

Digital Technologies and Virtual Environments in Language Learning

Orujlu Afsana Mirjalal

Lecturer, PhD Candidate, Nakhchivan State University, Nakhchivan, Azerbaijan

Email: afsanaoruclu@ndu.edu.az

ORCID: <https://orcid.org/0009-0008-9137-2964>

© 2026 Afsana Mirjalal Orujlu. This is an open-access article published in *Global Spectrum of Research and Humanities (GSRH)*. Licensed under **Creative Commons Attribution 4.0 International (CC BY 4.0)** — free to share, adapt and use commercially, with attribution.

<https://creativecommons.org/licenses/by/4.0/>

DOI: <https://doi.org/10.69760/gsrh.0260304028>

Abstract

Digital technologies can expand opportunities for language practice, feedback, and access to learning materials, but their educational value depends on how they are integrated into instruction. This conceptual article draws on a selective narrative review to examine automatic speech recognition, conversational agents, machine translation, and virtual environments in language learning. It distinguishes the functions of these technologies from the learning outcomes that teachers must evaluate. The discussion identifies opportunities for pronunciation practice, sustained interaction, reflective revision, and contextualized communication, alongside problems involving unreliable feedback, cognitive demands, unequal access, and personal data. Alpha School is considered briefly as an institutional illustration of adaptive learning and human guidance, rather than as evidence of foreign-language learning effectiveness. Four proposed classroom applications show how teachers can connect digital activity to observable language performance. The article argues that responsible integration requires clear learning objectives, appropriate scaffolding, human review of feedback, and opportunities to demonstrate learning beyond the original tool. Its contribution is a practical framework for evaluating instructional use; the proposed applications require subsequent classroom investigation.

Keywords: *language learning; digital technologies; automatic speech recognition; conversational agents; machine translation; virtual environments*

1. Introduction

Language learners can now encounter unfamiliar vocabulary, receive automated feedback, and practise conversations through digital systems outside scheduled lessons. These opportunities extend the learning environment, but access to a tool does not explain what a learner gains from using it. A translated paragraph may help a student understand a text while concealing weaknesses in independent reading. A chatbot may sustain an exchange without showing whether the student can manage the same interaction with another person. The educational question is therefore how digital activity contributes to identifiable language abilities.

Research provides reasons to investigate these possibilities while remaining attentive to differences between tools and tasks. Bibauw et al. (2022) report a positive overall effect of

dialogue-based computer-assisted language learning on second-language proficiency. Nickolai et al. (2024) identify evidence supporting automatic speech recognition in pronunciation instruction. Research on virtual environments also reports benefits, although the technologies, participants, and outcomes differ across studies (Hua & Wang, 2023). These findings support purposeful instructional experimentation; they do not establish that every application will work equally well for every learner.

This article examines the relationship between technological functions and pedagogical decisions. It addresses two questions: How can selected digital technologies support specific language-learning activities? What instructional and ethical conditions should guide their use? The central argument is that technology becomes educationally useful when teachers connect its functions to a learning objective, guide students' responses to feedback, and evaluate performance beyond simple task completion. The discussion retains the wider concerns of inclusion, privacy, and human interaction while concentrating on language education.

2. Scope and Review Approach

The article is a conceptual discussion informed by a selective narrative review. Targeted web searches used combinations of terms including automatic speech recognition, pronunciation learning, chatbots, language learning, machine translation, virtual reality, and educational artificial intelligence. Journal and publisher records, accessible research texts and abstracts, institutional repositories, and UNESCO publications were consulted. Sources were selected for direct relevance to language-learning processes or to the conditions governing educational technology use. Reviews and meta-analyses provide an overview of research, while an empirical pronunciation study illustrates learner behaviour during practice. Publication details were checked against source records.

This process was not an exhaustive database review, and the article makes no claim of systematic coverage or a pooled estimate of effectiveness. The selected research is primarily from 2021–2024, supplemented by a technical reference and an institutional website. Sources are interpreted according to their purpose: research syntheses inform the discussion of educational evidence, policy guidance informs ethical considerations, and Alpha School's website describes its own model. Findings from earlier systems are not treated as direct evaluations of every current generative tool.

The analysis groups technologies by the activities they support and examines the relationship between task design, feedback, learner action, and assessment. The classroom applications are proposed examples developed for this discussion. No classroom intervention, participant recruitment, or measurement of learning outcomes was conducted.

3. Digital Environments and Language-Learning Functions

A digital learning environment combines materials, communication channels, tools, and instructional arrangements. It may consist of a course platform and shared documents, a conversational application, or an immersive simulated setting. A virtual environment is therefore not necessarily a headset-based experience. In this article, immersive virtual reality refers specifically to systems that place users within a perceptually surrounding simulated environment; other digital activities may be conducted through an ordinary screen.

Language technologies also perform distinct operations. Automatic speech recognition converts speech into a textual representation; text-to-speech systems generate spoken output from text; machine translation produces a rendering in another language; and conversational systems generate or select responses to user input. These distinctions are consistent with the

technical organization of speech and language processing described by Jurafsky and Martin (2026). Several operations may be combined in one application, but their combination does not eliminate the need to examine each output critically.

Pedagogically, the important distinction is between what a system produces and what a learner does with it. Feedback can support learning when students understand it, judge its relevance, and act on it. A correction that is accepted without understanding may improve a submitted sentence without developing the student's ability to produce a similar sentence independently. Accordingly, the applications below are organized around opportunities for noticing, practice, revision, and transfer to another task.

4. Applications in Language Learning

4.1. Speech recognition and pronunciation practice

Automatic speech recognition can give learners an immediate written representation of their spoken production. Comparing intended words with the displayed transcript creates an opportunity to investigate communication problems. Nickolai et al. (2024), in a synthesis of 50 studies, identify substantial support for ASR-based pronunciation instruction and evaluation. Their review also highlights the diversity of study designs and research claims, which limits simple generalization from an individual application.

Learner behaviour is especially important. Inceoglu et al. (2024) examined 46 learners of English across six sessions of autonomous pronunciation practice. They investigated responses to unsuccessful transcription, including repeated attempts, ignored problems, and problems that remained unresolved. Learners' initial pronunciation attitudes were associated with their willingness to try again. This directs attention to what happens after feedback appears: access to repeated practice does not guarantee that learners will interpret or use it productively.

For classroom use, the transcript should be treated as a diagnostic prompt rather than a complete judgment of pronunciation. The teacher can ask students to replay the recording, compare it with the intended message, and consider whether the problem concerns production, recording conditions, or recognition. Human listening remains useful for assessing whether the message is understandable and whether stress or phrasing helps communication. The objective is successful speech, not simply obtaining a preferred transcript through repeated guessing.

4.2. Conversational agents and interaction

Conversational agents can provide a partner for additional practice when classroom speaking time is limited. However, a scripted dialogue, a voice assistant, and a generative chatbot should not be assumed to provide equivalent learning opportunities. Huang et al. (2022) identify uses of chatbots as interlocutors, simulations, and sources of support, together with challenges involving technical limitations, novelty, and cognitive load. Their review helps explain why apparent engagement needs to be examined alongside the quality of the interaction.

Bibauw et al. (2022) synthesize evidence from 17 studies of dialogue systems and report a medium overall effect on second-language proficiency. This supports dialogue practice as a promising area, but the estimate concerns the studies included in that analysis. More recent attention to generative systems adds questions about how teachers and learners evaluate outputs. Kohnke et al. (2023) discuss both the possibilities of ChatGPT for language education and the digital competencies required for ethical and effective use.

A productive task should require a communicative action, such as requesting clarification, explaining a preference, or resolving a misunderstanding. The teacher can specify the learner's level, the situation, and the purpose of the exchange, then examine whether the responses remain appropriate. Students should subsequently practise with peers so that they encounter turn-taking and repair beyond the patterns of one system. A fluent automated response is useful input only when it is accurate enough for the task and understandable to the learner.

Evidence also has contextual limits. Lai and Lee's (2024) review of 32 studies found that the included work was concentrated in Asian EFL settings. Such concentration makes replication in other educational contexts valuable. It would be premature to assume that differences in curriculum, access, language background, and classroom expectations have no influence on outcomes.

4.3. Machine translation and reflective revision

Machine translation can help learners access unfamiliar texts and compare alternative expressions. Its use also raises a question about which parts of the intellectual work the learner is expected to perform. Jolley and Maimone (2022) organize three decades of research around learner use, attitudes, effects on learning, and instructional responses. Their review shows that the educational issue extends beyond deciding whether translation tools should be allowed.

A useful classroom response is to make the learner's reasoning visible. Students can prepare an initial version, examine a machine-generated alternative, and explain which changes they accept. They might compare word choice, sentence structure, register, or the preservation of the original message. A disagreement between versions then becomes material for analysis. This proposed approach shifts attention from the final polished paragraph to the decisions that produced it.

The degree of permitted assistance should follow the assessment purpose. If a task assesses independent writing, unrestricted translation may obscure the ability being measured. If a task assesses revision or critical language awareness, a documented comparison may be appropriate. Teachers should explain these differences before the activity and use follow-up questions or short independent tasks to check whether students understand the language they submit.

4.4. Virtual environments and contextualized communication

Virtual environments can situate language use within an activity: finding a destination, asking for assistance, describing an exhibit, or cooperating on a task. Chen et al. (2022) report a positive overall effect of extended reality on language learning in a meta-analysis of 17 studies. Their analysis also identifies variation associated with the target language. Because extended reality includes different technologies, this result should not be interpreted as evidence that a headset is always preferable to a screen-based activity.

Hua and Wang (2023) review 38 empirical studies and identify both cognitive and affective benefits of virtual reality-assisted language learning. They also document problems including technical difficulties, distraction, time demands, and unsuitable materials. These findings favour close attention to the learning task and the quality of the environment rather than treating immersion as an educational outcome in itself.

Makransky and Petersen's (2021) Cognitive Affective Model of Immersive Learning provides a broader explanatory framework. It relates presence and agency to processes including motivation, cognitive load, and self-regulation. Applied to language teaching, this

suggests that the experience of being in a simulated setting needs instructional direction. A learner who devotes most attention to navigation may have little attention available for listening or formulating a response.

Teachers can reduce this competition by introducing controls before the language task, limiting unnecessary visual detail, and providing a clear communicative goal. A short preparation phase can introduce essential expressions; a follow-up discussion can revisit what learners said and what they could not express. Where immersive equipment is unavailable or unsuitable, a screen-based scene or printed visual can preserve the same communicative objective.

5. Teacher Guidance and Adaptive Learning

Digital personalization usually depends on information that a system can capture, such as submitted answers or interaction history. Teachers can consider additional evidence, including hesitation, misunderstanding of instructions, and the social dynamics of group work. Kuhail et al. (2023), in a systematic review of educational chatbots, distinguish different design principles, roles, and interaction styles across educational fields. This diversity reinforces the need to select a tool for a defined instructional function rather than assuming that personalization has one consistent meaning.

Alpha School offers an institutional illustration of adaptive learning combined with human support. Its website describes a model that concentrates core academic work into approximately two hours and allocates other time to life-skills activities, with adults supporting learners (Alpha School, n.d.). These are descriptions supplied by the institution. They do not establish that the model improves foreign-language learning, and the website is not treated here as an independent evaluation. The example is relevant because it makes the allocation of work between digital systems and educators visible. Its transfer to language education would require separate evidence about interaction, linguistic development, and learners' needs.

For language teachers, a practical division of responsibility is to use digital tools for selected practice and information while retaining responsibility for task design, interpretation of difficulties, and educational judgment. Students also need explicit guidance on asking for clarification, recognizing uncertain feedback, and deciding when to seek human help. Independence develops through supported decision-making, not simply through working alone with a device.

6. Access, Reliability, and Ethical Responsibilities

Digital access includes more than possession of a phone. Learners may differ in connectivity, available time, accessibility needs, and confidence using unfamiliar interfaces. The Global Education Monitoring Report Team (2023) examines technology in relation to equity, teacher preparation, and governance, emphasizing that educational value depends on the conditions of implementation. For language courses, this means that an activity should have a feasible alternative when its technical requirements prevent participation.

Accessibility arrangements should respond to the task. Transcripts can support access to spoken material, while audio can support access to written material. Yet an accommodation must be considered alongside the ability being assessed. A teacher needs to distinguish providing access to instructions from changing the construct of a listening assessment. Such decisions require pedagogical judgment and attention to the learner's circumstances.

The handling of personal data is another instructional responsibility. UNESCO's guidance recommends a human-centred approach to generative AI, including attention to privacy, inclusion, and the validation of educational uses (Miao & Holmes, 2023). A practical implication for language activities is to use fictional situations where personal details are unnecessary. Teachers should review institutional arrangements before requiring students to upload identifiable recordings, private correspondence, or assessed work to external platforms.

Reliability also affects how much authority should be assigned to automated feedback. Teachers can ask learners to identify the reason for a correction and compare it with an appropriate reference or teacher explanation. When the reason is unclear or contested, the uncertainty should be discussed. This practice makes feedback an object of inquiry and supports the critical digital competencies emphasized by Kohnke et al. (2023).

7. Proposed Classroom Applications and Evaluation

The following applications illustrate how the preceding analysis can inform teaching. They are proposals, not reports of tested interventions. Each connects a digital function to an observable language action and includes an opportunity to assess performance beyond the initial output.

First, a pronunciation activity can ask learners to record directions to a familiar place, inspect an automatic transcript, and select two passages for revision. They explain why they repeated or changed those passages. A partner then listens to a new recording and identifies the destination without seeing the transcript. The teacher can examine whether revision helped a listener understand the message, rather than recording only whether the software recognized the words.

Second, a conversational task can place an intermediate learner in the role of a visitor requesting a change to a reservation. The learner must explain the problem, ask a follow-up question, and confirm the agreed arrangement. After a short chatbot exchange, pairs perform a related situation with a different constraint. The teacher observes whether students can adapt the language and repair misunderstanding instead of reproducing the earlier exchange verbatim.

Third, a writing task can require a short original invitation followed by comparison with a machine-translated alternative. Students mark changes in tone, time reference, and clarity, then justify two accepted or rejected suggestions. A subsequent invitation addressed to a different audience is written independently. This sequence makes revision decisions inspectable and provides a separate indication of what the learner can transfer.

Fourth, a virtual visit can require pairs to choose an exhibit and explain their choice to another pair. Learners receive the communicative goal and essential vocabulary before entering the environment. They subsequently produce a brief spoken description without the visual display. Where equipment is unavailable, the teacher can use images of the same objects. Keeping the objective comparable makes it easier to evaluate whether the technological arrangement adds useful support.

Across these activities, evaluation should distinguish participation, perceived usefulness, and language performance. Completing an activity or enjoying it is valuable information, but neither alone demonstrates learning. Teachers can retain a small sample of initial and revised work and use a later comparable task to investigate retention. Any formal study would additionally need a defined research design, appropriate consent and data procedures, and transparent reporting of the assessment criteria.

8. Discussion and Limitations

The analysis suggests that digital technologies support language education through different mechanisms. Speech recognition makes aspects of spoken production visible, dialogue systems extend opportunities to respond, translation tools provide material for comparison, and virtual environments supply situational context. Their educational contribution depends on the learner's engagement with these functions and the teacher's ability to connect activity with assessment.

An important implication is that automated assistance and learner achievement should be examined separately. Better output during assisted practice may be a useful step, but the educational aim often includes performing a related task with less support. The proposed applications therefore incorporate explanation, peer interaction, or an independent follow-up. This is a design recommendation arising from the analysis, not an experimentally established model.

The article has several limitations. Source selection was selective, and the discussion relies substantially on existing reviews rather than a new appraisal of all underlying studies. Different reviews may include overlapping research, so their findings should not be counted as wholly independent confirmations. Tools also change over time, while evidence may concern earlier versions or different forms of automation. Finally, the proposed activities have not been evaluated in a particular institution or learner population. Future research should compare clearly specified tasks, document teacher support, and examine transfer and retention as well as immediate performance.

9. Conclusion

Digital technologies and virtual environments can extend the settings in which language learners encounter input, practise communication, and respond to feedback. Their value becomes clearer when each tool is associated with a specific learning purpose and an appropriate way of checking progress. Effective use requires teachers to guide interpretation, preserve opportunities for human interaction, and ensure that access and data practices do not exclude or disadvantage learners.

The practical contribution of this article is a set of applications that connects digital assistance to learner decisions and subsequent performance. These applications provide a basis for classroom investigation while retaining a clear distinction between proposed practice and demonstrated effects. The priority for implementation is to select manageable tools, explain their role to learners, and evaluate the language abilities that the instruction is intended to develop.

Declarations

Funding: This research received no external funding.

Conflicts of Interest: The author declares no conflict of interest.

Ethics and Data Availability: This article draws on published sources and publicly available institutional information. It involved no human participants and generated no empirical dataset. The classroom applications are illustrative proposals and have not been presented as observed results.

Use of Generative AI: Generative AI assisted with structural organization, language editing, and reference formatting; all content was reviewed and verified by the author.

Author Contributions: Conceptualization, analysis, drafting, and revision: A.M.O.

How to Cite:

Orujlu, A. M. (2026). Digital technologies and virtual environments in language learning. *Global Spectrum of Research and Humanities*, 3(4), 279-287. <https://doi.org/10.69760/gsrh.0260304028>

References

- Alpha School. (n.d.). *Overview*. Retrieved October 29, 2026, from <https://alpha.school/overview/>
- Bibauw, S., Van den Noortgate, W., François, T., & Desmet, P. (2022). Dialogue systems for language learning: A meta-analysis. *Language Learning & Technology*, 26(1), 1–24. <https://doi.org/10.64152/10125/73488>
- Chen, J., Dai, J., Zhu, K., & Xu, L. (2022). Effects of extended reality on language learning: A meta-analysis. *Frontiers in Psychology*, 13, Article 1016519. <https://doi.org/10.3389/fpsyg.2022.1016519>
- Global Education Monitoring Report Team. (2023). *Global education monitoring report 2023: Technology in education: A tool on whose terms?* UNESCO. <https://doi.org/10.54676/UZQV8501>
- Hua, C., & Wang, J. (2023). Virtual reality-assisted language learning: A follow-up review (2018–2022). *Frontiers in Psychology*, 14, Article 1153642. <https://doi.org/10.3389/fpsyg.2023.1153642>
- Huang, W., Hew, K. F., & Fryer, L. K. (2022). Chatbots for language learning — Are they really useful? A systematic review of chatbot-supported language learning. *Journal of Computer Assisted Learning*, 38(1), 237–257. <https://doi.org/10.1111/jcal.12610>
- Inceoglu, S., Chen, W. H., & Lim, H. (2024). Monitoring student behavior in autonomous automatic speech recognition-based pronunciation practice. *System*, 124, Article 103387. <https://doi.org/10.1016/j.system.2024.103387>
- Jolley, J. R., & Maimone, L. (2022). Thirty years of machine translation in language teaching and learning: A review of the literature. *L2 Journal*, 14(1), 26–44. <https://doi.org/10.5070/L214151760>
- Jurafsky, D., & Martin, J. H. (2026). *Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition with language models* (3rd ed. draft, August 19). <https://web.stanford.edu/~jurafsky/slp3/>
- Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for language teaching and learning. *RELC Journal*, 54(2), 537–550. <https://doi.org/10.1177/00336882231162868>

- Kuhail, M. A., Alturki, N., Alramlawi, S., & Alhejori, K. (2023). Interacting with educational chatbots: A systematic review. *Education and Information Technologies*, 28, 973–1018. <https://doi.org/10.1007/s10639-022-11177-3>
- Lai, W. Y. W., & Lee, J. S. (2024). A systematic review of conversational AI tools in ELT: Publication trends, tools, research methods, learning outcomes, and antecedents. *Computers and Education: Artificial Intelligence*, 7, Article 100291. <https://doi.org/10.1016/j.caeai.2024.100291>
- Makransky, G., & Petersen, G. B. (2021). The Cognitive Affective Model of Immersive Learning (CAMIL): A theoretical research-based model of learning in immersive virtual reality. *Educational Psychology Review*, 33, 937–958. <https://doi.org/10.1007/s10648-020-09586-2>
- Miao, F., & Holmes, W. (2023). *Guidance for generative AI in education and research*. UNESCO. <https://doi.org/10.54675/EWZM9535>
- Nickolai, D., Schaefer, E., & Figueroa, P. (2024). Aggregating the evidence of automatic speech recognition research claims in CALL. *System*, 121, Article 103250. <https://doi.org/10.1016/j.system.2024.103250>

Received: 1 September 2026

Accepted: 29 October 2026

Published: 31 October 2026