MEETING STUDENTS WHERE THEY ARE: A COMMUNITY COLLEGE APPROACH TO ASYNCHRONOUS PRONUNCIATION INSTRUCTION THROUGH H5P-INTEGRATED SPEECH RECOGNITION

Patryk Mrozek, Portland Community College
Lara Mendicino, Portland Community College
Kate Carney, Portland Community College
Annie Karas, Portland Community College

This study reports on the development, implementation, and evaluation of an English pronunciation training tool integrated into Portland Community College's Learning Management System (LMS) for asynchronous homework practice. The tool uses H5P-based automatic speech recognition to support suprasegmental instruction for students in ESOL Communication courses. A quasi-experimental study with 12 cohorts ($N = 198$) compared control and experimental groups over a 10-week term across two tasks: read-aloud and spontaneous speech. Results showed no statistically significant between-group differences, although descriptive patterns suggest consistently greater gains for students using the tool, particularly at the intermediate level and on the read-aloud task. These findings warrant further controlled investigation.

INTRODUCTION

Institutional Context

Portland Community College (PCC) is the largest postsecondary institution in the state of Oregon (USA), serving a student population of approximately 53,000—6% of whom are enrolled in PCC's ESOL classes (Portland Community College, n.d.-a). The ESOL program spans eight levels of proficiency across numerous skills-based courses at four campuses and multiple associated centers. All programs at PCC participate in an ongoing Learning Assessment process. As part of this process, the ESOL program identified pronunciation as an area for program-wide improvement and selected it as the department's Learning Assessment focus beginning in 2022.

During the 2022–2023 academic year, the department conducted a large-scale diagnostic assessment to better understand the needs of its diverse student population. Instructors analyzed student recordings across levels, focusing on comprehensibility, defined here as "the ease or difficulty a listener experiences in understanding an utterance" (Derwing & Munro, 2015, p. 5), and recurring pronunciation errors. The analysis revealed diverse error types reflecting learners' varied L1 backgrounds. This variability underscored the need for an instructional approach that could support learners across many linguistic profiles (Gabriel & Thiele, 2017; Tronnier & Zetterholm, 2014).

Mrozek, P., Mendicino, L., Carney, K., & Karas, A. (2026). Meeting students where they are: A community college approach to asynchronous pronunciation instruction through H5P-integrated speech recognition. In P. Trofimovich, A. Bodea, T.-N. N. Le, R. Suzuki, & C. Teló (Eds.), *Proceedings of the 16th Pronunciation in Second Language Learning and Teaching Conference* (pp. 1–15). Concordia University, Montreal, Canada, October 2025. https://doi.org/10.31274/psllt.21366

Pronunciation Helper: Background and Motivation

Following the needs assessment, suprasegmental features of speech (such as lexical stress, phrasing, and intonation) emerged as the most actionable targets for improving learner outcomes, as they represented areas of difficulty shared across the entire sample. This focus was further prompted by evidence demonstrating the strong impact of suprasegmentals on comprehensibility and intelligibility (Derwing et al., 1998; Diaz, 2017; Gordon & Darcy, 2022)—the instructional priorities the department selected after reviewing relevant research (Chau & Huensch, 2025; Levis, 2018).

In 2023–2024, ESOL faculty members undertook the development of a tool to support suprasegmental pronunciation instruction with automatic speech recognition (ASR): The Pronunciation Helper. The goal was to create a resource accessible to all students enrolled in PCC's ESOL Communication courses, regardless of instructor, schedule, or campus. The asynchronous design of the tool followed from evidence of the legitimacy of homework-based pronunciation training (Martin, 2020; Stoughton & Kang, 2024) and addressed two longstanding challenges: limited instructional time for pronunciation and substantial variation in instructors' preparedness to teach pronunciation effectively (Baker, 2014; Foote et al., 2011; Gabriel & Thiele, 2017).

The tool was integrated into PCC's learning management system (Brightspace D2L), enabling streamlined access for students and allowing instructors to monitor engagement and completion. Its design relied on widely available technologies used across the college: H5P for interactive activities, Kaltura for video hosting, and Google Slides for presentation materials. The final product took the form of a dedicated D2L shell composed of eight instructional modules (see Figure 1), each targeting suprasegmental features relevant to intelligibility (Guskaroska et al., 2024; Levis, 2018). The order of the modules was planned to enable early units to scaffold later learning and to ensure logical continuity across topics.

Figure 1

Pronunciation Helper Module Topics

Module 1: Word Stress in Two-Syllable Words
Module 2: Word Stress in Longer Words with Prefixes and Suffixes
Module 3: Compound Nouns
Module 4: Sentence Stress–Content and Function Words
Module 5: Thought Groups and Focus Words
Module 6: Linking Words Together
Module 7: Reductions
Module 8: Review

H5P and ASR Integration

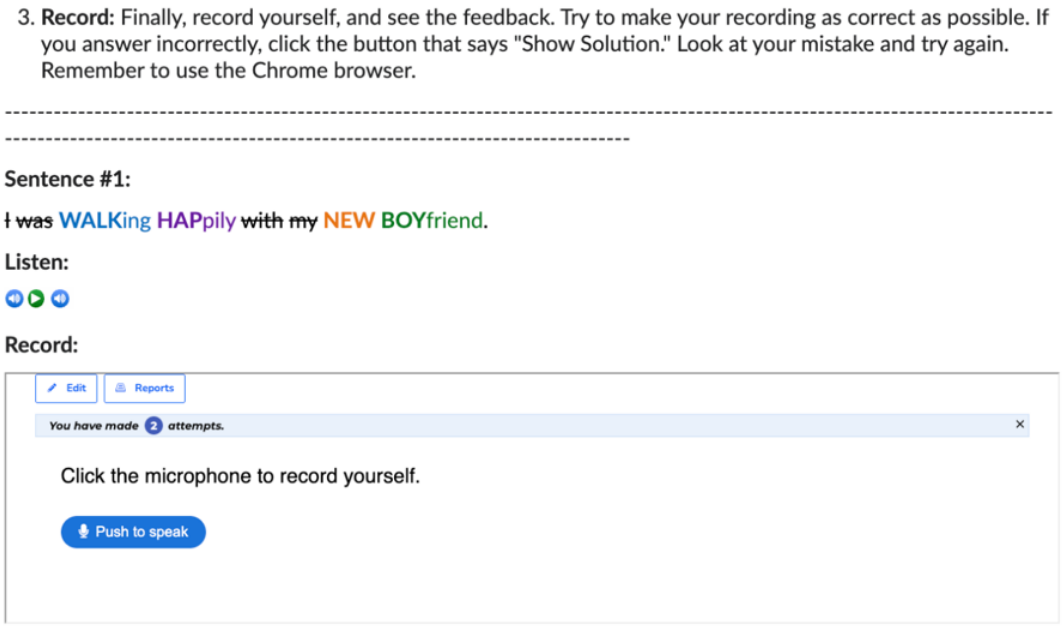
H5P is an open-source platform for creating interactive learning content widely used in LMS environments (Sbardella, 2023). Research has consistently documented high levels of student satisfaction and engagement with H5P-based activities (Doran & Briggs, 2024; Wicaksono et al., 2021). A key feature leveraged by PCC's Pronunciation Tool is H5P's "Speak the Words"

content type, which incorporates Google's ASR dictation subtype. This type of ASR transcribes user speech into text (Guskaroska, 2024), which enables learners to receive immediate feedback as a "match" or "mismatch" between their production and the target. Research suggests that although dictation-based ASR tools vary in accuracy and lack phonetic detail, they support practice, promote learner autonomy, and can improve comprehensibility (Evers & Chen, 2022; Liakin et al., 2014; Liu et al., 2019; McCrocklin, 2016, 2019; Mroz, 2018; Sun, 2023).

Tool Design

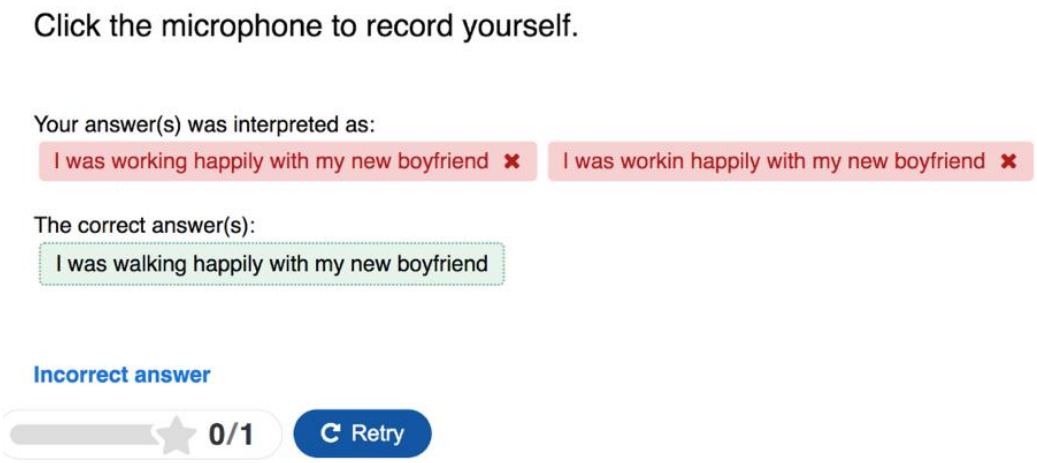
Each Pronunciation Helper module includes two core components. The Learn section provides instructor-created video lectures, fully captioned, paired with a deck of presentation slides explaining English prosody with examples. The Practice section offers students ten sentences for imitation practice. Sentences have been annotated to draw attention to relevant suprasegmental cues using different colors, capitalization, and punctuation as shown in Figure 2, and their vocabulary has been selected out of the 1,000 most frequent word families from the Corpus of Contemporary American English (COCA). Each sentence is accompanied by three model recordings from both native and non-native speakers, reflecting principles of high-variability phonetic training—an approach which argues that exposure to target sounds from multiple speakers in varied contexts enhances learning, promotes generalization, and leads to lasting gains in pronunciation and perception (Thomson, 2018).

Figure 2
Example Sentence-Recording Practice Activity



Students practice by recording their speech using H5P's "Speak the Words" activity. The built-in ASR dictation capability provides immediate feedback by displaying either a "Correct" notification or the ASR's interpretation of what was heard, offering learners an opportunity to compare, adjust, and repeat their production (see Figure 3).

Figure 3
Example of "Speak the Words" Feedback



Pre-Study Pilot: Student and Instructor Feedback

For three terms, ESOL Communication instructors piloted the tool both in class and as homework. Feedback was gathered from students and faculty with Google Forms that included both multiple-choice and open-ended questions about the Pronunciation Helper's perceived value and ease of use, which informed ongoing refinements. Survey data from 72 students indicated overall positive perceptions of the Pronunciation Helper. Most students found both the video lessons (77.8%) and pronunciation exercises (83.3%) useful, and 75% reported that they genuinely enjoyed completing the activities. Engagement was high, with multiple attempts reported. Open-ended comments reflected strong appreciation for the opportunity to practice pronunciation independently, with many students requesting continued access to the tool after the term ended. However, some students noted frustration when the ASR system did not accurately recognize their speech, and a smaller number reported occasional technical issues or difficulty understanding how their work was graded.

Instructor feedback echoed these themes. Teachers observed that students were enthusiastic about the modules and that the videos effectively clarified key concepts. At the same time, instructors noted that ASR feedback occasionally required interpretation that not all students were qualified to make, and they highlighted challenges with using the tool by lower-level learners—particularly related to linguistic complexity of the sample practice sentences and digital literacy requirements. This feedback motivated the research team's decision to explore the relationship between learners' proficiency and gains in this study.

Research Questions

To evaluate learner outcomes and satisfy the requirements of PCC's Learning Assessment process, a one-term, 10-week quasi-experimental study was conducted. Three research questions guided the investigation:

1. Do students who use the Pronunciation Helper tool show greater improvement on the measures of comprehensibility and accentedness than students who do not use the tool?
2. To what extent are learners' proficiency levels associated with gains in comprehensibility and accentedness?
3. Are gains in comprehensibility and accentedness consistent across read-aloud tasks and spontaneous speech?

Accentedness is defined as how closely the pronunciation of an utterance approximates that of a native speaker (Derwing & Munro, 1997). While not a teaching priority at the college, accentedness was of interest to the research team for institutional reasons: to illustrate to the faculty and students that comprehensibility and accentedness are partially independent concepts, where speech can be strongly accented but still easy to understand (Levis, 2018).

METHOD

Participants

Participants were students (n = 198) enrolled in twelve sections of PCC's ESOL Communication courses. Students were enrolled in Levels 5, 6, and 7 of the ESOL sequence, corresponding approximately to CEFR B1 (Level 5), B2 (Level 6) and C1 (Level 7). The participating sections were assigned to either a control group (six cohorts) or an experimental group (six cohorts) based on each participating cohort instructor's choice. Table 1 illustrates exact participant numbers across treatment groups, cohorts, and levels.

Table 1
Study Participants Across Treatment Groups, Levels, and Cohorts

Group	**Control (n = 101)**						**Experimental (n = 97)**					
Level	5	5	6	6	6	7	5	6	6	6	7	7
Cohort	1	2	3	4	5	6	7	8	9	10	11	12
No. of students	16	25	12	9	19	20	13	19	20	11	11	23

Note. The total numbers of students at each proficiency level are as follows: Level 5 = 54, Level 6 = 90, Level 7 = 54.

Additionally, 81 naive raters were recruited from PCC's Communication Studies program. These students were enrolled in three undergraduate-level Public Speaking courses, taught by the same instructor. Raters were predominantly L1 English speakers, although exact demographic information (L1s, ages, genders) was not collected to comply with college policies and FERPA regulations.

Procedure

The study followed a 10-week quasi-experimental design aligned with PCC's academic quarter. All cohorts received regular in-class pronunciation instruction consistent with PCC's Course

Content and Outcome Guidelines (CCOGs), which emphasize segmental features oriented toward the goal of producing "mostly understandable English in an academic setting" (Portland Community College, n.d.-b).

The control and experimental groups differed in their homework assignments. Control cohorts completed pronunciation homework at the discretion of the instructor, using materials and activities of their choosing. Typical homework included segmental-focused activities (e.g., phoneme identification, minimal pairs, and recorded practice). In contrast, experimental cohorts were assigned the Pronunciation Helper as weekly asynchronous homework. Thanks to smooth integration of the H5P exercises within the LMS, instructors were able to track students' progress to ensure that all students completed the eight modules of the tool.

Data Collection

Students completed a pretest in the first week of the term and a posttest during the final week by submitting two types of speech samples on each occasion (for prompts, see Appendix A):

1. Read-aloud task: Students read ten sentences selected directly from the Pronunciation Helper practice activities.
2. Spontaneous speech task: Students responded orally to a short prompt. The prompts were chosen to ensure they would not be overly demanding for any of the three levels and to systematically vary cognitive, linguistic, and discourse demands while remaining personally meaningful and culturally inclusive.

Recordings were collected anonymously through JotForm, an online survey platform. Raters evaluated each recording for comprehensibility and accentedness using 9-point Likert scales adapted from Derwing and Munro (1997), where a rating of 1 signified generally low comprehensibility and high accentedness, and a rating of 9 pointed towards high comprehensibility and a native-like accent (see Appendix B). Raters received limited training in the form of a 30-minute, pre-recorded instructional video and provided their responses via anonymous Google Forms.

Data Overview

A total of 296 unique speech samples were evaluated: 161 read-aloud and 135 spontaneous speech. Table 2 illustrates sample distribution across proficiency levels.

Table 2
Speech Samples by Group and Level

Group	Level 5	Level 6	Level 7	Total
Control	39	52	37	128
Experimental	48	65	55	168

The raters were assigned samples in a way that each sample would be rated by three raters. However, after removing incomplete, mislabeled, or unidentifiable responses, it transpired that

each recording received between one and three ratings. In total, the dataset consisted of 671 usable ratings. 163 pretest ratings and 133 posttest ratings were collected from the control group, while the experimental group was represented with 205 pretest ratings and 170 posttest ratings. The resulting dataset was unbalanced due in part to student anonymity and variable submission patterns; the raters submitted their ratings both in class and as homework, with some individuals failing to complete the task.

Data Analysis

All analyses were conducted in the R statistical environment using linear mixed-effects models. A series of models were fit to examine changes in comprehensibility and accentedness. For each measure, these included (a) a model with time as the sole predictor, (b) a group × time model to assess differential gains, and (c) two separate models for level and task as moderators of the group × time effect. In each model, random intercepts were included for cohort, in order to account for instructor and class-level variability, as well as recording ID, to account for multiple ratings of the same speech sample. To address the unbalanced structure of the dataset, estimated marginal means (EMMs) were obtained for all key comparisons. Importantly, EMMs were used to conduct planned contrasts aligned with the research questions, rather than for post hoc exploratory comparisons.

RESULTS

Effects of the Pronunciation Helper on Comprehensibility and Accentedness

Table 3 shows raw means and standard deviations for comprehensibility. At pretest, the control group showed slightly higher comprehensibility scores than the experimental group; however, this difference was not statistically significant ($p > .05$) and corresponded to a small effect size ($d = 0.21$). Unless otherwise noted, effect sizes (d) are reported as post − pre, such that positive values indicate improvement.

Table 4 shows model-adjusted gains for comprehensibility. In the experimental group, comprehensibility increased from 5.42 to 6.00, a gain of 0.58 points; this within-group change was statistically significant (post–pre = 0.58, $t(264) = 2.81$, $p = .0053$) and showed a small-to-moderate effect size ($d = 0.34$, 95% CI [0.08, 0.51]).

Table 3
Raw Means and Standard Deviations for Comprehensibility by Group and Time

Group	Time	*M*	*SD*
Control	Pre	5.75	1.81
Control	Post	5.94	1.80
Experimental	Pre	5.38	1.98
Experimental	Post	5.98	1.81

Table 4

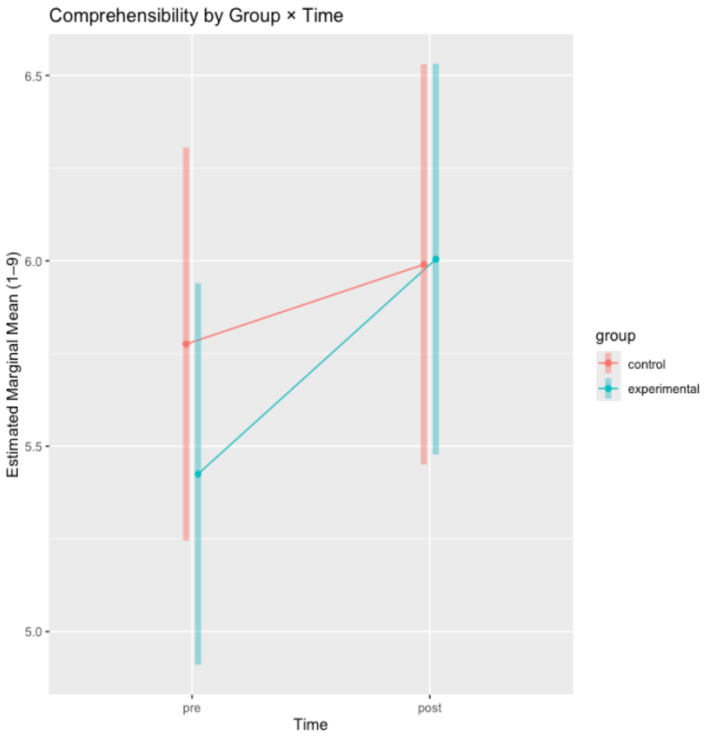
Estimated Gains in Comprehensibility by Group

Group	M_{diff}	*SE*	df	*t*	*p*	*d*	95% CI
Control	+0.215	0.233	248	0.920	.358	0.13	[–0.17, 0.42]
Experimental	+0.580	0.206	264	2.809	.005**	0.34	[0.08, 0.51]

While the pattern of estimated marginal means suggested a modest advantage for the experimental group (illustrated in Figure 4), these descriptive patterns should be interpreted with caution and cannot be taken as evidence of a reliable differential effect of the intervention. A key finding of this study is that the group × time interaction for comprehensibility was not statistically significant, $t(271.11) = 1.17$, $p = .243$. The standardized effect size for the interaction was small ($d = 0.22$, 95% CI [−0.15, 0.58]).

Figure 4

Comparison of Comprehensibility Gains for Control and Experimental Groups



Note. Points represent model-estimated marginal means from the linear mixed-effects model. Error bars indicate 95% confidence intervals.

For the measure of accentedness, the model showed no significant interaction between group and time, $t(255.53) = 0.70$, $p = .484$, corresponding to a very small effect size ($d = 0.12$, 95% CI [−0.22, 0.46]). Descriptive patterns indicated minimal change in accentedness across both groups, with no consistent evidence of improvement over time.

Effects by Proficiency Level

In terms of level-specific gains, estimated marginal means showed numerically largest improvement in comprehensibility for experimental students at Level 5 (illustrated by Table 5 and Figure 5). However, the three-way interaction (group × time × level) for the measure of comprehensibility was not statistically significant (all *p*s ≥ .48), with small, standardized effect sizes (*d*s ≈ −0.31 to −0.34).

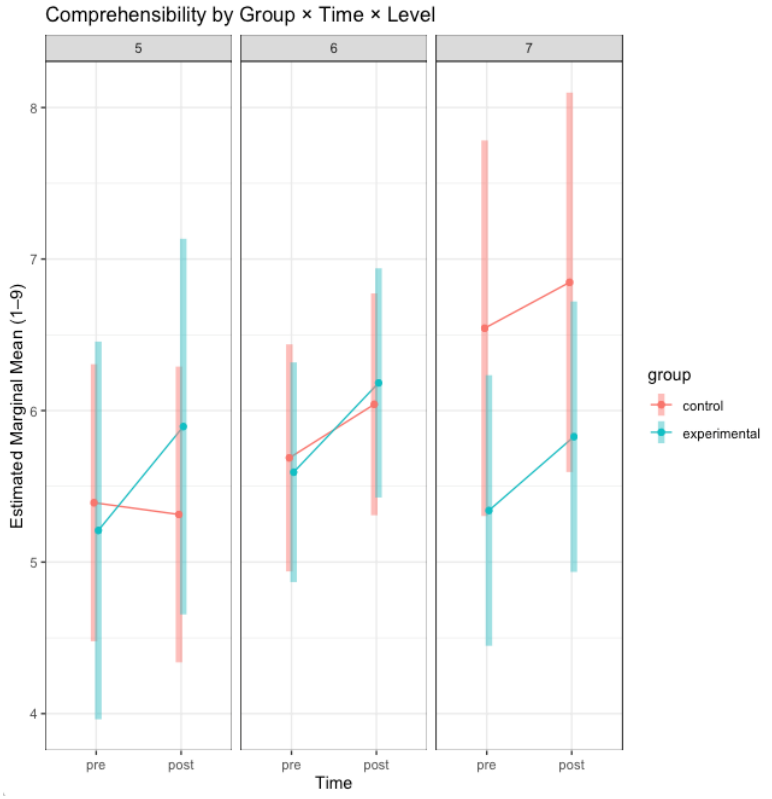
Table 5

Estimated Gains in Comprehensibility Across Levels of Proficiency

Level	Group	Pre	Post	Gain	*p*	*d*	95% CI
5	Control	5.39	5.31	−0.08	.86	−0.05	[−0.62, 0.53]
	Experimental	5.21	5.89	+0.69	.08	0.41	[−0.19, 1.00]
6	Control	5.69	6.04	+0.35	.32	0.21	[−0.24, 0.66]
	Experimental	5.59	6.18	+0.59	.08	0.35	[−0.09, 0.79]
7	Control	6.54	6.85	+0.30	.52	0.18	[−0.50, 0.86]
	Experimental	5.34	5.83	+0.49	.19	0.29	[−0.22, 0.79]

Figure 5

Estimated Gains in Comprehensibility Across Levels of Proficiency



For the measure of accentedness, the three-way interaction was not statistically significant either, $t(242.16) = -1.33$, $p = .19$, with small-to-moderate but non-significant estimates across levels (ds ≈ -0.23 to -0.55).

Effects of Task Type

The three-way interaction (group × time × task) was not statistically significant, $t(265.19) = -1.31$, $p = .19$. However, estimated marginal means were examined to explore potential patterns in the unbalanced data. The EMM-adjusted effect size of the three-way interaction was moderate ($d = 0.49$, 95% CI [−0.25, 1.23]), indicating a tendency for larger gains in the read-aloud task compared to spontaneous speech; this pattern was also visible in within-group gains (see Table 6).

Table 6
Comparison of Gains in Comprehensibility Across Tasks (EMM-adjusted)

Task	Group	Pre	Post	Gain	*p*	*d*	95% CI
Read-aloud	Control	5.76	6.01	+0.25	0.45	0.15	[−0.25, 0.55]
	Experimental	5.21	6.18	+0.97	0.0011	0.57	[0.21, 0.93]
Spontaneous speech	Control	5.80	6.09	+0.29	0.39	0.17	[−0.23, 0.58]
	Experimental	5.66	5.84	+0.18	0.54	0.11	[−0.26, 0.48]

Model analysis of accentedness revealed no significant main effects or interactions involving task, including the group × time × task interaction, $t(256.37) = -0.16$, $p = .874$, with a negligible effect size ($d = 0.06$).

DISCUSSION

Addressing the first research question, the ASR-integrated Pronunciation Helper did not lead to statistically significant between-group differences in comprehensibility or accentedness. However, estimated marginal means showed consistently larger within-level gains in the experimental group, aligning with prior ASR research (Evers & Chen, 2022; Liakin et al., 2014; Liu et al., 2019; McCrocklin, 2016, 2019; Mroz, 2018; Sun, 2023). These patterns might also reinforce prior findings that suprasegmental-focused instruction can positively influence comprehensibility (Diaz, 2017; Gordon & Darcy, 2022).

Addressing the second research question, although not statistically significant, exploratory patterns indicated somewhat larger within-level gains observed at Level 5. One possible explanation is that learners at the intermediate level may be optimally positioned to benefit from suprasegmental-focused instruction, as they have sufficient linguistic proficiency to attend to prosodic features while still benefiting from explicit scaffolding. However, given the lack of statistical significance, this interpretation remains tentative.

Regarding the final research question, which asked whether gains were consistent across task types, the group × time × task interaction was not significant, but the within-group gains pattern

suggests that the Pronunciation Helper had its clearest impact on the sentence-reading task. These findings would corroborate core research (Derwing et al., 1998; Saito & Lyster, 2012) showing that pronunciation instruction improves accuracy in controlled contexts but does not guarantee spontaneous-use gains.

Finally, looking at the measure of accentedness across all three RQs, the lack of meaningful gains across groups and conditions validates findings from multiple earlier studies and meta-analyses (Gordon & Darcy, 2022; Lee et al., 2015; Munro & Derwing, 1999), which suggest that accentedness outcomes are the least responsive to instruction relative to intelligibility and comprehensibility.

Study Limitations

Several design features limit the strength and generalizability of the findings. Instructor variability was built into the project, and although cohort was included as a random factor, differences in teachers' comfort and emphasis on pronunciation instruction could not be fully separated from the tool's effects. Institutional constraints required anonymous student recordings, preventing individual pre–post comparisons and allowing only group-level trends. Remaining constraints stemmed from anonymous rater submissions. The absence of inter-rater reliability estimates further limits the interpretability of the ratings, as variability in rater judgments may have introduced additional noise into the data.

Implications and Future Directions

Although the results did not yield statistically significant between-group differences, the consistent descriptive patterns favoring the experimental group suggest a potentially meaningful effect that may have been obscured by design constraints. As such, these findings provide a strong rationale for replication under more controlled conditions, where the sources of variability outlined above can be minimized. At the institutional level, the project demonstrates a replicable, cost-effective model for community colleges seeking to strengthen pronunciation support. Because the tool relies on widely available technologies, it can be implemented with minimal financial investment.

Looking ahead, several improvements and research directions emerge. Future research should aim to replicate the present study under more controlled conditions, including balanced datasets, repeated-measures designs tracking individual learners over time, and more consistent in-class pronunciation instruction across cohorts. Future versions of the Pronunciation Helper should explore ASR systems capable of providing explicit feedback on suprasegmental features of speech, as opposed to ASR dictation models. Additionally, future studies should ensure that assessment materials are independent from instructional content. In the present study, the inclusion of sentences from the Pronunciation Helper in the pre- and post-tests likely introduced practice effects, particularly in the read-aloud task, and may have inflated observed gains for sentences drawn directly from the instructional materials. Using novel test items would help better isolate the effects of the intervention.

ACKNOWLEDGMENTS

The authors thank the instructors in the ESOL and Communication Studies departments at Portland Community College for their collaboration and the many students who participated. We also wish to thank Portland Community College's Innovation Hub for awarding us the Techquity grant, which provided the essential support and resources needed to develop the tool.

ABOUT THE AUTHORS

Patryk Mrozek, M.A. TESOL (patmrozek@gmail.com), is a part-time ESOL instructor at Portland Community College and Co-Coordinator for CATESOL's Teachers of Pronunciation (TOP) interest group. His research interests include technology-enhanced pronunciation instruction and the teaching of suprasegmental features of speech.

Lara Mendicino, M.A. TESOL (lara.mendicino@pcc.edu), is a full-time ESOL instructor and current Faculty Department Chair at Portland Community College. She is interested in the relationship between oral reading and silent reading fluency, the connection between prosody and reading comprehension, and oral reading strategies like shadowing and echoing.

Kate Carney (therealkatecarney@gmail.com) worked as an ESOL instructor for 34 years. She retired from PCC in 2024.

Annie Karas, M.A. TESOL (annie.karas@gmail.com), is a full-time ESOL instructor at Portland Community College. She is interested in supporting low-level learners who were introduced to education later on in life. She also enjoys helping students find the connection between their studies in ESOL and their future careers.

REFERENCES

Baker, A. (2014). Exploring teachers' knowledge of second language pronunciation techniques: Teacher cognitions, observed classroom practices, and student perceptions. *TESOL Quarterly*, *48*(1), 136–163. https://doi.org/10.1002/tesq.99

Chau, T., & Huensch, A. (2025). The relationships among L2 fluency, intelligibility, comprehensibility, and accentedness: A meta-analysis. *Studies in Second Language Acquisition*, *47*(1), 282–307. https://doi:10.1017/S0272263125000014

Davies, M. (n.d.). *Corpus of Contemporary American English (COCA)*. Retrieved October 20, 2025, from https://www.english-corpora.org/coca

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://www.jstor.org/stable/44488664

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

Derwing, T. M., Munro, M. J., & Wiebe, G. (1998). Evidence in favor of a broad framework for pronunciation instruction. *Language Learning*, *48*, 393–410. https://doi.org/10.1111/0023-8333.00047

Diaz, A. (2017). Towards improving Japanese EFL learners' pronunciation: The impact of teaching suprasegmentals on intelligibility. *APU Journal of Language Research*, *2*, 69–84. https://doi.org/10.34409/apujlr.2.0_69

Doran, J., & Briggs, N. (2024). H5P interactive materials for flipped learning: Engagement, effectiveness and challenges. In *Proceedings of the Thailand TESOL International Conference 2024: Reconnecting ELT Professionals for Glocal Sustainability*. Thailand TESOL. https://thailandtesol.org/upload/pdf/1728270532.pdf

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*, 1869–1889. https://doi.org/10.1080/09588221.2020.1839504

Foote, J. A., Holtby, A. K., & Derwing, T. M. (2012). Survey of the Teaching of Pronunciation in Adult ESL Programs in Canada, 2010. *TESL Canada Journal*, *29*(1), 1–22. https://doi.org/10.18806/tesl.v29i1.1086

Gabriel, C., & Thiele, S. (2017). Learning and teaching of foreign language pronunciation in multilingual settings: A questionnaire study with teachers of English, French, Italian and Spanish. In M. K. P. Meyer, & A. Deppermann (Eds.), *Migration, Mehrsprachigkeit und Inklusion* (pp. 267–294). Universität Liechtenstein Press.

Gordon, J., & Darcy, I. (2022). Teaching segmentals and suprasegmentals: Effects of explicit pronunciation instruction on comprehensibility, fluency, and accentedness. *Journal of Second Language Pronunciation*, *8*(2), 168–195. https://doi.org/10.1075/jslp.21042.gor

Guskaroska, A. (2024). Exploring technology acceptance of automatic speech recognition (ASR) for pronunciation practice [Master's thesis, Iowa State University]. https://dr.lib.iastate.edu/server/api/core/bitstreams/027e0329-8f63-46b4-b431-9da7876ff76c/content

Guskaroska, A., Zawadzki, Z., Levis, J., Challis, K., & Prikazchikov, M. (2024). *Teaching pronunciation with confidence: A resource for ESL/EFL teachers and learners*. Iowa State University Digital Press. https://doi.org/10.31274/isudp.2024.161

Lee, J., Jang, J., & Plonsky, L. (2015). The effectiveness of second language pronunciation instruction: A meta-analysis. *Applied Linguistics*, *36*(3), 345-366. https://doi.org/10.1093/applin/amu040

Levis, J. M. (2018). *Intelligibility, oral communication, and the teaching of pronunciation*. Cambridge University Press.

Liakin, D., Cardoso, W., & Liakina, N. (2014). Learning L2 pronunciation with a mobile speech recognizer: French /y/. *CALICO Journal*, *32*(1), 1–25. https://doi:10.1558/cj.v32i1.25962

Liu, X., Xu, M., Li, M., Han, M., Chen, Z., Mo, Y., Chen, X., & Liu, M. (2019). Improving English pronunciation via automatic speech recognition technology. *International Journal of Innovation and Learning*, *25*(2), 126–140. https://doi.org/10.1504/IJIL.2019.097674

Martin, I. A. (2020). Pronunciation can be acquired outside the classroom: Design and assessment of homework-based training. *Modern Language Journal*, *104*(2), 457–479. https://doi.org/10.1111/modl.12638

McCrocklin, S. M. (2016). Pronunciation learner autonomy: The potential of automatic speech recognition. *System*, *57*, 25–42. https://doi.org/10.1016/j.system.2015.12.013

McCrocklin, S. M. (2019). ASR-based dictation practice for second language pronunciation improvement. *Journal of Second Language Pronunciation*, *5*, 98–118. https://doi.org/10.1075/jslp.16034.mcc

Mroz, A. P. (2018). *Noticing gaps in intelligibility through Automatic Speech Recognition (ASR): Impact on accuracy and proficiency.* Paper presented at 2018 Computer-Assisted Language Instruction Consortium (CALICO) Conference, Urbana, IL, United States.

Munro, M. J., & Derwing, T. M. (1999). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, *49*(s1), 285–310. https://doi.org/10.1111/0023-8333.49.s1.8

Portland Community College. (n.d.-a). *Demographics*. Portland Community College. Retrieved October 30, 2025, from https://www.pcc.edu/about/demographics

Portland Community College. (n.d.-b). *Course content and outcome guides*. Portland Community College. Retrieved October 30, 2025, from https://www.pcc.edu/ccog

Saito, K., & Lyster, R. (2012). Effects of form-focused instruction and corrective feedback on L2 pronunciation development of /ɹ/ by Japanese learners. *Language Learning*, *62*, 595–633. https://doi.org/10.1111/j.1467-9922.2011.00639.x

Sbardella, T. (2023). Enhancing language learning with H5P: Examples, challenges and suggestions. In *ICERI Proceedings* (pp. 7454–7474).

Stoughton, A. M., & Kang, O. (2024). A systematic review of empirical mobile-assisted pronunciation studies through a perception–production lens. *Languages*, *9*(7), Article 251. https://doi.org/10.3390/languages9070251

Sun, W. (2023). The impact of automatic speech recognition technology on second language pronunciation and speaking skills of EFL learners: A mixed methods investigation. *Frontiers in Psychology*, *14*, 121–187. https://doi.org/10.3389/fpsyg.2023.1210187

Thomson, R. I. (2018). High variability [pronunciation] training (HVPT): A proven technique about which every language teacher and learner ought to know. *Journal of Second Language Pronunciation*, *4*(2), 208–231. https://doi.org/10.1075/jslp.17038.tho

Tronnier, M., & Zetterholm, E. (2014). *Pronunciation difficulties in L2 and the multilingual classroom*. 16–17. Abstract from Multilinguality in Speech Research: Data, Methods and Models (International Workshop). https://lup.lub.lu.se/search/files/6097101/4927002.pdf

Wicaksono, J. A., Setiarini, R. B., Ikeda, O., Novawan, A., & Hariri, N. A. (2021). The use of H5P in teaching English. In *Proceedings of the First International Conference on Social Science, Humanity, and Public Health (ICOSHIP 2020)* (pp. 227–230). Atlantis Press. https://doi.org/10.2991/assehr.k.210101.049

APPENDIX A

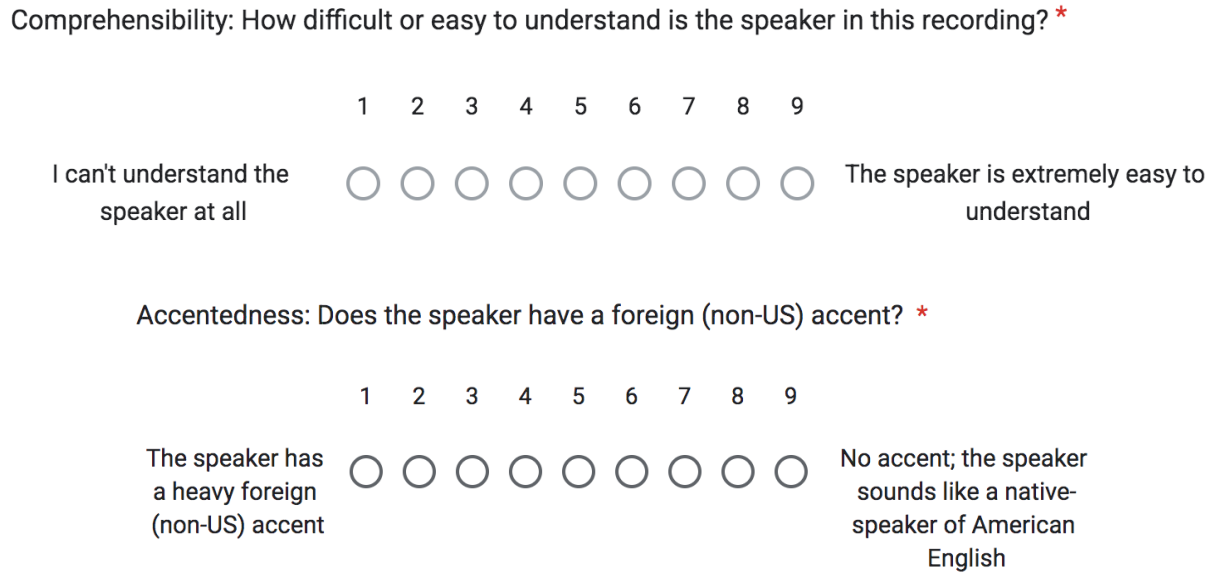
Speech Sample Prompts

Required Recording #1
Click "Record" and read the sentences below:

1. Local produce often comes from rural areas.
2. It's time to decide on your perfect career.
3. She was successful at the conclusion of the academic year.
4. Athletic achievements require dedication and hard work.
5. This weekend, I plan to finish my homework early.
6. Their wedding was a beautiful celebration of their love.
7. My neighbors decided that it was time for a new car.
8. With her class schedule, she has no social life.
9. It's a great exhibit: you should check it out.
10. They're *gonna* eat dinner when he gets home.

Required Recording #2
Choose one (1) of the questions below. Click "Record" and answer it. Speak in full sentences for at least 1 minute.

1. What is your favorite place to visit, and why? Describe it to someone who has never been there.
2. Explain how to cook your favorite dish or prepare a traditional food from your culture.
3. What advice would you give to someone moving to the United States for the first time?

APPENDIX B

Evaluation Scales for Raters