

## Mobile-Assisted Pronunciation Instruction in Bangladesh: Challenges and Opportunities

Md. Absar Uddin

Assistant Professor, Department of English Language and Literature  
International Islamic University Chittagong, Bangladesh

---

### ABSTRACT

This review brings together published studies on the use of mobile tools for learning English pronunciation in Bangladesh and compares them with global evidence. The findings show that mobile devices can increase practice opportunities and build learner confidence, but access is often limited by shared devices, high data costs, and weak connectivity. Most studies highlight listening and repeating with model voices, but progress is often small when feedback is not clear or actionable. The review is organized around the following main themes: access and motivation, feedback and assessment, task design, teacher mediation, and equity. It argues that effective use of mobile tools requires short practice routines linked to classroom teaching, feedback that learners and teachers can understand, and a balance between important sounds and basic features of rhythm and stress. Materials that use little data or work offline are especially valuable. Overall, mobile-assisted pronunciation instruction can offer fair and practical benefits when designed for local conditions.

---

**Keywords:** Mobile-assisted language learning (MALL), Mobile-assisted pronunciation instruction (MAPI), Bangladesh, Automated speech recognition (ASR), Corrective feedback, Teacher mediation

---

### INTRODUCTION

Pronunciation is a key part of speaking because it affects both intelligibility and comprehensibility. *Intelligibility* is listeners' actual understanding of a speaker's message, typically measured by word or sentence transcription accuracy. It reflects what was understood, not how hard it was to understand. [24] [6] and *Comprehensibility* is the perceived ease or difficulty of understanding—how effortful the speech is for listeners—even when the content is understood. It is usually captured with listener ratings on ordinal scales [24] [37].

Learners can be understood even if they have a noticeable accent, but speech becomes difficult to follow if pronunciation falls below a certain level [5] [24]. In Bangladesh, many learners face persistent difficulties with both sounds (segmentals) and features of stress and intonation (suprasegmentals). Research reports problems with vowel contrasts, aspiration, syllable structure, word stress, and intonation, which all reduce the clarity of speech in both secondary and tertiary contexts [16] [20] [22]. At the same time, classroom studies suggest that pronunciation often receives limited attention because of time pressure, lack of materials, and limited teacher training [27]. The learning environment in Bangladesh is becoming more mobile. National and industry reports show a steady rise in mobile subscriptions and internet use, though problems of affordability and network coverage remain [12] [13]. This situation highlights both the opportunities and inequalities in mobile access.

Mobile-assisted language learning (MALL) is therefore a practical option for supporting pronunciation because it allows short practice, flexible feedback, and learning at any place and time—if the design matches local needs. Globally, MALL is now well established as an area of research, defined as the use of mobile technologies for language learning where portability provides unique benefits [18]. Within pronunciation teaching, Mobile-Assisted Pronunciation Instruction (MAPI) often uses short tasks such as shadowing, minimal pairs, or rhythm practice with feedback delivered by text, audio, or automated speech recognition (ASR). Evidence from international studies shows that such tools can improve pronunciation and learner motivation, though results are not always consistent [21] [33]. Meta-analyses focusing on ASR also report medium overall effects, with stronger outcomes when feedback is clear and corrective [38] [34]. Still, suprasegmental features like rhythm and intonation receive less attention than segmental ones, and links to pedagogy remain uneven [33].

In Bangladesh, studies on mobile learning have increased but evidence is scattered across theses, local journals, and international publications. Few focus directly on pronunciation. For example, Begum [2] reports a classroom case study with first-year English majors at a Bangladeshi public university, using SMS micro-tests and teacher feedback, triangulated with observation and questionnaires; outcomes indicate basic phone functions can bolster EFL activities despite policy bans and technical

limits. Ibna Seraj et al. [31] employ an explanatory sequential mixed-methods design (survey of tertiary EFL learners followed by interviews), finding high smartphone readiness for oral English alongside barriers such as scarce authentic materials, limited speaking opportunities, and large classes. Extending to intervention, Ibna Seraj & Klimova [32] implement a nine-week smartphone-apps-based method featuring WhatsApp, calls, and voice recording; students report greater flexibility and increased speaking practice, while teachers note heavier assessment load and classroom noise. Mumu's [23] findings show mobile-assisted language learning boosts vocabulary, motivation, and reading comprehension when tasks are short, multimodal, and feedback is timely. Outcomes improve with accessible, low-bandwidth design and teacher orchestration; however, weak connectivity, device inequality, and limited training constrain scalability and consistent achievement gains.

The growing presence of commercial ASR-based apps such as ELSA Speak and Parlo also shows that pronunciation technology is reaching learners outside classrooms, though cost and connectivity continue to be barriers [3] [7].

This review therefore brings together Bangladeshi research on mobile-assisted pronunciation instruction (MAPI) and examines it in relation to global evidence. It aims to: (i) describe the phonological and instructional context that shapes learners' needs, (ii) review published work on mobile tools for pronunciation in Bangladesh, (iii) identify features of design and classroom use linked to positive outcomes, and (iv) highlight research gaps, especially in relation to equity, feedback transparency, and the teaching of suprasegmentals. The goal is to provide context-sensitive guidance for making MAPI practical and pedagogically meaningful in Bangladesh.

#### **Context: Bangladeshi EFL & Mobile Access**

English language teaching in Bangladesh has developed under the influence of communicative language teaching (CLT). This method was expected to make English learning more interactive and effective, but its results in the country have been limited. Scholars argue that many barriers, such as very large classes, limited teaching resources, exams that do not match classroom practice, and insufficient teacher training, have weakened the success of CLT. These barriers are especially strong in the case of pronunciation, which often does not receive enough time or attention in lessons [14]. More recent reviews describe English teaching as "bogged down," meaning that progress is slow and often ineffective. They also call for approaches that are better adapted to the Bangladeshi context instead of importing international methods without changes.

A key difficulty comes from the large phonological gap between Bangla and English. The two languages differ in both segmental and suprasegmental features. Segmental features are the individual sounds of a language, while suprasegmentals refer to features such as stress, rhythm, and intonation. Research shows that Bangladeshi learners often face challenges with vowel contrasts, aspirated consonants such as /p<sup>h</sup>, t<sup>h</sup>, k<sup>h</sup>/, and syllable structures. They also struggle with word stress and sentence-level intonation. These problems reduce intelligibility, meaning that learners may not be easily understood even if their grammar is correct [16] [22]. Such evidence makes clear that explicit and systematic pronunciation teaching is necessary. Without it, learners may continue to face communication difficulties that affect both classroom learning and real-life interaction.

At the same time, mobile phone use has become a central feature of everyday life in Bangladesh. According to national reports, the number of mobile subscriptions reached around 188 million by 2025, which is almost equal to the population size. However, this number also reflects the common practice of using multiple SIM cards rather than every individual owning a phone [1] [4]. Household internet access has also grown, with more than half of homes now connected. Still, mobile internet subscriptions have dropped in recent years because of the rising cost of data, showing that access is not always secure or sustainable (BBS as reported in [8] [9]). These patterns highlight the tension between the wide reach of mobile phones and the fragile nature of actual access.

This issue is not limited to Bangladesh. Across South Asia, According to [12] [13]. affordability, digital skills, and network coverage as the most important barriers to mobile internet adoption. These challenges directly affect how mobile-assisted language learning (MALL) can be designed and implemented in Bangladesh [12] [13]. If learning tools require high amounts of data or stable connections, they may exclude a large group of learners. On the other hand, tools that are low-cost, offline-capable, and easy to use have the potential to make learning more inclusive.

Bangladesh also has valuable experience with mobile learning through large-scale educational projects. One well-known example is the English in Action programme, which delivered mobile-based content and classroom audio to teachers. This initiative showed that mobile technologies can support new ways of teaching and improve student performance when pedagogy and logistics are carefully aligned. It also proved that even in resource-limited schools, small, predictable mobile-based tasks can be sustained over time [39]. This experience

demonstrates that mobile delivery is not only possible but also effective when it fits local classroom realities. Taken together, these points show that Bangladeshi English teaching faces two connected realities. On the one hand, there is a strong need for systematic pronunciation support because of the phonological distance between Bangla and English. On the other hand, there is an opportunity to use mobile technologies, since mobile phones are widespread and can support flexible practice. The challenge is that access remains fragile, and without careful design, mobile learning can increase inequalities rather than reduce them. To be successful, mobile-assisted pronunciation instruction must take account of these conditions by being low-cost, teacher-supported, and sensitive to the specific needs of Bangladeshi learners.

### Research Gap and Rationale

Evidence on mobile-assisted pronunciation instruction (MAPI) in Bangladesh is still limited and uneven. Most studies use short interventions with small convenience samples and assess accuracy rather than listener-based outcomes such as intelligibility and comprehensibility. Suprasegmental feature --word stress, rhythm, connected speech, and intonation—receive far less attention than individual sounds. Little is known about how transparent and reliable ASR feedback is for Bangladeshi English, including accent sensitivity and noise effects typical of classrooms. Reports rarely detail classroom integration, teacher roles, formative assessment, or workload. Equity issues—device access, data costs, rural connectivity, and accessibility for learners with disabilities—are seldom examined, and cost-effectiveness and sustainability remain unclear. These gaps limit practical guidance for teachers and policy makers. The present study addresses this need by synthesizing available evidence, highlighting context-specific pronunciation challenges, and outlining low-cost, classroom-ready approaches that prioritize intelligibility, comprehensibility, and transparent feedback within the Bangladeshi EFL context.

### METHODOLOGY

This study examined only published work on mobile-assisted pronunciation learning. Sources comprised peer-reviewed journals, academic books, and conference proceedings accessible through recognized databases and publisher sites. Items entered the pool when they investigated mobile devices or mobile-delivered audio for pronunciation practice or classroom instruction. Publications focused solely on grammar, vocabulary, or non-mobile technologies were set aside. Relevant items were located through targeted keyword searches combining terms for mobile learning, pronunciation, and second-language learning, and through checks of reference lists. Only full-text items in English were considered, and search dates covered all available

publication years. After screening, the selected publications were organized by common concerns: access and motivation; feedback and ways of checking progress; task design; teacher support; and fairness and inclusion. Within each group, points of agreement and disagreement were compared, with likely reasons—such as setting, learner profile, or activity type—noted. The aim was to provide clear, organized takeaways for teaching and research rather than to calculate pooled effect sizes.

### FINDINGS AND DISCUSSION

When examined together, research from Bangladesh and international reviews shows that mobile-assisted pronunciation instruction (MAPI) can be very effective in producing medium gains in L2 pronunciation, including ASR-supported interventions [26] [19]. The most successful cases provide learners with frequent opportunities to practice, supply feedback that is quick and clear, and connect mobile practice with classroom teaching. Students and teachers often view mobile tools positively because they make practice easier, less stressful, and more flexible. However, the benefits are not always stable. Many learners face barriers such as shared phones, costly data packages, and unreliable internet access, which reduce the possibility of sustained practice [13] [30].

In Bangladesh, most reported activities focus on listening and repeating after model recordings, also called shadowing. While these tasks help learners to become more aware of pronunciation, they often fail to explain what exactly is wrong with a student's speech or how to improve it. By contrast, global reviews consistently show that the strongest learning gains occur when learners receive explicit and understandable corrective feedback. Automated speech recognition (ASR) tools or teacher-led corrections that clearly identify the error and demonstrate the correct form are especially effective. Without such feedback, progress is slower and less visible [21] [33] [38] [34].

The phonological distance between Bangla and English also strongly influences outcomes. Because the two languages differ in both vowels and consonants, learners often struggle with aspiration, vowel contrasts, and syllable structures. Predictable transfer errors reduce intelligibility and make communication difficult. At the same time, suprasegmental features such as stress, rhythm, and intonation are often neglected, even though they are essential for being understood in real conversations. Both local and international studies argue that these features deserve more attention in mobile practice and classroom teaching [16] [22].

### **Access and Motivation**

Evidence from Bangladesh clearly shows that mobile phones create new opportunities for learners to practice pronunciation. Students report that phones allow them to practice privately and repeatedly, which reduces anxiety about speaking and increases their confidence. They also value the flexibility of mobile practice, since they can use small amounts of time during commutes, breaks, or after class. Teachers appreciate that mobiles enable them to assign short pronunciation tasks without requiring computer labs or other heavy infrastructure [30].

This pattern matches findings in international research. Studies show that brief, focused mobile tasks not only improve pronunciation but also motivate learners to continue practicing. The sense of progress and the ability to repeat tasks many times help to build persistence and positive attitudes toward speaking [21] [33].

However, motivation is fragile when access is weak. Many students in Bangladesh share phones with family members, face difficulties with charging or storage, or cannot afford continuous data plans. These problems interrupt regular practice and reduce the chance of long-term improvement [13]. In addition, when learners are not given clear guidance on what to focus on, practice becomes unfocused. Students may repeat tasks without improvement, which reduces motivation over time. Research suggests that motivation depends not only on access to a device but also on the design of practice tasks. Tasks should be short, clearly targeted, and available in offline or low-data formats to ensure that practice is realistic in local conditions [13] [30].

Bangladesh has shown in the past that large-scale mobile learning projects can work when technology and pedagogy are aligned. The English in Action programme used simple audio delivered through teacher phones to increase speaking practice in classrooms. It demonstrated that small, predictable tasks can be sustained even in schools with limited resources, leading to noticeable improvements in classroom participation and learning [39]. This experience suggests that the same principles can be applied to pronunciation: access alone is not enough; practice must be carefully designed and teacher-led to remain meaningful.

### **Feedback and Assessment**

Feedback is one of the weakest areas in Bangladeshi MAPI studies. Most feedback consists of students listening to a model and repeating it. This approach helps raise awareness but does not explain exactly which sound is wrong, how it should be corrected, or why it matters. Without this diagnostic element, progress is often limited.

International evidence points in another direction. Reviews and meta-analyses show that feedback is most effective when it is explicit, immediate, and easy for learners to understand. Mobile tools that integrate ASR, when used transparently, provide learners with direct comparisons between their speech and the target. Teacher-led corrective feedback can also achieve strong results when it clearly names the error and demonstrates the correct production [21] [33].

Quantitative studies confirm this advantage. ASR-supported pronunciation training generally produces medium improvements, and the effects are larger when learners receive corrective feedback that identifies specific errors [38]. Qualitative studies also show motivational benefits: learners feel more engaged when they can see or hear the difference between their own production and the model [34]. However, this only works when the feedback is transparent. If a system uses colors, scores, or waveforms, both learners and teachers must know what these signals mean. Otherwise, feedback becomes confusing, and students may return to simple repetition without improvement.

Assessment practices in Bangladesh also need attention. Many studies rely on self-reports or general speaking exams, which do not specifically measure changes in pronunciation. International research recommends distinguishing between three key outcomes: intelligibility, comprehensibility and accentedness. *Accentedness* is the perceived degree of “foreign accent,” i.e., how different the speech sounds from a target-language norm; it is distinct from intelligibility and comprehensibility and can be high even when speech is understandable [24] [6].

Selecting appropriate tasks and measures for each construct makes progress easier to see [28] [5] [24]. In classroom practice, simple and short assessment methods can be effective, such as rubric-guided read-alouds, short pre/post probes on specific features, or quick intelligibility ratings. These tools make learning visible without creating extra workload for teachers.

### **Task and Design Features**

The types of tasks used in Bangladeshi studies are mainly shadowing, minimal pairs, and short repetition activities. These are effective to some extent, but research shows they are more powerful when combined with structured cycles and clear goals. International evidence indicates that short, repeated tasks improve fluency and comprehensibility, even if accentedness changes more slowly. Teachers should therefore set realistic goals, focusing first on making learners easier to understand rather than expecting rapid “native-like” pronunciation [10] [11].

Design quality matters as much as the task itself. Effective mobile practice usually follows a cycle: input (model), learner attempt, feedback, and immediate retry. This cycle helps learners notice errors and correct them. Research also supports sequencing tasks: starting with high-frequency segmental problems, such as aspiration and vowel contrasts, and later moving to prosody tasks, including stress, rhythm, and intonation [38] [34]. In Bangladesh, design must also respond to local access challenges. Tasks should be downloadable, work offline, and have small file sizes so that learners can practice in short “found time” moments. When tasks are directly linked to assessments and are followed by quick classroom consolidation, learners see them as meaningful and are more motivated to continue [13] [39].

### **Teacher Mediation and Classroom Integration**

Research highlights that teachers play a crucial role in making mobile practice effective. Evidence shows that MAPI works best when it fits into a predictable routine that teachers can manage within normal lesson time. A common cycle is: (i) short pre-teaching of the target feature, (ii) mobile micro-practice, and (iii) in-class consolidation through guided repetition or a short communicative task. This structure helps students link mobile practice with course goals and makes pronunciation instruction a regular part of learning [39].

In contrast, when mobile practice is treated as an optional extra, it quickly disappears because teachers are pressured to cover other parts of the syllabus. This was already observed in evaluations of communicative language teaching in Bangladesh, where pronunciation was often neglected due to time and assessment demands [14].

Teachers also need support to use feedback effectively. If ASR is used, teachers must be able to explain what the feedback means and how to guide students in their next attempt. If only model-based recordings are available, teachers can still create effective feedback loops by modeling sounds, asking students to listen carefully, and prompting corrections. International studies confirm that effective pedagogy is more important than advanced technology; even simple tools can work when guided by skilled teachers [34] [38].

### **Equity and Localization**

Equity is one of the most important factors in the success of MAPI. In South Asia, major barriers to mobile learning include affordability, limited digital skills, and poor coverage [13]. In Bangladesh, additional issues such as phone sharing, limited charging facilities, and small storage space make access even more fragile [30]. Tools that require high bandwidth or constant connectivity risk excluding

many learners. Therefore, MAPI must provide low-data or offline-capable options, with short sessions that can be completed even on shared devices.

Localization is equally important. Because Bangla and English have predictable phonological differences, tasks should directly target these areas. For example, learners need practice with aspiration contrasts (/p<sup>h</sup>, t<sup>h</sup>, k<sup>h</sup>/), vowel length and quality, and syllable structure. Short, L1-aware prompts that explain these issues in simple terms can make practice more memorable and actionable [16] [22]. International research also shows that suprasegmentals strongly influence comprehensibility, yet they are often overlooked. Even short tasks on word stress, rhythm, and intonation can improve how easily learners are understood [33] [5]. Adding these features into MAPI can raise communicative payoff without requiring much extra time.

Finally, transparency and ethics are part of equity. When ASR is used, learners should know what is being recorded, how the system produces scores, and what the results mean. Teachers should also feel confident to challenge incorrect scores and provide human correction. Where ASR is not possible, teacher-mediated feedback can still provide fair and high-quality support [34].

The evidence from both Bangladeshi and international studies shows that mobile-assisted pronunciation instruction (MAPI) has strong potential, but its success depends less on technology itself and more on how it is used in classrooms. The findings suggest that access to devices and apps creates new opportunities, but opportunities only turn into learning when practice is purposeful, feedback is understandable, and teaching routines are consistent. In short, MAPI in Bangladesh works best when it is designed for the realities of learners and teachers rather than for the features of a tool [13] [30].

### **Access, Motivation, and Fragility**

One of the strongest benefits of mobile learning in Bangladesh is that it increases access to practice. Learners value the chance to practice privately, repeat tasks many times, and work outside class hours. Teachers also appreciate the flexibility of assigning small, manageable tasks without relying on computer labs or extra class sessions. As seen in earlier projects like English in Action, mobile tools can support regular speaking practice in classrooms that otherwise lack resources [39].

Yet this access is fragile. Learners often share devices, cannot afford steady internet data, and face charging or storage problems. These challenges interrupt learning and reduce motivation. International evidence confirms that motivation depends not only on having a device but also on the design of practice. Short, targeted tasks that work offline or with low

data help maintain focus and persistence [13] [30]. If practice feels unfocused or too costly, students lose interest. For MAPI to be effective, design must take account of these fragile conditions by ensuring activities are short, low-bandwidth, and clearly guided.

### **Feedback as the Key to Learning**

Feedback emerges as the most critical factor shaping pronunciation improvement. In Bangladesh, most MAPI studies rely on listening to model recordings and repeating them. While these activities increase awareness, they rarely identify exactly which sound is incorrect or how to fix it. As a result, progress often remains limited.

Global research shows a different picture. Studies consistently report that feedback is most effective when it is explicit, immediate, and easy to understand. Automated speech recognition (ASR) can be powerful when it shows learners how their speech differs from a model, but only if the system is transparent and its signals (scores, colors, or waveforms) are easy to interpret. Teacher-led correction can also be very effective when it clearly names the error and models the correct production [21] [33].

Meta-analyses confirm these benefits. ASR-supported training tends to show medium-sized improvements, and these are larger when feedback highlights specific errors [38]. Learners also report feeling more motivated when they can hear or see the difference between their output and the target [34]. The lesson for Bangladesh is clear: feedback should never be a “black box.” Whether it comes from technology or from teachers, learners must understand what went wrong and how to improve on the next attempt.

### **Task Design and Sequencing**

The kinds of tasks used in Bangladeshi studies—shadowing, minimal pairs, and repetition—are consistent with international practice. However, the effectiveness of these tasks depends heavily on design. Short, repeatable cycles of 2–5 minutes are far more effective than long, occasional drills. Tasks should also follow a logical sequence: first focusing on high-frequency segmental problems such as aspiration contrasts and vowel quality, then gradually adding suprasegmental features like stress, rhythm, and intonation [10] [11] [38].

Another design principle is the input–practice–feedback cycle. Learners should listen to a model, attempt the target, receive feedback, and immediately try again. This sequence helps them notice errors, correct them, and build accuracy. Without the retry step, feedback loses much of its effect. In Bangladesh, this structure can be adapted to low-data

environments by using small audio packs, offline-capable exercises, and paper-based rubrics combined with teacher guidance [13] [39].

### **Teacher Mediation and Classroom Integration**

Teachers play a decisive role in whether mobile practice actually leads to learning. Evidence shows that when mobile tasks are embedded into predictable lesson routines, they are more likely to be sustained. A simple cycle—brief teaching of a target, short mobile micro-practice, and classroom consolidation—helps learners connect mobile work with course goals. This structure also makes pronunciation a regular, visible part of lessons instead of an optional extra [39].

By contrast, when pronunciation tasks are treated as add-ons, they quickly fade away under syllabus pressure. This problem has long been noted in evaluations of communicative language teaching in Bangladesh, where pronunciation was often ignored due to time and assessment constraints [14]. To avoid this, teachers need lightweight routines that can be delivered within normal class time and do not add unsustainable workload.

Teachers also need support to interpret feedback. If ASR is used, they must be able to explain what scores or symbols mean and guide students on how to improve. If only model-based tasks are available, teachers can still create effective feedback loops through modeling, perception checks, and guided correction. International studies highlight that coherent pedagogy is more important than advanced apps; even simple mobile tasks can be effective if teachers are well prepared [34] [38].

### **Equity, Inclusion, and Localization**

Equity is not an optional consideration but a condition for success. Across South Asia, affordability, coverage, and digital literacy remain the main barriers to mobile internet use [13]. In Bangladesh, these challenges are compounded by shared phones, charging problems, and limited storage [30]. If MAPI assumes stable, high-bandwidth connections, it risks leaving behind large groups of learners. Instead, tools must be designed to work offline, use small downloads, and allow practice in short “found time” sessions.

Localization is equally essential. Bangla and English differ in predictable ways, and these differences should shape task selection. Learners often need focused practice with aspiration (/p<sup>h</sup>, t<sup>h</sup>, k<sup>h</sup>/), vowel length, and syllable structure. Short, L1-aware explanations can make these features easier to remember and correct [16] [22]. At the same time, suprasegmentals should not be ignored. Although they are often neglected, stress, rhythm, and intonation strongly affect comprehensibility. Even

short activities, such as clapping for rhythm or practicing sentence stress, can make learners much easier to understand [33] [5].

Ethical transparency is also part of equity. If ASR is used, learners should know what data is recorded, how it is analyzed, and what the results mean. Teachers should also feel able to challenge incorrect scores and provide human correction. Where ASR is not possible, teacher-led feedback can still ensure fair and effective learning [34].

### **Towards a Research Agenda**

The current evidence base in Bangladesh has several gaps. Many studies focus more on attitudes than on measurable outcomes. Few are classroom-embedded interventions, and most are conducted at the university level. There is a strong need for studies in schools, where class sizes are larger and resources more limited. Future research should test short, repeatable routines that can realistically fit into weekly lessons.

Another gap concerns prosody. While segmental errors have been widely studied, suprasegmentals such as stress, rhythm, and intonation remain under-researched, even though they are critical for comprehensibility. Mobile tasks that target these features in simple and engaging ways should be designed and tested in the Bangladeshi context [33]. Feedback models also require more careful evaluation. Studies should not only measure whether ASR, teacher feedback, or hybrid approaches improve pronunciation but also assess how transparent and understandable the feedback is for learners and teachers. Mixed-method designs that combine quantitative results with classroom observations and interviews would help explain why certain approaches succeed or fail [38] [34].

Finally, research should consider long-term durability. Many interventions show short-term gains, but little is known about whether improvements in pronunciation are maintained over time. Follow-up assessments, even simple ones, can help answer this question. Reporting should also include detailed descriptions of tasks, feedback types, assessment methods, and resource demands, so that other teachers and researchers can replicate or adapt the interventions.

### **The Broader Lesson: Designing for Conditions**

The overall lesson from the evidence is that MAPI in Bangladesh must be designed for conditions, not for tools. Access to devices and apps alone does not guarantee learning. What matters is whether tasks are short and repeatable, feedback is transparent and actionable, teachers are supported to integrate practice into lessons, and equity is ensured through low-cost and offline delivery.

This conclusion echoes broader patterns in English language teaching in Bangladesh. Past innovations, such as communicative language teaching, often struggled because they did not match local realities of class size, exams, and teacher preparation [14]. MAPI will face the same risks if it is introduced without careful alignment to context. By focusing on conditions—pedagogy, feedback, assessment, equity—Bangladesh can turn widespread mobile use into meaningful gains in pronunciation and communication.

### **Implications for Practice**

The evidence from this review suggests that mobile-assisted pronunciation instruction (MAPI) can help Bangladeshi learners if it is designed with classroom realities, teacher needs, and learner conditions in mind. The practical challenge is not only about providing tools but about ensuring that these tools are used in ways that are feasible, fair, and pedagogically meaningful. Below, several practical implications are highlighted, focusing on lesson design, feedback practices, assessment, task sequencing, teacher support, and equity.

#### **Embedding MAPI into Lesson Routines**

Experience with communicative language teaching in Bangladesh has shown that new methods fail when they are not aligned with assessment systems, class time, and teacher preparation [14]. For this reason, pronunciation instruction through mobile tools must not be treated as an extra activity but as part of weekly teaching routines. A practical structure involves three steps: (i) explicit pre-teaching of a sound or feature, (ii) short mobile micro-practice that learners complete individually or in pairs, and (iii) in-class consolidation through guided repetition or a communicative task.

Such routines can be delivered in less than ten minutes and can become a predictable part of weekly lessons. When students know that pronunciation will always be practiced and checked, they are more motivated to engage. The English in Action programme demonstrated that even low-cost mobile content can sustain classroom change when materials, training, and logistics are planned together [39]. For MAPI, the lesson is clear: pronunciation practice must be embedded into teaching routines to ensure continuity and visibility.

#### **Making Feedback Transparent and Actionable**

In MAPI, transparent feedback means learners can see and understand three things: clear scores linked to simple goals, visual waveforms (or spectrograms) that show where sounds match or differ, and specific error labels (sound, stress, timing) with next steps. This clarity shows goals, progress, and what to do next, improving self-regulation and teaching fit [15] [35] [25].

Feedback is central to pronunciation learning. International studies show that learners improve most when feedback is explicit, immediate, and easy to interpret [21] [33]. This has two practical implications for classrooms.

First, when ASR is used, teachers should explain clearly what scores, colors, or waveforms mean. Students must know whether a low score reflects a specific sound problem, rhythm, or stress, and they must be guided on how to retry. Without this, learners may feel confused and lose trust in the tool.

ASR systems often show accent sensitivity: error rates rise for L2 speakers, regional dialects, and racialized varieties due to training-data imbalance and acoustic-phonetic mismatch [17] [36]. Environmental noise—classroom chatter, traffic, reverberation, low-quality mics—distorts spectra/timing and degrades recognition, especially for nonstandard accents [40] [29].

Second, when ASR is not available, teachers can still make feedback actionable. Instead of simply repeating a model, they can identify one error, explain it in simple terms, and model the correction. A short cycle of “model → attempt → correction → retry” can help learners adjust quickly. Evidence shows that teacher-led feedback can be as effective as automated systems when it is specific and consistent [34] [38].

#### **Aligning Assessment with Learning Goals**

Assessment also needs to match pronunciation constructs. In many Bangladeshi contexts, speaking tests are broad and do not show whether pronunciation has improved. International guidance recommends focusing on comprehensibility—how easy the speech is to understand—rather than quick changes in accentedness [5] [28].

Teachers can use short, low-burden methods: a simple rubric for reading aloud, a quick rating of intelligibility by peers, or a pre/post probe on one sound feature. If ASR tools are available, scores can be linked with classroom rubrics to make them more meaningful. These small steps ensure that progress is visible to both teachers and learners, which increases motivation and provides evidence of learning.

#### **Designing Tasks for Local Phonology and Bandwidth**

Tasks should reflect the predictable difficulties that Bangladeshi learners face. Studies highlight problems with aspiration contrasts (/p<sup>h</sup>, t<sup>h</sup>, k<sup>h</sup>/), vowel length and quality, and syllable structure [16] [22]. Early practice cycles should focus on these high-yield targets. Once learners gain some confidence, short prosody tasks—such as practicing stress on content

words, clapping to rhythm, or repeating intonation patterns—should be added. International studies show that prosody has a strong effect on comprehensibility but remains under-taught in both local and global contexts [33].

Equally important is adapting tasks to bandwidth realities. In Bangladesh, many students use shared phones, pay high data costs, or experience unstable connections [13] [30]. For this reason, tasks should be low-data or offline-capable. Small downloadable audio packs, lightweight apps, and exercises that work without continuous internet access are more inclusive. By designing for these realities, MAPI can reach a wider group of learners.

#### **Supporting Teachers as Key Mediators**

Teachers are central to sustaining mobile practice. Evidence from Bangladesh and similar contexts shows that pronunciation learning is most effective when teachers introduce, monitor, and consolidate mobile tasks within class time [39]. For this to happen, teachers need simple training and ready-made materials.

Professional development should focus on three main areas:

**(i) How to teach pronunciation explicitly** – giving teachers the skills to model sounds and explain differences between Bangla and English.

**(ii) How to use feedback effectively** – showing teachers how to interpret ASR outputs or how to deliver clear corrections when ASR is not available.

**(iii) How to design light routines** – training teachers to include short pronunciation cycles in lessons without disrupting other syllabus requirements.

By supporting teachers in these ways, MAPI can become a sustainable part of classroom practice rather than an optional add-on.

#### **Ensuring Equity and Transparency**

Finally, equity must guide every design choice. In South Asia, affordability, coverage, and digital skills remain the main barriers to mobile learning [13]. In Bangladesh, these challenges are joined by device sharing, charging problems, and small storage capacity [30]. To avoid widening inequalities, materials should be short, offline-capable, and downloadable in small parts.

Localization is also essential. Feedback should address predictable transfer problems in ways that learners can understand. For example, a simple note like “add a puff of air for /p<sup>h</sup>/” can be more effective than a technical explanation. Ethical transparency is also part of equity. Learners should know when their speech is recorded, how scores are produced, and how to use them. Teachers should be confident to question incorrect scores and provide human corrections when needed [34].

## CONCLUSION

This review has shown that mobile-assisted pronunciation instruction (MAPI) in Bangladesh has clear promise but also clear challenges. The key insight is that technology alone does not guarantee learning. Mobile tools can increase practice opportunities and motivate learners, but benefits are fragile when access is costly, shared, or unreliable [13] [30]. The success of MAPI depends on how well practice, feedback, and teaching routines are aligned with local conditions. Across the evidence, three themes stand out. First, feedback is central. Listening and repeating model recordings raise awareness, but they do not always help learners know what is wrong. International research shows that explicit, immediate, and understandable feedback—whether from automated speech recognition (ASR) or from teachers—is much more effective [21] [33] [38] [34]. Without such guidance, practice quickly becomes repetition without improvement. Second, task design and teacher mediation make the difference between success and failure. Short, structured cycles—input, practice, feedback, retry—are more effective than long or irregular drills. Teachers play a crucial role in introducing these cycles, monitoring progress, and consolidating learning in class. When pronunciation is integrated into weekly lessons, it becomes visible and sustainable. When it is treated as an add-on, it disappears under syllabus pressure [14] [39].

Third, equity and localization must shape every decision. Learners in Bangladesh face barriers such as high data costs, shared devices, and limited digital skills. Tools that assume constant connectivity risk excluding many students. Low-data, offline-capable materials, small downloads, and localized feedback that addresses Bangla-English differences are essential for fairness and effectiveness [16] [22] [30]. The broader lesson is that MAPI should be designed for conditions, not for tools. Bangladesh does not need complex applications to improve pronunciation. What is needed are short, frequent, feedback-rich routines, linked to classroom teaching, supported by teachers, and accessible to learners with limited connectivity. Under these conditions, even modest interventions can produce durable gains in comprehensibility and participation..

## REFERENCES

[1] Bangladesh Telecommunication Regulatory Commission. *Mobile subscribers (monthly table)*. (2025) <https://btrc.gov.bd/site/page/0ae188ae-146e-465c-8ed8-d76b7947b5dd/Mobile-Subscribers>

[2] Begum, R. Prospect for cell phones as instructional tools in the EFL classroom: A case study of Jahangirnagar University, Bangladesh. *English Language Teaching*, 4(1), 105–115. (2011) <https://doi.org/10.5539/elt.v4n1p105>

[3] Daily Asian Age. AI-powered international English learning app ‘ELSA Speak’ officially launches in Bangladesh. (2025, July 25) <https://dailyasianage.com/news/340550/ai-powered-international-english-learning-app-elsa-speak-officially-launches-in-bangladesh>

[4] DataReportal. *Digital 2025: Bangladesh*. (2025) <https://datareportal.com/reports/digital-2025-bangladesh>

[5] Derwing, T. M., & Munro, M. J. Accent, intelligibility, and comprehensibility: Evidence from four L1s. *Studies in Second Language Acquisition*, 19(1), 1–16. (1997) <https://doi.org/10.1017/S0272263197001010>

[6] Derwing, T. M., & Munro, M. J. Second language accent and pronunciation teaching: A research-based approach. *TESOL Quarterly*, 39(3), 379–397. (2005) <https://doi.org/10.2307/3588486>

[7] The Daily Star. Parlo, an international English learning app launched in Bangladesh. (2024, November 29) <https://www.thedailystar.net/business/organisation-news/press-releases/news/parlo-international-english-learning-app-launched-bangladesh-3764411>

[8] The Business Standard. Mobile internet subscription drops to 20-month low. (2025a, March 9) <https://www.tbsnews.net/bangladesh/telecom/mobile-internet-subscription-drops-20-month-low-1088541>

[9] The Business Standard. Over 50% of Bangladeshi households are now internet users: BBS. (2025b, January 1) <https://www.tbsnews.net/bangladesh/over-50-households-use-internet-bangladesh-bbs-1032411>

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

[11] Fouz-González, J. Using apps for pronunciation training: An empirical evaluation of the English File Pronunciation app. *Language Learning & Technology*, 24(1), 62–85. (2020) <https://doi.org/10.64152/10125/44709>

[12] GSMA. *Accelerating digital transformation in Bangladesh*. (2021) [https://www.gsma.com/solutions-and-impact/connectivity-for-good/public-policy/wp-content/uploads/2021/05/GSMA\\_Accelerating\\_digital\\_transformation\\_in\\_Bangladesh\\_2021.pdf](https://www.gsma.com/solutions-and-impact/connectivity-for-good/public-policy/wp-content/uploads/2021/05/GSMA_Accelerating_digital_transformation_in_Bangladesh_2021.pdf)

[13] GSMA. *The state of mobile internet connectivity 2023*. (2023) <https://www.gsma.com/r/somic-2023/>

[14] Hamid, M. O., & Baldauf, R. B. Will CLT bail out the bogged down ELT in Bangladesh? *English Today*, 24(3), 16–24. (2008) <https://doi.org/10.1017/S0266078408000254>

[15] Hattie, J., & Timperley, H. The power of feedback. *Review of Educational Research*, 77(1), 81–112. (2007) <https://doi.org/10.3102/003465430298487>

- [16] Jakir, M. Pronunciation of the aspirated English sounds by Bangladeshi EFL/ESL learners: An empirical study. *Jahangirnagar University Journal of Humanities*, 10(1), 250–271. (2020)[https://www.jnu.ac.bd/journal/assets/pdf/10\\_1\\_250.pdf](https://www.jnu.ac.bd/journal/assets/pdf/10_1_250.pdf)
- [17] Koenecke, A., Nam, A., Lake, E., Nudell, J., Quartey, M., Mengesha, Z., Toups, C., Rickford, J. R., Jurafsky, D., & Goel, S. Racial disparities in automated speech recognition. *Proceedings of the National Academy of Sciences*, 117(14), 7684–7689. (2020) <https://doi.org/10.1073/pnas.1915768117>
- [18] Kukulska-Hulme, A. Mobile-assisted language learning. In C. A. Chapelle (Ed.), *The concise encyclopedia of applied linguistics* (pp. 665–672). Wiley. (2018) <https://oro.open.ac.uk/57023/>
- [19] Lee, J., Jang, J., & Plonsky, L. The effectiveness of second language pronunciation instruction: A meta-analysis. *Applied Linguistics*, 36(3), 345–366. (2015) <https://doi.org/10.1093/applin/amu040>
- [20] Maniruzzaman, M. Phonetic and phonological problems encountered by the Bengali-speaking EFL learner: How can they be overcome? *JU Review (Part-C)*. (2006) <https://www.researchgate.net/publication/317061468>
- [21] Metruk, R. Mobile-assisted language learning and pronunciation instruction: A systematic literature review. *Education and Information Technologies*, 29, 16255–16282. (2024) <https://doi.org/10.1007/s10639-024-12453-0>
- [22] Mostafa, T. A contrastive analysis between Bangla and English phonology: Some pedagogical recommendations. In *Proceedings of the European Conference on Language Learning 2013* (pp. 1–12). (2013)[https://papers.iafor.org/wp-content/uploads/papers/ecll2013/ECLL2013\\_0162.pdf](https://papers.iafor.org/wp-content/uploads/papers/ecll2013/ECLL2013_0162.pdf)
- [23] Mumu, S. S. *Mobile-assisted language learning in secondary level schools in Bangladesh* (Undergraduate thesis, BRAC University). (2021). <https://dspace.bracu.ac.bd/handle/10361/16618>
- [24] Munro, M. J., & Derwing, T. M. Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, 45(1), 73–97. (1995) <https://doi.org/10.1111/j.1467-1770.1995.tb00963.x><https://oro.open.ac.uk/32444/>
- [25] Neri, A., Cucchiari, C., & Strik, H. Feedback in computer assisted pronunciation training: Technology push or demand pull? *Proceedings of ICSLP 2002*, 1209–1212. (2002) <https://doi.org/10.21437/ICSLP.2002-246>
- [26] Ngo, T. T.-N., Chen, H. H.-J., & Lai, K. K.-W. The effectiveness of automatic speech recognition in ESL/EFL pronunciation: A meta-analysis. *ReCALL*, 36(1), 4–21. (2024) <https://doi.org/10.1017/S0958344023000113>
- [27] Qader, M. A., Ashraf, T. A., Monira, S., Islam, M. A., & Rahman, M. R. M. Challenges of pronunciation practices in the ESL curriculum within the CLT framework in Bangladesh: A systematic review. *World Journal of Education*, 14(3), 137–144. (2024) <https://doi.org/10.5430/wje.v14n3p137>
- [28] Saito, K., & Plonsky, L. Effects of second language pronunciation teaching revisited: A proposed measurement framework and meta-analysis. *Language Learning*, 69(3), 652–708. (2019). <https://doi.org/10.1111/lang.12345>
- [29] Seltzer, M. L., Yu, D., & Wang, Y. An investigation of deep neural networks for noise robust speech recognition. *2013 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 7398–7402. IEEE. (2013) <https://doi.org/10.1109/ICASSP.2013.6639064>
- [30] Ibna Seraj, P. M. I., Klimova, B., & Habil, H. Use of mobile phones in teaching English in Bangladesh: A systematic review (2010–2020). *Sustainability*, 13(10), 5674. (2021) <https://doi.org/10.3390/su13105674>
- [31] Ibna Seraj, P. M., Habil, H., & Hasan, M. K. Investigating EFL learners’ readiness of using smartphone and problems for learning oral English communication skills at tertiary level in an EFL context. *International Journal of Interactive Mobile Technologies*, 14(14), 148–163. (2020) <https://doi.org/10.3991/ijim.v14i14.14759>
- [32] Ibna Seraj, P. M., & Klimova, B. Smartphone apps-based teaching method to develop oral English communication skills at the tertiary level in an EFL context. *World Journal of English Language*, 15(2), 250–262. (2025). <https://doi.org/10.5430/wjel.v15n2p250>
- [33] Stoughton, A. M., & Kang, O. A systematic review of empirical mobile-assisted pronunciation studies through a perception–production lens. *Languages*, 9(7), 251. (2024) <https://doi.org/10.3390/languages9070251>
- [34] Sun, W. 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, 1210187. (2023) <https://doi.org/10.3389/fpsyg.2023.1210187>
- [35] Shute, V. J. Focus on formative feedback. *Review of Educational Research*, 78(1), 153–189. (2008) <https://doi.org/10.3102/0034654307313795>
- [36] Tatman, R., & Kasten, C. Effects of talker dialect, gender & race on accuracy of Bing Speech and YouTube automatic captions. *Interspeech 2017*, 934–938. (2017) <https://doi.org/10.21437/Interspeech.2017-1746>
- [37] Trofimovich, P., & Isaacs, T. Disentangling accent from comprehensibility. *Bilingualism: Language and Cognition*, 15(4), 905–916. (2012) <https://doi.org/10.1017/S1366728912000168>
- [38] Tseng, W.-T., Chen, S., Wang, S.-P., Cheng, H.-F., Yang, P.-S., & Gao, X. A. The effects of MALL on L2 pronunciation learning: A meta-analysis. *Journal of Educational Computing Research*, 60(6), 1492–1520. (2022) <https://doi.org/10.1177/07356331211058662>

- [39] Walsh, C. S., Shrestha, P., & Hedges, C. Mobile phones for professional development and English teaching in Bangladesh. *International Journal of Innovation and Learning*, 13(2), 183–200. (2013) <https://doi.org/10.1504/IJIL.2013.052287>
- [40] Wassink, A. B., Gansen, C., & Bartholomew, I. Uneven success: Automatic speech recognition and ethnicity-related dialects. *Speech Communication*, 140, 50–70. (2022)<https://doi.org/10.1016/j.specom.2022.03.009>