Language in India www.languageinindia.com ISSN 1930-2940 Vol. 25:9 September 2025

The Benefits of Internet-Based Language Tools for Teaching English Pronunciation to Non-Native English Speakers in Thailand: A Synopsis

Benjamin Leo Balanoff
Assumption University Hua Mak
Bangkok, Thailand
benjaminbalanoff@gmail.com

Abstract

This extended summary establishes the core arguments, evidence, and implications of the paper while preserving its structure and emphasis. It traces the logic from problem framing through literature synthesis, benefits and implementation strategies, limits and risks, and closes with role-specific recommendations and future research directions that are tailored to Thailand's English-language education system.

Introduction: Purpose and Rationale

The paper addresses a long-standing imbalance in Thai English education: learners' written proficiency frequently outpaces their spoken intelligibility, with pronunciation being the primary bottleneck. While English occupies a central place in Thailand's national curriculum and in students' academic trajectories, workplace mobility, and regional integration (ASEAN), systematic pronunciation teaching is often uneven. In practice, time and assessment lean toward grammar, vocabulary, and reading—skills that align with national testing systems—while

pronunciation often appears as occasional imitation or read-aloud practice with little feedback and limited assessment.

Through this framework, the paper investigates whether internet-based tools—including mobile-assisted language learning (MALL) apps, automatic speech recognition (ASR) platforms, and phonetic visualization software—can help close the gap between curricular ambition and classroom reality. It advances the proposition that these tools will not replace teachers, nor should they, but that they can systematically supply what large, exam-oriented classrooms struggle to provide—individualized, immediate feedback, abundant low-anxiety practice, multimodal input and data that guides instruction.

The scope is deliberately broad. The paper treats pronunciation as a composite of segmental accuracy (e.g., /l/–/r/, dental fricatives, final consonants, vowel distinctions) and suprasegmental control (stress, rhythm, and intonation)—with the latter frequently neglected yet tightly linked to intelligibility. It situates Thailand's context within current SLA frameworks and it engages the sociolinguistic reality that English today is a lingua franca: intelligibility across diverse accents matters more than native-like imitation. The analysis stretches from classroom practice to policy (curriculum standards and assessment), teacher education (phonetics and digital literacy), and equity (rural access, device/data inclusion), because technology only produces impact when embedded within coherent systems.

Problem Framing

The Stakes of Intelligibility

Global communication hinges on speakers being understood the first time. Research summarized in the paper shows that listeners' judgments of competence and comprehensibility are more sensitive to pronunciation than to grammatical accuracy. A Thai graduate who writes well but cannot make themselves clearly understood in a meeting, interview, or classroom setting faces a disadvantage unrelated to their conceptual ability. Therefore, pronunciation is not a superficial enhancement. It is structural to opportunity.

Thai Systemic Constraints

Three systemic factors negatively impact pronunciation outcomes in Thailand:

1. **Curricular Emphasis and Assessment.** Even where official syllabi invoke "communicative competence," practical assessment systems reward grammar and reading, with a limited focus on oral practice. Pronunciation targets are rarely built into outcomes or rubrics, and teachers often lack time and tools to grade oral production.
2. **Teacher Preparation and Confidence.** Many teachers report limited formal training in articulatory phonetics and in diagnosing errors typical of Thai-to-English transfer (e.g., /l/–/r/, final stops, vowel length). When teachers are unsure of modeling, they understandably revert to safer, text-driven tasks. Large classes (40–50 students) compound this: whole-class speaking practice is noisy, time-consuming, hard to monitor, and difficult to assess fairly.
3. **Access and Exposure.** Outside Bangkok and a few urban centers, consistent exposure to proficient English speech is limited. Students seldom hear sustained, varied models, especially for prosody (stress timing and intonation patterns) that are not obvious in print.

Limited exposure makes it harder to build broader perceptual categories, particularly for contrasts absent in Thai.

Thai L1 Transfer: Segmentals and Suprasegmentals

The paper synthesizes a decade of findings to outline a predictable difficulty profile:

- **Segmental issues:** There is persistent /l/–/r/ confusion and the substitution of dental fricatives /θ/ and /ð/ with /t/ and /d/, frequent final consonant deletion or unreleased stops and vowel merging (e.g., ship/sheep, full/fool) due to fewer Thai categories and length differences. These errors affect accent and, depending on the context, intelligibility.
- **Suprasegmental issues:** Thai is tonal, while English is stress-timed. Learners often produce flat intonation, misplace lexical stress and maintain syllable timing that obscures information structure. Research cited in the paper emphasizes that mistimed stress and monotonous or misplaced pitch movement can lower intelligibility even when single phonemes are correct.

Sociolinguistic Factors

Learners frequently report pronunciation anxiety and fear of ridicule, which reduces willingness to attempt extended speech. In addition, accent ideology—the elevation of a single "native" target—can demotivate learners who, despite progress, still diverge from that standard. The paper reframes the goal as intelligibility within English as a Lingua Franca (ELF): Thai learners must understand and be understood by speakers from diverse linguistic backgrounds, not only by North American or British L1 speakers.

Theoretical Foundations for Technology-Mediated Pronunciation

The paper aligns the functions of internet-based tools with well-established SLA mechanisms:

- **Interaction Hypothesis:** Learning accelerates through interaction that provides feedback and opportunities to repair misunderstandings. ASR role-plays, voicebots, and structured dialogues simulate turn-taking and corrective moves, allowing learners to practice more interaction than class time permits.
- **Noticing Hypothesis:** Errors must become perceptually noticeable for acquisition to occur. Visualization (spectrograms and pitch contours) and highlighted contrasts make subtle misunderstandings visible and audible, helping learners "see" stress and intonation or compare their waveform to a target.
- **Output Hypothesis:** Push learners to produce and self-correct. Production done under constraints reveals gaps that input alone may not. ASR-based drills trigger repeated attempts until the signal matches the target.
- **Self-Determination Theory:** Motivation rises with autonomy (self-paced practice), competence (immediate feedback and visible progress), and relatedness (peer sharing or class-linked goals). Gamified progression helps to harness these drivers, especially for shy students who prefer to avoid public speaking.

By mapping features (instant feedback, repetition, visualization, autonomy) to mechanisms (feedback, noticing, pushed output, intrinsic motivation), the paper explains why internet-delivered pronunciation work can yield learning conditions rarely achievable at scale inside a typical Thai classroom.

Literature Synthesis

Across recent analyses and classroom studies summarized in the paper, several patterns emerge:

1. **ASR-Supported Training Improves Segmentals.** Multiple studies report that learners receiving real-time phoneme-level feedback (e.g., on /l/–/r/, final stops) make larger gains than students in imitation-only or delayed-feedback conditions. Thai studies echo this: frequent, short app sessions correlate with clearer segmental production and greater practice volume.
2. **Visualization Lifts Prosodic Control.** Tools that display pitch contours and syllable prominence help learners reshape intonation patterns. For Thai speakers, seeing a rising yes/no question contour or a stress shift is often the moment of realization needed to internalize a different rhythm.
3. **Autonomy and Motivation Matter.** Learners sustain practice when they can choose modules, track progress, and repeat safely. Thai university associates using mobile apps report lower anxiety and higher willingness to communicate, consistent with the affective curriculum.
4. **Blended Beats Purely Digital or Purely Traditional.** The strongest outcomes appear when teachers explicitly integrate app tasks with lesson targets and then use class time for communicative application (role-plays, presentations, peer feedback). In such models, technology supplies a density of practice and feedback and the teachers supply the meaning, discourse, culture, and assessment.

5. **Gaps Persist.** The paper highlights the under-representation of Thai contexts in long-horizon, controlled studies and the limited analysis of rural deployment and low-resource schools. It also highlights the weak diagnostics for suprasegmentals in mainstream apps and the scarce research into how feedback phrasing affects Thai learners' uptake and motivation.

Benefits of Internet-Based Tools in the Thai EFL Context

1) Immediate, Individualized Feedback

Where one teacher cannot feasibly correct 45 students in real time, ASR scoring and model playback give every learner their own feedback loop. The paper emphasizes that specificity matters: "You said /l/; the target is /r/" + slow-motion model + articulatory tip can break entrenched habits more effectively than general advice like "speak more clearly." Frequent micro-attempts build motor patterns and reduce linguistic fossilization.

2) Autonomy and Practice Volume

Thai students' class time is limited and oriented towards tested skills. Apps extend practice to commuting time, after-school slots, and home use, multiplying speaking attempts per week. The sense of control—choosing modules, setting streak goals—converts idle minutes into cumulative exposure. Over weeks, this raises both perceptual acuity and production stability.

3) Visualization and Multimodal Feedback

For tonal-language speakers, English prosody is counterintuitive. Visualization renders invisible features visible: a flat line versus a rising contour, light vs. heavy stress and vowel length patterns. When paired with slowed audio and articulatory animations, visual channels promote noticing and increase self-correction. Multimodal redundancy (audio + visual + text) also supports learners with different processing strengths.

4) Accessibility and Equity

The paper argues that mobile-first, low-bandwidth tools can be equalizers where trained phonetics teachers are scarce. Downloadable modules allow practice in areas where internet connectivity is low and shared devices or school-managed labs can spread access. Critically, apps can include multiple accent models, widening exposure beyond a single status norm and preparing learners for ELF realities.

5) Differentiation at Scale

Apps adaptively target weak points (e.g., final /t/ release, /v/ vs. /w/) and adjust difficulty. Teachers can assign phoneme or pattern-specific tasks by group and review class dashboards to spot common problem areas (e.g., "half the class isn't marking question rises"). This pushes individualized support into large classes without radically increasing teacher workload.

6) Confidence and Affective Safety

Private, repeatable practice reduces social risk. Visible progress—scores, contour alignment, unlocked levels—translates effort into competence signals. Studies summarized in the paper link

short daily practice with higher willingness to communicate and lower speaking anxiety, especially for learners who are reluctant to speak up in class due to shyness or introversion.

Implementation for Thailand: From Concept to Routine

The paper turns from "can" to "how," outlining system-compatible ways to embed pronunciation technology.

Curriculum and Assessment Alignment

Pronunciation gains durability when it is planned for and assessed. The paper proposes that national and school-level curricula define CEFR-aligned descriptors for both segmentals and suprasegmentals (e.g., B1: maintains primary word stress, uses rising intonation in yes/no questions, sustains final stop consonants in high-frequency words). When outcomes exist, teachers have permission (and pressure) to allocate time, and students recognize pronunciation as graded work, not extra credit.

Rubrics can incorporate tool-generated items (audio portfolios, app progress logs) and blend them with live tasks (role-plays, presentations). The key is to construct alignment: if intelligibility is the goal, assessment must capture clarity in spontaneous speech, not just within app scores.

Rural Inclusion and Infrastructure

The paper acknowledges infrastructure constraints—patchy Wi-Fi, limited devices, charging logistics, data costs—and responds with layered strategies:

- Offline/low-data modules to protect online practice from bandwidth shocks.
- Device pools (school carts, library checkout) to serve students without smartphones.
- Charging stations and basic maintenance plans to keep devices usable.
- Partnerships with telecoms for educational data bundles.

Because rural schools often tend to have fewer trained teachers, thoughtfully delivered apps can disproportionately benefit these learners and provide additional access and support systems.

Teacher Professional Development

Technology impact rests on teacher agency. The paper recommends practice-rich professional development that integrates:

- Phonetics essentials tailored to Thai transfer issues (articulatory contrasts, common error diagnosis).
- Hands-on tool use (setting up classes, interpreting dashboards, assigning targeted modules).
- Lesson design that connects app tasks to weekly objectives and in-class communicative activities.
- Coaching cycles to sustain adoption: try a routine, review learner data with a mentor, adjust.
- Sample language for giving feedback that is supportive, specific, and actionable ("Your /r/ is close; watch the tongue and avoid lateral airflow. Try again on these three minimal pairs.").

When teachers see before/after audio and class-level trend lines, skepticism tends to give way to instructional curiosity.

Blended Learning Routines

The paper favors repeatable routines over unplanned usage, for example:

- **Flipped micro-practice (10–15 min):** before class, learners complete an app module on a target (e.g., final stops, question intonation).
- **In-class consolidation (15–20 min):** teacher demos, drills, and communicative tasks (info-gap dialogues, role-plays) that require the target pattern.
- **Post-class reflection (5–10 min):** learners record a short monologue or dialogue, submit to the app and to a class portfolio; teacher samples a few clips to plan next steps.

Over weeks, this routine builds muscle memory (via app density) and transfer (via classroom communication). Critically, it normalizes pronunciation as a weekly strand, not a one-off unit.

Challenges and Limitations

The paper offers a balanced appraisal of risks.

1. **Infrastructure and Access.** Without devices, charging, or data, technology amplifies inequality. The remedy is policy-level support—hardware refresh cycles, shared-device programs, and subsidized data—plus instructional designs that do not presume that students have access to universal, 24/7 connectivity.

2. **Teacher Resistance and Cognitive Load.** Some teachers fear being displaced or even replaced. Others may worry about the competence of their own pronunciation. Others may feel overwhelmed by one more platform. Professional development must explicitly address mindset, model design, doable routines, and honor teacher expertise: technology is a teaching assistant, not a replacement.
3. **Overreliance and Misalignment.** App scores can become the goal, diluting communication aims. ASR can also inaccurately score heavily accented or beginner speech. Therefore, the paper insists on teacher mediation: use app data to target instruction but also evaluate progress with live, communicative tasks and intelligibility rubrics.
4. **Suprasegmental Diagnostics Remain Weak.** Many tools excel on phonemes but are rough on rhythm and intonation. Teachers should pair apps with explicit prosody instruction (thought groups, nuclear stress, contrastive stress, question tunes) and use visualization as a support, not as the sole authority.
5. **Learner Variability and Motivation.** Autonomy helps many learners while others struggle without structure. Competitive gamification can demotivate those at the bottom of leaderboards. The solution is goal-based, mastery-oriented framing (effort and improvement count) and teacher check-ins to sustain habits.
6. **Data Privacy and Sustainability.** While it is not the paper's central focus, sustainable adoption implies clear data policies, transparent pricing, and avoidance of vendor lock-in. Schools need exit ramps and content portability in the event that online tools change.

Recommendations

For Educators (Classroom and Program Level)

- **Make pronunciation a feature.** Allocate weekly minutes, name the target (segmental and prosodic), and cement it in goals learners can recognize in their own speech.
- **Tie app tasks to communicative use.** Assign a focused app module (e.g., /l/–/r/ minimal pairs, rising question intonation) and immediately use that target in meaningful, short tasks (surveys, role-plays, peer interviews).
- **Use analytics as formative assessment.** Scan class dashboards to find common errors and plan micro-lessons and targeted listening discrimination work for the next class.
- **Build learner autonomy deliberately.** Teach how to set weekly goals and reflect on feedback ("What changed in your pitch contour?"), and track progress with an audio portfolio.
- **Normalize error as information.** Frame feedback as instructional, not punitive. Celebrate micro-gains (clearer final /t/, a correctly rising yes/no question).
- **Design inclusive routines.** Offer non-competitive challenges (e.g., class streaks) and flexible ways to demonstrate progress to avoid demotivating learners who dislike leaderboards.

For School Leaders and Policymakers

- **Codify outcomes.** Add CEFR-aligned pronunciation descriptors to standards and local syllabus, explicitly covering prosody alongside phonemes.

- **Fund access and resilience.** Budget for device pools, charging carts, and connectivity solutions and negotiate education data bundles with telecom providers while also prioritizing rural schools.
- **Invest in sustained professional development.** Provide coaching-oriented professional learning that blends phonetics with practical, tool-integrated lesson design. Recognize and reward teacher leadership in pronunciation instruction.
- **Pilot, evaluate, scale.** Run structured pilots with comparison groups; collect intelligibility-focused evidence (teacher-rated speech samples, listener judgments) alongside app metrics and gauge what proves effective.
- **Adopt plural accent models.** Encourage tools and materials that expose learners to diverse Englishes, supporting the ELF target of clear, flexible intelligibility rather than rigid native-like speech.

For Researchers

- **Run long-term Thai studies.** Track learners across semesters to assess retention and transfer to spontaneous speech, not just word-list performance.
- **Diversify comparisons by learner profile.** Compare rural-urban, adolescent-adult, beginner-intermediate, and high- vs. low-anxiety groups to understand moderators of effect.
- **Probe feedback quality**. Examine ASR accuracy for Thai transfer patterns and test which feedback phrasings and visualizations lead to successful self-repair and sustained motivation.

- **Compare implementation models.** Experimentally contrast traditional, purely digital, and blended designs on intelligibility, motivation, and cost-effectiveness.
- **Document teacher change.** Study how professional development models alter teacher beliefs, routines, and capacity to diagnose and respond to prosodic vs. segmental issues.

For EdTech Providers and Partners

- **Strengthen prosody analytics.** Improve detection and explanation of stress, rhythm, and intonation and provide interpretable visualizations and targeted practice at the phrase level.
- **Design for low-resource contexts.** Provide offline modes, small-footprint updates, device-diagnostic performance, and local language support for teacher dashboards.
- **Support teacher workflows.** Class dashboards that provide actionable patterns, not just scores and exportable evidence for portfolios, including simple ways to assign targeted practice.
- **Respect privacy and portability.** Clear data policies and the easy export of learner audio and progress alongside pricing that schools can sustain year to year.

A Concrete Scenario

A lower-secondary school in Isaan decides that every Grade 9 class will spend 15 minutes per week on pronunciation. The department defines a 12-week cycle (weeks 1–4: final stops and plural -s; weeks 5–8: question intonation and sentence stress; weeks 9–12: contrastive

stress). Each week, students complete a 10-minute app module at home (downloadable), then, in class, the teacher runs a 15-minute routine:

1. Micro-demo of the target (2–3 minutes) with quick perception checks.
2. Communicative pair task requiring the target (7–8 minutes), with teacher roaming and coaching.
3. Two volunteer recordings played for the class to notice stress/intonation or final consonants, framed positively.
4. Exit reflection: students note one micro-gain and one next step in their log.

Every other week, students record a 30–45-second monologue (e.g., weekend plans, a short story retell). The teacher samples five recordings, tags common issues in the dashboard, and plans the next week's micro-demo accordingly. The school maintains a device cart for learners without smartphones and the telecom partner provides an education data plan for after-school use. After one term, teachers report that more students volunteer to speak, final consonant deletion has declined markedly, and question intonation is audibly more consistent. Administrators hear cleaner speech in morning assemblies and students say they feel less intimidated to speak English.

Conclusion

The paper's central claim is cautious and optimistic in equal measure. Internet-based tools are not magic, but they are exceptionally well-suited to the precise weaknesses of pronunciation teaching in Thailand: they create countless private, low-stakes practice

opportunities while delivering immediate, individualized feedback that teachers cannot scale. They make prosody visible and track growth in ways that inform instruction. When these affordances are harnessed within blended routines, grounded to clear outcomes, and supported by teacher learning and equitable access, Thai learners can move from hesitant, text-bound English users to speakers whose pronunciation (segmental and prosodic) enables them to be understood clearly across diverse contexts.

Crucially, the target is not a single "native" accent. In an ELF world, intelligibility, flexibility, and confidence are the real currency. The paper reframes success accordingly and supplies practical pathways for classrooms, schools, and ministries to act. If Thailand aligns curriculum, assessment, professional development, and infrastructure around this reframing, pronunciation can shift from a neglected add-on to a core pillar of communicative competence, opening doors for learners academically, professionally, and personally.

References

- Amrate, M., & Tsai, P.-H. (2024). Computer-assisted pronunciation training: A systematic review. *ReCALL, 36*(1), 1–20. Cambridge University Press. https://doi.org/10.1017/S0958344023000203
- Blake, J. (2024). An open CAPT system for prosody practice: Practical steps toward a multilingual setup. *Languages, 9*(1), 27. https://doi.org/10.3390/languages9010027
- Chaisiri, P. (2023). Promoting Thai EFL learners' willingness to communicate via technology-mediated oral tasks. *LEARN Journal: Language Education and Acquisition Research Network, 16*(1), 97–120.

- Fouz-González, J. (2025). Prosody visualization in computer-assisted pronunciation training: Current practices and future directions. *System, 119*, 103213. https://doi.org/10.1016/j.system.2024.103213
- Hongnaphadol, W., & Attanak, A. (2022). Reducing Thai EFL students' pronunciation anxiety through a CAPT-based Reading Progress application. *Journal of Liberal Arts, Prince of Songkla University, 14*(1), 83–122. https://doi.org/10.14456/jlapsu.2022.4
- International Telecommunication Union. (2023). *Measuring digital development: Facts and figures 2023.* Geneva: ITU. https://www.itu.int/itu-d/reports/statistics/facts-figures-2023
- Jaiprasong, S., & Pongpairoj, N. (2020). L2 production of English word stress by L1 Thai learners. *LEARN Journal: Language Education and Acquisition Research Network, 13*(2), 143–156.
- Lou, L., Chen, W., & Zhao, Y. (2024). Intelligent ASR-based speaking systems for EFL learners: A bibliometric approach and meta-analysis. *Computer Assisted Language Learning.* Advance online publication. https://doi.org/10.1080/09588221.2024.2321965
- Mohsen, M. A. (2025). A scientometric study of computer-assisted pronunciation training (CAPT): Trends, challenges, and future directions. *Humanities and Social Sciences Communications, 12*, 391. https://doi.org/10.1057/s41599-025-04474-y
- Pinweha, S. (2025). The effects of differentiated pronunciation instruction via a virtual learning environment on primary-school Thai EFL teachers. *PASAA, 70*, 164–203.

- Songyai, P., & Suwanarak, K. (2022). Thai EFL teachers' beliefs and reported practices: An investigation of English pronunciation instruction. *Psychology and Education, 59*(1), 1650–1661.
- Suntornsawet, J. (2022). A systemic review of Thai-accented English phonology. *PASAA, 63*, 351–369.
- World Bank, & Equitable Education Fund. (2021). *Thailand education: Reimbursable advisory services—Final report.* Bangkok: EEF/World Bank.