International Journal of Social Science and Human Research

ISSN (print): 2644-0679, ISSN (online): 2644-0695

Volume 09 Issue 10 October 2026

DOI: 10.47191/ijsshr/v9-i10-70, Impact factor- 8.361

Page No: 9186- 9190

Making Word Stress Audible: Peer Feedback and Pronunciation Development in a Vietnamese University EFL Classroom

Nguyen Ngoc Anh
Thang Long University – Hanoi - Vietnam

ABSTRACT: This proposed study will examine how structured peer feedback contributes to the monitoring and revision of English word stress in a Vietnamese university EFL classroom. It will distinguish accurate stress placement from intelligibility and perceived comprehensibility. Approximately 40 intermediate learners will participate in six instructional sessions combining perception practice, individual recording, peer feedback, and renewed production. Baseline, immediate, and delayed assessments will include sentence reading and a short independent speaking task. Assessment words will include items practised during instruction and items withheld from instruction. Two trained assessors will judge primary stress placement, while independent English users will transcribe selected words and rate the ease of understanding short speech samples. Recorded peer exchanges and feedback sheets will support qualitative analysis of noticing, explanation, verification, and uptake. Descriptive summaries will be integrated with individual pronunciation trajectories, including unsuccessful corrections and uncertain cases. The study will investigate the conditions under which peer feedback supports justified self-monitoring rather than repeated imitation or unsupported correction. Because instruction combines several components and includes no comparison group, observed changes will not establish a causal effect of peer feedback. The proposal specifies a feasible collection and assessment framework without reporting results or assuming that pronunciation will improve.

KEYWORDS: pronunciation teaching; word stress; peer feedback; intelligibility; comprehensibility; university EFL

INTRODUCTION

A learner may repeat a word successfully immediately after a teacher and still fail to make it recognisable during a later utterance. Pronunciation instruction therefore faces a problem of interpretation: successful imitation can indicate temporary alignment without demonstrating independent control. Peer practice creates more opportunities to speak, but its educational value depends on what learners hear, identify, and explain to one another. A corrected repetition alone provides limited evidence of what has changed.

Word stress offers a focused way to investigate these processes. In multisyllabic words, the distribution of prominence contributes to how listeners recognise lexical items. Field (2005) examined lexical stress from the listener's perspective and showed that its relationship with intelligibility depends on the nature of the stress displacement and accompanying vowel changes. This supports a targeted investigation without implying that every stress difference causes misunderstanding or that stress alone determines intelligibility.

The proposed study will investigate structured peer feedback in one Vietnamese university EFL class. It will connect feedback episodes with recordings produced independently before, immediately after, and several weeks after instruction. Its goal is to explain how learners respond to feedback and whether changes are visible beyond a single corrected attempt. It will not treat a Vietnamese accent as a deficiency or assume that removing an accent is the purpose of teaching.

The project will distinguish instructional targets, evidence of learning processes, and listener outcomes. It will also examine incorrect feedback and unsuccessful uptake, rather than selecting only exchanges that appear effective. These choices are intended to make the study useful for teachers deciding how to organise peer pronunciation work and how to evaluate its limits.

LITERATURE REVIEW AND CONCEPTUAL FRAMEWORK

Intelligibility as a teaching objective

Levis (2005) contrasts orientations towards native-like pronunciation and intelligibility in pronunciation teaching. For this proposal, the instructional objective will be speech that listeners can understand with manageable effort. A reference pronunciation will support practice and assessment, but conformity to one speaker's voice will not define success. Accepted stress variants will be identified before scoring rather than retrospectively treated as errors.

Derwing and Munro (1997) distinguish accentedness, intelligibility, and comprehensibility through listener-based assessment. The present study will operationalise intelligibility as successful recovery of a target word, and comprehensibility as perceived ease of

understanding an utterance. Accent strength will not be rated. A speaker may produce the intended stress pattern while remaining difficult to understand for other reasons; conversely, contextual support may allow recognition despite an unexpected stress pattern.

Instructional outcomes and measurement

Saito and Plonsky (2019) distinguish pronunciation measures according to their construct, scoring procedure, and degree of production control. Their framework cautions against interpreting improvement on a controlled task as equivalent to general speaking development. Accordingly, this project will assess a specific feature in sentence reading and obtain separate listener judgements on independent speech. Neither task will stand for pronunciation ability as a whole.

The instructional focus will be primary lexical stress. Learners will practise identifying a prominent syllable, comparing productions with reference recordings, and attending to the contrast between stronger and weaker syllables. Duration, pitch, intensity, and vowel quality can contribute to perceived prominence; no single cue will be equated with stress. The study will not prescribe exaggerated loudness or assume that every unstressed vowel must become schwa.

Peer feedback and self monitoring

Martin and Sippel (2021) investigated providing and receiving pronunciation feedback among learners of German, making the feedback provider's activity relevant to pronunciation learning. Huang, Lee, and Ballinger (2023) examined peer feedback with learners of English and highlighted the importance of training. These studies motivate examining what learners do while evaluating peers, while differences in target language, setting, and delivery limit direct predictions for the proposed classroom.

The conceptual framework will connect noticing a possible mismatch, explaining it, checking it against an appropriate reference, and attempting a revision. These are proposed analytical categories rather than stages every learner must follow. Immediate uptake will mean an observable response to feedback; successful uptake will require a defensible revision of the target feature. Delayed independent performance will remain a separate outcome, because accepting feedback does not demonstrate retention.

Research aim and questions

The study will investigate the relationship between structured peer feedback, changes in primary word stress, and listener understanding. RQ1: What changes occur in stress placement, word recognition, and comprehensibility across baseline, immediate, and delayed assessment? RQ2: How do learners identify, explain, verify, accept, or reject pronunciation feedback during paired practice? RQ3: How do feedback episodes relate to immediate revisions and later independent production, including words withheld from instruction?

The questions are exploratory and do not presuppose improvement. The study will distinguish words practised during instruction from words assessed but not explicitly taught. It will examine associations and situated trajectories rather than estimate the separate causal contribution of peer feedback.

PROPOSED METHODOLOGY

Design and participants

A longitudinal classroom process study will combine descriptive performance measures with qualitative analysis of recorded feedback. Approximately 40 undergraduate EFL learners aged 18 or above will be invited from one course. Intermediate proficiency will be documented through an institutional placement record or available assessment, with the actual measure and score range reported. The recruitment target reflects the capacity to manage recordings and repeated assessments, not a statistical power calculation.

Approximately 20 stable pairs will participate in six focal sessions. Pair allocation will consider scheduling and willingness to work together, with the procedure documented. Both members will alternate between speaker and feedback provider. Attendance, pair changes, outside pronunciation practice, and withdrawals will be recorded. The convenience sample and single institution will limit generalisation beyond the participating context.

Instructional materials and training

Materials will include multisyllabic words, short sentences, reference audio, stress marking sheets, and a brief feedback checklist. Two pronunciation instructors will review lexical familiarity, syllable structure, stress variants, and the suitability of each reference. Words with unresolved disagreement about acceptable stress will be excluded. The teaching variety will be declared, while documented alternatives will remain acceptable in assessment.

The instructional bank will contain approximately 48 words distributed across six sessions, with planned recycling. The assessment list will contain eight words drawn from this bank and eight words withheld from explicit instruction. The sets will be reviewed for comparable familiarity and syllable complexity. Meanings will be checked without providing pronunciation models for the withheld items. The final lists will be fixed after piloting, before the main study begins.

An initial training workshop will model how to identify the syllable heard as prominent, express uncertainty, request a replay, and consult the reference before recommending a change. Students will practise with materials outside the assessment list. Feedback

will concern an audible feature rather than a judgement about the speaker. Vietnamese or English explanations will be permitted so that feedback quality is not restricted by English metalanguage.

Teaching sequence and timetable

Week 1 will contain baseline recording before the feedback workshop. Six instructional sessions will follow in Weeks 2–7, with approximately 25 minutes allocated to the focal sequence in each session. Immediate assessment will occur in Week 8; delayed assessment will occur in Week 10, approximately three weeks after the final instructional session. Exact intervals and deviations will be reported.

Each sequence will begin with short perception practice using nonassessment examples. Learners will then record a first attempt at eight session words in short sentences. Their partners will listen without initially seeing stress markings, identify uncertain productions, and explain their perception. The pair will consult the reference audio and marking guide before the speaker records another attempt. Roles will reverse, and each learner will complete a brief reflection about one feedback decision.

The teacher will monitor the procedure and address unresolved pronunciation questions after the paired exchange. Teacher interventions will be logged, including whether a proposed peer correction is confirmed or challenged. Reference access will be standardised across pairs. This arrangement combines perception training, repetition, reference comparison, peer interaction, and teacher support; any performance change cannot be attributed exclusively to one component.

Independent production assessments

At each assessment, learners will read 16 short sentences containing the target words and complete a 60-second description of an everyday situation. Sentence frames will provide limited semantic prediction of the target while remaining natural. Different but reviewed frames will be used across assessments, with order varied to reduce memorisation. Recording conditions, microphone distance, preparation time, and permitted repetitions will be standardised after piloting.

The same target words will recur across assessments. The withheld items are therefore untrained in the instructional sessions, not entirely unexposed. Test experience and independent practice may influence later production, and these possibilities will be recorded rather than described as pure transfer. Independent descriptions will use parallel prompts without requiring target vocabulary; they will provide broader speech samples but cannot measure spontaneous use of every target word.

With complete participation, the sentence task would yield 1,920 target-word tokens across three assessments, alongside 120 short descriptions. These figures are collection targets rather than results. Absent assessments, technical failures, and unusable clips will be reported separately. A flawed first take will not be replaced selectively because its pronunciation appears weak; repeat rules will concern recording quality or procedural interruption.

Assessment by trained raters and independent listeners

Two trained assessors will independently judge the location of primary stress in sentence recordings, using accepted variants specified in advance. Recordings will be coded and presented in random order without assessment dates. Unclear prominence will receive an uncertainty code rather than a forced decision. A third assessor will review disagreements after the independent ratings are retained for agreement analysis.

Approximately 12 proficient English users outside the class will be recruited as independent listeners. Their language backgrounds, hearing-related access needs, and familiarity with Vietnamese-accented English will be documented. They will not be classmates or teaching staff responsible for the speakers. Their responses will represent a defined listener sample, not universal English intelligibility.

For word recognition, listeners will hear a complete short sentence without its written script and transcribe what they hear. Target recovery will be scored separately from spelling accuracy when the intended word is identifiable. Each clip will receive two independent transcriptions through a balanced allocation plan. A listener will hear only one assessment version of a given speaker–word combination, limiting recognition from repeated exposure to the same item.

Comprehensibility will be rated on the independent descriptions using a nine-point scale from very difficult to very easy to understand. Each description will receive three ratings, with assignment balanced across assessment occasions and speakers. Listeners will practise with unrelated examples and receive instructions about ease of understanding rather than accent preference. Short blocks, breaks, and randomised order will limit fatigue and order effects.

Feedback records and stimulated recall

A separate device for each consenting pair will record the feedback exchange. Files will be linked to learner, pair, session, word, first attempt, and revised attempt identifiers. Feedback sheets will record the perceived prominent syllable, the suggestion made, verification source, and the speaker's decision. These records will allow an apparently successful repetition to be examined alongside the explanation that preceded it.

Approximately 10 learners will be invited for short stimulated-recall interviews within 48 hours of a selected session where practicable. Selection will seek varied trajectories, including rejection, uncertainty, and inaccurate correction. Following Gass and

Mackey (2017), replayed segments will prompt retrospective accounts rather than be treated as direct access to cognition. Actual recall intervals will be reported, and interviews will not presume that the feedback helped.

DATA ANALYSIS PLAN

Pronunciation and listener outcomes

Table 1 distinguishes the planned measures. Stress placement will be summarised as the proportion of scorable tokens matching an accepted pattern, separately for trained and withheld words. Unclear and missing tokens will be reported outside that denominator. Word recognition will be summarised from independent listener responses, preserving cases where the same clip receives different transcriptions.

Table 1. Planned pronunciation measures and scoring

Measure	Evidence source	Scoring and interpretation
Stress placement	Target words in sentence recordings	Accepted primary stress location versus another location. Unclear or missing tokens are reported separately. This is a feature-specific measure.
Word intelligibility	Independent sentence transcription	Recovery of the intended target word, scored separately for each listener. Identifiable spelling errors will not count as failed recognition.
Comprehensibility	Independent 60-second descriptions	Nine-point scale from 1 very difficult to 9 very easy to understand. Ratings concern the entire sample, not stress placement alone.
Feedback uptake	First and revised practice attempts	Retention, defensible revision, unsuccessful revision, or unresolved response, examined with the associated feedback episode.

Note. All measures are planned. No scores are reported. Accent strength will not be assessed, and no single measure will be treated as a complete index of pronunciation ability.

Comprehensibility distributions will be described by assessment occasion and speaker, with the scale direction stated. Initial interrater agreement will be reported before adjudication for stress judgements. An appropriate interrater reliability estimate and its assumptions will be specified for the listener ratings after piloting establishes the allocation structure. Neither agreement nor reliability will be described as already established.

Within-learner trajectories will be examined alongside class summaries and paired practice histories. Item, listener, speaker, and pair dependencies will remain visible; token counts will not be treated as independent participants. The primary plan is descriptive. Any later inferential extension will require a justified model and a documented change to the analysis plan before outcome-driven selection of tests.

Feedback episodes and integration

An episode will begin when a production is evaluated or questioned and end when the pair revises, retains, or leaves it unresolved. Coding will distinguish perceived mismatch, specific explanation, reference checking, alternative modelling, uncertainty, acceptance, rejection, and renewed production. Whether a correction is defensible will be judged against the reference and accepted variants, independently of how confidently it is delivered.

A second analyst will review approximately one quarter of episodes after calibration. Code definitions, boundaries, overlapping codes, and disagreements will be documented. Published extracts will preserve relevant language switching and uncertainty, with checked English translations where needed. No invented exchanges will substitute for unclear or missing recordings.

An integrated case matrix will connect feedback decisions, first and revised attempts, assessment stress ratings, word recognition, and recall accounts. Later changes will be described as associated trajectories, not proof that a particular comment caused learning. Negative cases will include accurate advice followed by an unsuccessful revision, inaccurate advice accepted, and immediate improvement that does not persist.

Feedback training will include an explicit distinction between a pronunciation difference and a communication problem. A partner who understands a word should not automatically label it incorrect because it differs from the reference voice. Equally, successful guessing from sentence meaning should not establish that the acoustic form was clear. Learners will be encouraged to report what they heard before proposing a correction, allowing the feedback record to preserve perception and evaluation as separate decisions. Listener preparation will include examples that differ in accent strength and ease of understanding, without suggesting that one accent is preferable. The researcher will check whether listeners understand the transcription and rating tasks before study recordings begin. Listening blocks and playback rules will be finalised during piloting. The same playback allowance will apply across

assessment occasions, and any replay will be logged so that later interpretation does not confuse immediate recognition with recognition after repeated exposure.

Disagreement between measures will be retained as evidence requiring explanation. A stress revision may accompany poorer word recognition if another sound changes, while comprehensibility may improve through slower, more organised delivery. The integrated analysis will examine these possibilities without assigning every change to stress placement. Where the recording cannot support a defensible interpretation, the report will retain an uncertainty judgement rather than impose a successful learning narrative.

Research quality and ethics

Materials and scoring rules will be piloted outside the main sample. Assessment order will be concealed where practicable, and raters will not receive class participation information. Calibration materials will be separate from study clips. Recording failures, exclusions, and rating disagreements will be retained in an audit trail. Explanations of unexpected outcomes will consider vocabulary knowledge, segmental changes, speech rate, and listener familiarity alongside stress.

Institutional review and classroom permission will precede recruitment. Participation and recording consent will be voluntary and separate from course assessment. Students who decline will receive equivalent teaching without inclusion of their data. An independent colleague will manage consent where feasible if the teacher is also the researcher. Withdrawal arrangements, secure storage, access restrictions, and retention periods will be specified before collection.

Voice recordings can remain identifiable even after names are removed. They will therefore not be described as fully anonymous or openly shared without separate permission. Public outputs will use coded transcripts and limited authorised examples. Independent listeners will also provide consent. No recruitment, ethical approval, interview, or assessment is claimed to have occurred in this proposal.

Anticipated contribution and limitations

The project is intended to clarify how peer feedback operates in a focused pronunciation sequence and how immediate revisions relate to independent performance. It may support practical guidance on training students to hear a specific feature, explain their uncertainty, and verify suggestions. Such implications will depend on the collected evidence; improvement, favourable attitudes, and successful transfer are not predetermined outcomes.

The single-class design cannot separate feedback from repetition, teacher support, or ordinary course learning. Trained assessors and independent listeners address different questions, and neither represents every communication setting. Repeated assessment can create familiarity, while short descriptions may contain few relevant stress opportunities. These limits will constrain interpretation and motivate later controlled research rather than be resolved through stronger causal wording.

CONCLUSION

The proposed study will investigate word stress instruction through linked evidence from peer interaction, repeated production, and independent listeners. Its central distinction is between an accepted correction, a defensible pronunciation revision, and speech that another person understands. By examining all three, including unsuccessful and uncertain trajectories, the project will provide a context-sensitive account of classroom pronunciation practice. Data collection and analysis will follow the stated ethical and methodological procedures before any empirical conclusions are drawn.

REFERENCES

1) Derwing, T. M., & Munro, M. J. (1997). Accent, intelligibility, and comprehensibility: Evidence from four L1s. Studies in Second Language Acquisition, 19(1), 1–16. https://doi.org/10.1017/S0272263197001010
2) Field, J. (2005). Intelligibility and the listener: The role of lexical stress. TESOL Quarterly, 39(3), 399–423. https://doi.org/10.2307/3588487
3) Gass, S. M., & Mackey, A. (2017). Stimulated recall methodology in applied linguistics and L2 research (2nd ed.). Routledge. https://doi.org/10.4324/9781315813349
4) Huang, Y., Lee, A. H., & Ballinger, S. (2023). The characteristics and effects of peer feedback on second language pronunciation. Journal of Second Language Pronunciation, 9(1), 47–70. https://doi.org/10.1075/jslp.22034.hua
5) Levis, J. M. (2005). Changing contexts and shifting paradigms in pronunciation teaching. TESOL Quarterly, 39(3), 369–377. https://doi.org/10.2307/3588485
6) Martin, I. A., & Sippel, L. (2021). Is giving better than receiving? The effects of peer and teacher feedback on L2 pronunciation skills. Journal of Second Language Pronunciation, 7(1), 62–88. https://doi.org/10.1075/jslp.20001.mar
7) Saito, K., & Plonsky, L. (2019). Effects of second language pronunciation teaching revisited: A proposed measurement framework and meta-analysis. Language Learning, 69(3), 652–708. https://doi.org/10.1111/lang.12345

There is an Open Access article, distributed under the term of the Creative Commons Attribution – Non-Commercial 4.0 Intenational (CC BY-NC 4.0) (https://creativecommons.org/licenses/by-nc/4.0/), which permits remixing, adapting and building upon the work for non-commercial use, provided the original work is properly cited.