-regulation.
4. Fluency with Final Sound Maintenance	Speech is frequently disrupted; final sounds are often dropped when trying to speak fluently.	Fluency is uneven; maintaining final sounds causes noticeable pauses or breakdowns.	Mostly smooth delivery; maintains final sounds with occasional hesitation (esp. clusters).	Smooth, natural flow while maintaining final sounds accurately, even under time pressure.
5. Confidence in Producing Word Endings	Strong hesitation/self-doubt; avoids, mumbles, or rushes endings.	Some hesitation is evident; confidence varies with speed/word difficulty.	Generally confident; minor hesitation does not strongly affect performance.	Confident, stable production of word endings; minimal hesitation and strong control.

Appendix C. Post-Intervention Questionnaire

This questionnaire elicits learners' perceptions of eJOY-supported shadowing practice for improving English pronunciation, with a focus on word-final consonants.

Please read each statement and select one response that best matches your opinion (✓).

Statements	Strongly disagree (1)	Disagree (2)	Neutral / Not sure (3)	Agree (4)	Strongly agree (5)
1. I think practicing pronunciation using eJOY with the shadowing technique is enjoyable.					

2. I think the shadowing technique has made me more aware of final consonant sounds when speaking English.					
3. I think my pronunciation of words ending in consonants has improved after shadowing practice.					
4. I think using eJOY keeps me motivated to improve my English pronunciation.					
5. I think I am more confident in pronouncing English words after using eJOY with shadowing.					
6. I think eJOY makes shadowing practice easier and more convenient.					
7. I think I will continue using eJOY and shadowing for future English pronunciation practice.					