



# ICT Tools for Teaching Speaking Skills: A Comprehensive Pedagogical Review

Dr. Gollapalli Shalom, Assistant Professor, Dept. of BS&HSS, Andhra University College of Engineering for Women. ORCID: 0000-0001-9898-1078

## Abstract

The integration of Information and Communication Technology (ICT) into English language teaching has fundamentally transformed the pedagogical landscape for developing speaking skills, particularly in contexts where learners face systemic barriers to oral proficiency. This paper presents a comprehensive review of prominent ICT tools — encompassing podcasting platforms, video-conferencing applications, language learning apps, artificial intelligence-powered pronunciation coaches, digital storytelling environments, and gamified speaking tools — and evaluates their pedagogical efficacy within established theoretical frameworks, including Communicative Language Teaching (CLT), the Common European Framework of Reference for Languages (CEFR), and Vygotsky's Zone of Proximal Development. Drawing on empirical studies and classroom-based research, the paper argues that strategically deployed ICT tools can substantially reduce affective barriers to spoken communication, promote authentic language use, and accelerate proficiency gains. Furthermore, the paper highlights the critical need to combine ICT integration with differentiated instructional approaches to address the linguistic heterogeneity characteristic of contemporary classrooms. Challenges including digital equity, teacher preparedness, and the risk of decontextualised technology use are critically examined. The paper concludes with implications for practitioners, curriculum designers, and policy makers committed to leveraging technology for meaningful oral language development.

**Keywords:** ICT tools, speaking skills, language teaching, technology-enhanced learning, oral communication, CEFR, communicative competence

## 1. Introduction

Among the four macro-skills of language learning — listening, speaking, reading, and writing — oral communication occupies a position of primary importance in academic, professional, and social life. The ability to communicate effectively in spoken English is increasingly regarded as a prerequisite for participation in globalized economies and knowledge societies. As Gollapalli Shalom (2026) observes, speaking skills are the most sought-after competencies in a person's life, and professional, academic, and personal growth all fundamentally depend upon effective communication. Despite this acknowledged centrality, speaking remains arguably the most neglected skill in formal language education, particularly in the Indian subcontinent, where examination-focused pedagogies have historically privileged reading and writing at the expense of oral fluency (Rajagopalan, 2004).

The widespread integration of Information and Communication Technology (ICT) into education has opened new possibilities for addressing this pedagogical gap. Digital tools and platforms offer learners unprecedented access to authentic spoken language, interactive practice environments, and immediate formative feedback — elements that are difficult to replicate within the constraints of large, resource-limited classrooms. Over the past two decades, a substantial body of research has documented the potential of technology-enhanced language learning (TELL) to foster oral proficiency, reduce language anxiety, and motivate continued engagement with the target language (Warschauer & Kern, 2000; Chapelle, 2003).

However, ICT integration does not automatically translate into improved speaking outcomes. The effectiveness of any digital tool is contingent upon thoughtful pedagogical design, alignment with learner proficiency levels, and the capacity of teachers to embed technology within broader communicative frameworks. This paper undertakes a systematic review of the major ICT categories currently employed in teaching speaking skills, evaluates their pedagogical underpinnings, and situates them within current debates in applied linguistics and educational technology. The paper also underscores that sustainable gains in oral proficiency require ICT to be embedded within differentiated instructional frameworks that account for individual learner readiness, interests, and learning profiles (Tomlinson, 2014).

## 2. Theoretical Framework

### *2.1 Communicative Language Teaching*

The theoretical bedrock of speaking instruction in contemporary applied linguistics is Communicative Language Teaching (CLT), which emerged in the 1970s as a reaction against structuralist, grammar-translation approaches. CLT posits that the primary goal of language learning is the development of communicative competence — the ability to use language appropriately and effectively in real-world contexts (Hymes, 1972; Canale & Swain, 1980). Rather than treating accuracy as the sole metric of success, CLT prioritizes fluency, pragmatic appropriateness, and the negotiation of meaning in interaction.

Within this framework, speaking activities must simulate authentic communicative contexts, provide opportunities for spontaneous language production, and incorporate meaningful feedback.

### ***2.2 Zone of Proximal Development and Scaffolding***

Vygotsky's (1978) concept of the Zone of Proximal Development (ZPD) — defined as the distance between what a learner can accomplish independently and what can be achieved through guided interaction with a more capable peer or instructor — provides an essential lens for understanding how ICT tools can function as mediating scaffolds. Digital platforms, when well-designed, can continuously calibrate task difficulty to maintain learners within their ZPD, presenting speaking challenges that are neither overwhelming nor insufficiently stimulating. This dynamic scaffolding function is particularly visible in AI-powered speaking tools and adaptive language learning applications.

### ***2.3 The CEFR and Oral Proficiency Benchmarks***

The Common European Framework of Reference for Languages (Council of Europe, 2001) provides a widely adopted set of descriptors for oral proficiency, ranging from A1 (breakthrough) to C2 (mastery). CEFR-aligned speaking descriptors encompass interaction, monologue, and spoken production across domains of everyday communication, academic discourse, and professional interaction. ICT tools increasingly incorporate CEFR-referenced proficiency goals, enabling learners to identify their current level and chart a course of improvement. The empirical significance of aligning instruction with CEFR benchmarks is underscored by Gollapalli Shalom (2026), whose classroom-based study demonstrated that systematic, technology-supported instruction produced statistically significant gains in CEFR proficiency levels among undergraduate engineering students, with measurable upward mobility across proficiency bands.

## **3. Overview of ICT in Language Education**

The term ICT in language education encompasses a wide spectrum of hardware, software, and networked platforms. Early CALL (Computer Assisted Language Learning) research in the 1980s and 1990s focused primarily on discrete-point grammar and vocabulary exercises delivered via standalone computers (Levy, 1997). The advent of the internet, followed by the proliferation of smartphones, broadband connectivity, and cloud computing, dramatically expanded the affordances available to language teachers and learners. Contemporary ICT resources range from social media and streaming video to sophisticated artificial intelligence applications capable of real-time speech recognition, phonemic analysis, and natural language generation.

Warschauer (1996) proposed an influential tripartite model of CALL evolution: structural CALL (emphasising drill and practice), communicative CALL (encouraging meaning-focused interaction), and integrative CALL (embedding technology in authentic, learner-centred tasks). Contemporary ICT use in speaking instruction operates predominantly within the communicative and integrative phases, prioritising

interaction, negotiation of meaning, and authentic output over mechanical repetition. Research consistently demonstrates that technology-enhanced speaking activities produce favourable affective outcomes — lower language anxiety, higher motivation, greater willingness to communicate — alongside measurable gains in oral proficiency (Stockwell, 2012; Shadiev et al., 2017).

## **4. ICT Tools for Teaching Speaking Skills**

### ***4.1 Podcasting and Digital Audio Platforms***

Podcasting emerged in the mid-2000s as an accessible medium for learner-generated audio content, and its applications in speaking instruction have been extensively documented. Platforms such as Audacity, GarageBand, Anchor, and SoundCloud enable learners to record, edit, and publish spoken monologues, dialogues, interviews, and radio-style programmes. The act of preparing and recording a podcast requires learners to engage in careful planning, drafting, oral rehearsal, and self-evaluation — processes that mirror the metacognitive demands of authentic spoken communication.

Rosell-Aguilar (2007) reviewed podcasting in language education and identified three principal modes: listening to teacher-produced or professionally produced podcasts to develop aural models of pronunciation and discourse; creating learner-generated podcasts as a form of spoken output; and engaging in podcast exchange projects with international partner classes to foster intercultural communicative competence. Research by Ducate and Lomicka (2009) found that learner podcasters demonstrated improvements in pronunciation, fluency, and oral organization after a semester-long project, with participants reporting increased confidence and reduced anxiety. Importantly, the asynchronous nature of podcasting affords learners time to draft, revise, and re-record their contributions, mitigating the real-time performance pressure that many speakers find inhibiting.

### ***4.2 Video Conferencing and Virtual Communication Platforms***

The COVID-19 pandemic accelerated the adoption of video conferencing platforms — Zoom, Microsoft Teams, Google Meet, and Cisco Webex — as primary vehicles for spoken interaction in language classrooms. These tools replicate many dimensions of face-to-face communication, including visual paralinguistic cues, turn-taking, and spontaneous repair strategies, while offering additional affordances such as breakout rooms for small-group work, recording functions, and integrated chat. Research conducted during the pandemic transition period found that synchronous video-mediated communication maintained learner engagement and preserved opportunities for authentic interaction, though researchers noted challenges related to reduced gesture visibility and the suppression of backchanneling cues (Hampel & Stickler, 2005; Moser et al., 2021).

Beyond institutional video conferencing, telecollaborative platforms such as Tandem, HelloTalk, and Speaky facilitate language exchange partnerships between learners in different countries. These applications operationalize a key tenet of CLT — that authentic communicative need is the most powerful

driver of language development. Learners engaged in tandem exchanges have access to native-speaker input, meaningful interactional goals, and implicit corrective feedback, replicating conditions that classroom instruction alone cannot reliably provide (O'Dowd, 2018). Empirical studies on telecollaboration consistently report gains in oral fluency, pragmatic awareness, and intercultural competence (Kohn & Warth, 2011).

#### ***4.3 Language Learning Applications***

The proliferation of mobile language learning applications — most prominently Duolingo, Babbel, Rosetta Stone, and Pimsleur — has extended access to speaking practice beyond the classroom into learners' everyday environments. These platforms leverage the affordances of smartphones, including ubiquitous access, push notification systems, and gamified reward structures, to encourage frequent, distributed practice. From a pedagogical standpoint, the most valuable speaking features of these applications include oral repetition tasks, speech-recognition-evaluated pronunciation exercises, and structured dialogue simulations.

Research on the effectiveness of mobile applications for speaking development presents a nuanced picture. Vesselinov and Grego (2012) found that Rosetta Stone users achieved significant gains in listening and speaking proficiency following structured use. However, critics note that the controlled, discrete-point speaking tasks embedded in most applications fail to develop the spontaneous, pragmatically sensitive oral competence demanded by real-world communication (Godwin-Jones, 2011). Consequently, mobile applications are most productively viewed as supplementary tools for building phonemic accuracy and vocabulary for spoken use, rather than as standalone substitutes for communicative speaking practice.

#### ***4.4 AI-Powered Speech and Pronunciation Tools***

Among the most technically sophisticated ICT resources available to speaking teachers, AI-powered tools represent a frontier of growing pedagogical interest. Applications such as ELSA Speak (English Language Speech Assistant), Speechling, Google's Speech Recognition API, and Carnegie Speech's NativeAccent employ advanced automatic speech recognition (ASR) and neural acoustic modelling to provide learners with phoneme-level feedback on pronunciation, intonation, rhythm, and stress patterns. ELSA Speak, for instance, uses deep learning to identify specific phonemic errors and presents targeted remediation exercises, generating individualized practice programmes calibrated to each learner's identified weakness profile.

The pedagogical value of AI-powered pronunciation tools rests substantially on their capacity to deliver the kind of immediate, specific, and actionable feedback that Hattie and Timperley (2007) identified as the most powerful driver of learning gain. Traditional classroom settings rarely permit the level of individualized phonetic coaching that AI tools can offer; a teacher managing thirty or sixty students cannot provide sustained phoneme-by-phoneme corrective feedback to each learner. AI tools fill this gap at scale,

functioning as personalised pronunciation tutors available around the clock. Research by Neri et al. (2008) demonstrated that ASR-based pronunciation training produced gains equivalent to or exceeding those achieved through human-led pronunciation instruction in controlled experimental conditions, though researchers cautioned against overreliance on automated feedback at the expense of pragmatic and discursive dimensions of spoken language.

Generative AI tools, including large language models such as GPT-4 and conversational agents built on these architectures, have recently opened additional possibilities for speaking instruction. Chatbot-mediated conversation practice allows learners to engage in extended spoken or text-based dialogues on topics of interest, with the AI serving as a patient, non-judgmental interlocutor capable of adapting register, complexity, and topic focus to learner needs. Preliminary research on AI conversational partners in EFL contexts has reported positive effects on willingness to communicate, speaking confidence, and lexical breadth (Huang et al., 2022).

#### ***4.5 Digital Storytelling Tools***

Digital storytelling — the practice of creating multimodal narratives combining spoken narration with images, video, music, and text — has gained considerable traction as a pedagogically rich speaking activity. Platforms such as Adobe Spark, WeVideo, Storybird, and VoiceThread provide accessible interfaces for producing digital stories and enable learners to narrate personally meaningful content, thereby capitalising on the motivational benefits of self-selected topics and authentic communicative purpose. The multimodal scaffolding provided by images and music reduces the cognitive burden on speakers, allowing attention to be devoted to oral fluency and expressive delivery rather than memory retrieval.

Robin (2006) identified digital storytelling as a powerful vehicle for developing multiple language competencies simultaneously, including oral fluency, narrative structuring, vocabulary for spoken description, and prosodic control. Studies conducted in EFL contexts in Asia and the Middle East have documented improvements in speaking confidence, narrative coherence, and pronunciation clarity among learners who engaged in sustained digital storytelling projects (Abdel-Hack & Helwa, 2014). Crucially, digital storytelling foregrounds the social dimension of spoken language by producing artefacts intended for an audience, thus embedding oral production in a genuinely communicative context aligned with CLT principles.

#### ***4.6 Gamification and Game-Based Speaking Tools***

Gamified learning environments and game-based applications represent a further category of ICT with established applications in speaking instruction. Platforms such as Flipgrid (now Flip by Microsoft), Kahoot, Quizlet Live, and purpose-designed language games embed speaking tasks within game mechanics — points, leaderboards, collaborative challenges, and narrative progression — that sustain

learner motivation and reduce the social vulnerability associated with public oral performance. Flipgrid, in particular, has attracted substantial pedagogical interest for its capacity to support asynchronous video responses to teacher-posed discussion prompts, enabling reflective, low-stakes speaking practice with peer audience engagement.

Research on gamification in language learning suggests that competitive and collaborative game structures increase intrinsic motivation, prolong voluntary engagement with speaking tasks, and generate repeated cycles of oral production — a frequency of practice that is associated with fluency development (Yildirim, 2017). Deterding et al. (2011) conceptualised gamification as the application of game design elements in non-game contexts to drive engagement, and this principle translates productively into speaking pedagogy when game mechanics are aligned with substantive communicative objectives rather than employed as superficial motivational devices.

## 5. Pedagogical Implications

The review of ICT tools presented above reveals a consistent pattern: technology is most effective in developing speaking skills when it is embedded within coherent communicative pedagogical frameworks rather than deployed as an isolated supplement to conventional instruction. Several cross-cutting implications emerge from the evidence.

First, effective ICT integration for speaking must account for learner heterogeneity. Contemporary classrooms — particularly in Indian higher education — routinely accommodate students spanning multiple CEFR proficiency levels, diverse linguistic backgrounds, and varied learning styles. Gollapalli Shalom (2026) demonstrates that in heterogeneous classrooms, a one-size-fits-all instructional approach leaves fast learners under-challenged while slow learners remain unsupported; the solution proposed and empirically validated in that study is a combination of differentiated instruction with purposefully selected educational technology. This finding has direct implications for speaking instruction: ICT tools should be selected and sequenced to match learners' individual readiness profiles, with adaptive platforms, tiered digital tasks, and choice-based speaking activities enabling genuine differentiation.

Second, the affective dimension of speaking development demands sustained attention. Language anxiety — the apprehension specifically associated with speaking in a foreign language — is among the most robustly documented inhibitors of oral production (Horwitz et al., 1986). ICT tools, by creating low-stakes, private or semi-private speaking environments, can substantially reduce anxiety and increase willingness to communicate. Asynchronous platforms (podcasting, video journals, Flipgrid) afford learners the time and psychological safety to rehearse and refine their spoken output before committing it to a permanent or semi-permanent record, thereby attenuating the vulnerability of real-time performance.

Third, feedback quality is critical. The most productive ICT tools for speaking provide specific, timely, and actionable feedback — characteristics identified by Hattie and Timperley (2007) as the hallmarks of effective formative assessment. AI-powered pronunciation tools excel on this dimension, offering phoneme-level precision unavailable through typical classroom feedback. Teachers should critically evaluate the feedback mechanisms of any ICT tool before deployment, privileging those that move beyond binary correct/incorrect judgments toward explanatory, generative feedback that helps learners understand the nature of their errors and the strategies available for remediation.

Fourth, ICT use must be pedagogically purposeful. Research warns against what Selwyn (2016) termed "technicism" — the uncritical adoption of technology as an end in itself rather than a means to defined learning goals. The most effective ICT interventions for speaking are those in which the choice of tool is driven by a clear pedagogical rationale, integrated into a sequence of authentic communicative tasks, and evaluated systematically against measurable proficiency benchmarks.

## **6. Challenges in ICT Integration for Speaking Instruction**

Despite the considerable promise documented in the literature, ICT integration for speaking instruction faces a range of persistent challenges that require critical acknowledgment.

Digital equity constitutes a fundamental structural barrier. Access to the hardware, software, and reliable internet connectivity required for technology-enhanced speaking tasks is unevenly distributed across socioeconomic groups, geographic regions, and institutional types. In many Indian contexts, bandwidth limitations, inadequate device availability, and unreliable electricity supply impose severe constraints on the feasibility of synchronous video-mediated communication or data-intensive AI tools. Pedagogical recommendations for ICT integration must therefore be sensitive to infrastructural realities and should include provisions for asynchronous, lower-bandwidth alternatives where necessary (Bansode & Pujar, 2016).

Teacher preparedness represents a second critical challenge. The effective pedagogical deployment of ICT tools for speaking requires teachers to possess not only technical proficiency with the relevant platforms but also the capacity to align digital tools with communicative learning objectives, scaffold technology-mediated tasks appropriately, and critically evaluate learner output generated through digital means. Professional development programmes in educational technology for speaking instruction remain unevenly distributed and poorly sustained in many institutional contexts, leading to superficial or pedagogically misaligned technology use (Bax, 2003).

The risk of form-focused reductionism constitutes a third concern. Many commercially available speaking tools — particularly pronunciation apps and grammar-focused chatbots — inadvertently reinforce an accuracy-at-the-expense-of-fluency orientation that CLT sought to overcome. Learners who experience

technology exclusively as a forum for phonemic correction may develop what Skehan (1998) called an accuracy focus that inhibits the risk-taking and spontaneous production necessary for fluency development. Teachers must therefore balance the deployment of accuracy-oriented tools with ample opportunities for unmonitored, meaning-focused oral interaction.

## 7. Conclusion

This paper has reviewed the principal categories of ICT tools currently available for teaching speaking skills and evaluated their pedagogical efficacy within established theoretical frameworks. The evidence consistently supports the view that technology-enhanced speaking instruction can produce meaningful gains in oral fluency, phonemic accuracy, communicative confidence, and learner motivation — provided that digital tools are embedded within communicative pedagogical frameworks, aligned with learner proficiency and differentiated to address individual learning needs, and evaluated against clearly articulated speaking objectives.

Looking ahead, the rapid development of generative AI, immersive virtual reality environments, and multimodal conversational agents is likely to further transform the landscape of technology-enhanced speaking instruction. These emerging tools hold particular promise for creating the authentic, low-stakes interactional contexts in which spoken fluency is most naturally developed. However, as the preceding analysis makes clear, no technology — however sophisticated — can replace the informed professional judgment of a teacher who understands the social, affective, and cognitive dimensions of spoken language development.

Ultimately, the most productive orientation toward ICT in speaking instruction is neither uncritical enthusiasm nor reflexive scepticism, but rather principled technological integration: a stance in which digital tools are chosen, adapted, and evaluated on the basis of robust pedagogical theory, empirical evidence, and an unflinching commitment to equitable, inclusive, and genuinely communicative language learning.

## References

1. Abdel-Hack, E. M., & Helwa, H. S. A. A. (2014). Using digital storytelling and weblogs instruction to enhance EFL narrative writing and critical thinking skills among EFL majors at faculty of education. *Educational Research*, 14(9), 8–41.
2. Bansode, S. Y., & Pujar, S. M. (2016). ICT skills among library professionals: A study of university library professionals in Karnataka. *DESIDOC Journal of Library and Information Technology*, 36(5), 282–287.
3. Bax, S. (2003). CALL — past, present and future. *System*, 31(1), 13–28.
4. Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, 1(1), 1–47.

5. Chapelle, C. A. (2003). English language learning and technology: Lectures on applied linguistics in the age of information and communication technology. John Benjamins.
6. Council of Europe. (2001). Common European framework of reference for languages: Learning, teaching, assessment. Cambridge University Press.
7. Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From game design elements to gamefulness: Defining gamification. *Proceedings of the 15th International Academic MindTrek Conference*, 9–15.
8. Ducate, L., & Lomicka, L. (2009). Podcasting: An effective tool for honing language students' pronunciation? *Language Learning and Technology*, 13(3), 66–86.
9. Godwin-Jones, R. (2011). Emerging technologies: Mobile apps for language learning. *Language Learning and Technology*, 15(2), 2–11.
10. Gollapalli Shalom. (2026). Inclusive learning: Employing differentiated instruction for CEFR success. *International Journal of Creative Research Thoughts (IJCRT)*, 14(4), b605–b618.
11. Hampel, R., & Stickler, U. (2005). New skills for new classrooms: Training tutors to teach languages online. *Computer Assisted Language Learning*, 18(4), 311–326.
12. Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112.
13. Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign language classroom anxiety. *The Modern Language Journal*, 70(2), 125–132.
14. Huang, X., Zou, D., Cheng, G., Chen, X., & Xie, H. (2022). Trends, research issues, and applications of artificial intelligence in language education. *Educational Technology and Society*, 24(3), 238–255.
15. Hymes, D. H. (1972). On communicative competence. In J. B. Pride & J. Holmes (Eds.), *Sociolinguistics: Selected readings* (pp. 269–293). Penguin.
16. Kohn, K., & Warth, C. (2011). Message competence and lingua franca communication. In A. De Houwer & A. Wilton (Eds.), *English in Europe today* (pp. 181–200). John Benjamins.
17. Levy, M. (1997). *Computer-assisted language learning: Context and conceptualisation*. Oxford University Press.
18. Moser, K. M., Wei, T., & Brenner, D. (2021). Remote teaching during COVID-19: Implications from a national survey of language educators. *System*, 97, 102431.
19. Neri, A., Mich, O., Gerosa, M., & Giuliani, D. (2008). The effectiveness of computer-assisted pronunciation training for foreign language learning by children. *Computer Assisted Language Learning*, 21(5), 393–408.
20. O'Dowd, R. (2018). From telecollaboration to virtual exchange: State-of-the-art and the role of UNICollaboration in moving forward. *Journal of Virtual Exchange*, 1, 1–23.
21. Rajagopalan, K. (2004). The concept of World English and its implications for ELT. *ELT Journal*, 58(2), 111–117.
22. Robin, B. R. (2006). The educational uses of digital storytelling. In C. Crawford et al. (Eds.), *Proceedings of Society for Information Technology and Teacher Education International Conference* (pp. 709–716). AACE.

23. Rosell-Aguilar, F. (2007). Top of the pods: In search of a podcasting "podagogy" for language learning. *CALICO Journal*, 25(1), 471–492.
24. Selwyn, N. (2016). *Is technology good for education?* Polity Press.
25. Shadiev, R., Hwang, W.-Y., & Liu, T.-Y. (2017). A study of the use of wearable devices for healthy and enjoyable English language learning in authentic contexts. *Educational Technology and Society*, 20(4), 217–231.
26. Skehan, P. (1998). *A cognitive approach to language learning*. Oxford University Press.
27. Stockwell, G. (2012). *Computer-assisted language learning: Diversity in research and practice*. Cambridge University Press.
28. Tomlinson, C. A. (2014). *The differentiated classroom: Responding to the needs of all learners* (2nd ed.). ASCD.
29. Vesselinov, R., & Grego, J. (2012). *Duolingo effectiveness study*. City University of New York.
30. Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.
31. Warschauer, M. (1996). Computer-assisted language learning: An introduction. In S. Fotos (Ed.), *Multimedia language teaching* (pp. 3–20). Logos International.
32. Warschauer, M., & Kern, R. (2000). *Network-based language teaching: Concepts and practice*. Cambridge University Press.
33. Yildirim, I. (2017). The effects of gamification-based teaching practices on student achievement and students' attitudes toward lessons. *Internet and Higher Education*, 33, 86–92.