AI-NYELV
AI LANGUAGE

Co-funded by
the European Union

AI-assisted Foreign Language Education in Higher Education

D3.1: Comparative Analysis Report of AI Tools for Language Learning, Language Teaching, and Translation/Interpreting

Project data

Project	AI-ASSISTED FOREIGN LANGUAGE EDUCATION IN HIGHER EDUCATION
Project Acronym	AI-NYELV
Project number	2024-2-HU01-KA220-HED-000284589
Project Start Date	01-05-2025
Project End Date	31-10-2027
Coordinator	INSEDO OKTATÁSI ÉS SZOLGÁLTATÓ KORLÁTOLT
Partners	UNIVERSITAT DE BARCELONA UNIVERSITY OF CYPRUS GOETEBORGS UNIVERSITET SCUOLA SUPERIORE PER INTERPRETI E TRADUTTORI UNIWERSYTET WARSZAWSKI UNIVERSITY OF PANNONIA FIAS INTERNATIONAL

Deliverable data

Work Package	WP3 - STUDENT RESOURCES
Work Package Leader	GOETEBORGS UNIVERSITET
Deliverable number	3.1-EN
Deliverable name	Comparative Analysis Report of AI Toolsfor Language Learning, Language Teaching, and Translation/Interpreting
Dissemination level	Public
Date of release	22/04/2026
Status	Final

Versions

Version	Release date	Author(s)	Changes
V1 - draft	08/01/2026	SSML Carlo Bo	First draft of the document, 24 tools
V2 - draft	27/02/2026	SSML Carlo Bo	Improved analysis, minor corrections
V3 - final	22/04/2026	SSML Carlo Bo	Final version with 26 tools

Table of Contents

Executive Summary

This **comparative analysis** presents **26 artificial intelligence (AI) tools** identified and selected through the preceding desk research activity (*A3.1*) of the AI-NYELV project.

The evaluated tools span three primary stakeholder categories, with several versatile systems and applications serving multiple user groups and use cases — **language learning**, **language teaching**, and **translation/interpreting tools**.

The report provides **feature-based insights** into contemporary AI-enabled language technology and resources to **support informed decision-making processes and adoption** by language learners, educators, and translators/interpreters in higher education institutions and professional contexts.

Key findings demonstrate the growing prominence of personalisation and adaptive learning: contemporary language learning platforms leverage substantial advances in **AI-enabled personalisation**, employing adaptive algorithms which adjust difficulty and content based on learner performance.

Concurrently, **machine translation** (MT) by neural engines and large language models (LLMs) appears to be **nearing human quality** for high-resource languages, with computer-assisted translation (CAT) and interpreting (CAI) tools showing increasing sophistication and maturity, through the integration of automatic speech recognition (ASR) and transcription, AI-grounded glossary creation and terminology management, and real-time support functionalities.

Still, throughout all tool categories, **human oversight**, pedagogical curation, and expert verification remain essential, as AI-generated content may contain inaccuracies, errors, biases, or inappropriate formulations.

This analysis report also dedicates substantial attention to **ethical and pedagogical considerations** of AI tool adoption across several critical dimensions, including:

- **Data privacy**: information security, GDPR (General Data Protection Regulation) compliance, and protection measures vary; institutional and professional users bear responsibility for appropriate confidentiality safeguards.
- **Algorithmic biases**: AI systems may perpetuate linguistic, cultural, ethnic, or demographic biases present in training data, when this is not adequately diversified and sufficiently representative.
- **Digital equity**: commercial subscription models, technology requirements, and private ownership constraints may create accessibility barriers.

- **Over-reliance risks**: excessive delegation to or dependence on AI may undermine the development of foundational cognitive competence or the maintenance of independent capabilities, critical expertise, and robust skills.
- **Academic integrity**: clear guidelines and policies are needed for appropriate AI usage in learning environments, tutoring context, and assessment settings.

Transparency: limited explainability poses challenges for educational accountability and professional trustworthiness.

1. Introduction

1.1 Background and Context

The **integration of artificial intelligence (AI) into language education** represents one of the most significant pedagogical developments of the contemporary era. AI technology is fundamentally and profoundly transforming **how languages are learned, taught, and professionally applied** in translation and interpretation contexts.

Early research indicates **major disparities in AI literacy between students and educators**: according to a National Literacy Trust study[1], the percentage of 13–18-year-olds using generative AI surged from 37.1% in 2023 to 77.1% in 2024, while RAND Corporation surveys[2] indicate that teacher adoption still lags behind, reflecting a significant gap in potential engagement with these tools. Such a generational digital divide in the use of AI suggests that teachers are currently **unable to meet students' expectations of innovative teaching methods and technologies**.

The AI-NYELV project aims at responding to this transformation by systematically **investigating, evaluating, and disseminating knowledge about AI tools** relevant to language users and professionals.

Funded under the Erasmus+ programme for cooperation partnerships in higher education (KA220-HED), the project brings together eight partner institutions across six European countries (see *Table 1* below).

In particular, the project's work package n°3 (WP3) — led by the University of Gothenburg, co-led by SSML Carlo Bo, with the collaboration of FIAS International — has the primary objectives of identifying, evaluating, and presenting **student resources** on AI-based tools which can enhance the language learning experience.

WP3 directly contributes to achieving two specific project results: **enhancing learning outcomes** and **promoting critical thinking**. The comparative analysis of the tools ensures that pedagogically sound options are presented (also through a complementary online software gallery), to provide educators and learners with clear guidance on which AI tools are most effective for specific teaching and learning contexts, and improve the overall quality of language instruction.

[1] National Literacy Trust (2024) *Children, young people and teachers' use of generative AI to support literacy in 2024*, https://literacytrust.org.uk/research-services/research-reports/children-young-people-and-teachers-use-of-generative-ai-to-support-literacy-in-2024/.
[2] RAND Corporation (2025) *AI Use in Schools Is Quickly Increasing but Guidance Lags Behind: Findings from the RAND Survey Panels*, https://www.rand.org/pubs/research_reports/RRA4180-1.html.

Institution	Country	Role
INSEDO	Hungary	Coordinator
University of Gothenburg	Sweden	WP3 leader
SSML Carlo Bo	Italy	WP3 co-leader
FIAS International	Italy	Partner
University of Barcelona	Spain	Partner
University of Cyprus	Cyprus	Partner
University of Pannonia	Hungary	Partner
University of Warsaw	Poland	Partner

Table 1. AI-NYELV Project Partnership.

1.2 Purpose and Objectives

The primary purpose of this comparative analysis is to assist language learners, educators, and professionals in making informed decisions when selecting and using AI tools by systematically evaluating and **presenting functionalities, strengths, and limitations**.

Based on feature evaluations, the report is set to emphasise appropriate use cases, raise awareness of limitations rather than indiscriminate adoption, and ultimately **promote critical thinking**.

The analysis also has the objective of supporting AI fluency by offering conceptual frameworks to approach these tools, while sensitizing teachers and students on **responsible AI use**, providing categories to critically and ethically assess AI-produced information, and constructively build upon generated outputs as a starting point.

1.3 Methodology

The comparative analysis builds upon desk research conducted within *Activity A3.1*, evaluating the 26 selected representative AI tools (see *Table 2* below) through technical documentation, academic literature, user reviews, and hands-on testing by project partners. To facilitate reporting and reading, results are subsequently presented by means of tables, charts, use cases, and performance indicators.

For the analytical framework, 12 functional feature categories and 7 evaluation criteria were considered for each tool and user group:

- **Functional features**: core functionality, language skills covered, AI technologies used, multilingual support, real-time capabilities, customisation/adaptation, assessment tools, accessibility and flexibility (across space and time), data privacy/storage and security, collaboration support, integrability, and further limitations.

- **Evaluation criteria** for pedagogical/professional value: skill development, engagement (cognitive, affective, behavioural), ease of use, reliability/accuracy, AI transparency/explainability, autonomy, ethical considerations.

Babbel	BoothMate	ChatGPT
Coursebox	DeepL	Duolingo
ElevenLabs	ELSA Speak	Google Gemini
Google Translate	Grammarly	InterpretBank
LanguageTool	LearningStudioAI	MagicSchool
Matesub	Microsoft Translator	Mini Course Generator
NotebookLM	Notion	Perplexity
PONS Translate Pro	Talkpal	TeachAid
TTS Maker	Wordscope	

Table 2. List of AI tools identified and selected in A3.1 and analysed in A3.2 of the AI-NYELV project.

1.4 Scope and Limitations

This comparative analysis **focuses primarily on tool features**, functionalities, and documented capabilities, rather than empirical results or user data derived from controlled experiments or surveys — which are the object of other activities within the project instead.

As AI tool capabilities evolve rapidly, with new features, pricing models, and broader language support added on a continuous basis, the functionalities described in such a static report are **limited in time** and reflect the state of tool development at the period of analysis (end of the year 2025). Users are encouraged to verify the latest specifications, structures, and terms of service before making decisions on active adoption.

Moreover, most evaluated tools deliver stronger performance for high-resource languages compared to languages with lower resources in the digital space, which may benefit from narrower or under-performing support. This **disparity** reflects broader patterns in AI development, and has significant implications for digital language equality, especially across educational programmes.

Beyond functional evaluation, the report covers **ethical considerations and concerns** which have become increasingly central and urgent to discussions of AI in education and alignment with European values of human dignity, fairness, and transparency.

These include accessibility barriers, algorithmic bias, data privacy concerns, and the implications of over-reliance on technology in pedagogical and professional settings. By incorporating these dimensions, the analysis equips users with a broader perspective on the potential impacts, responsibilities, and risks involved in deploying AI solutions in language education settings.

1.5 Consultation Guide

While the final report following *Activity A3.1* desk research delivers exhaustive tool-by-tool description, this comparative analysis prioritises accessibility for stakeholders who typically approach AI tools from the perspective of their primary role — whether as language learners seeking practice, educators designing instruction, or translation/interpreting professionals managing workflows.

The report is, therefore, **organised by user category** to facilitate targeted navigation and consultation by readers with specific informational needs or professional interests. Various tools serve more use cases and appear over multiple groups, reflecting their **applicability to diverse relevant uses** (see *Figure 1* below).

Section 2, Tools for Language Learning, addresses students and independent learners seeking AI support for vocabulary acquisition, grammar improvement, writing development, and speaking/pronunciation practice. Tools examined include gamified applications, conversational AI, writing assistants, and speech synthesis platforms.

Section 3, Tools for Language Teaching, targets educators experimenting with AI assistance for lesson planning, content creation, assessment support, and differentiated instruction design. The section examines both general-purpose tools and systems specifically designed for education, with attention to pedagogical frameworks, integration with learning management systems (LMSs), and classroom management/collaboration features.

Section 4, Tools for Translation and Interpreting, serves both professionals and students preparing for professional practice. Coverage spans neural machine translation (NMT) engines, computer-assisted translation (CAT) and computer-assisted interpreting (CAI) tools, and audiovisual translation/subtitling software.

The Conclusion synthesises findings across categories, identifies emerging trends, outlines user group-specific recommendations, and offers cross-cutting ethical and pedagogical considerations, promoting a framework for responsible AI observation, evaluation, and adoption.

The section also includes consultation and decision-support tables harmonising findings and trends by learner proficiency and autonomy (*Table 10*), integration context (*Table 11*), and risk profiles (*Table 12*).

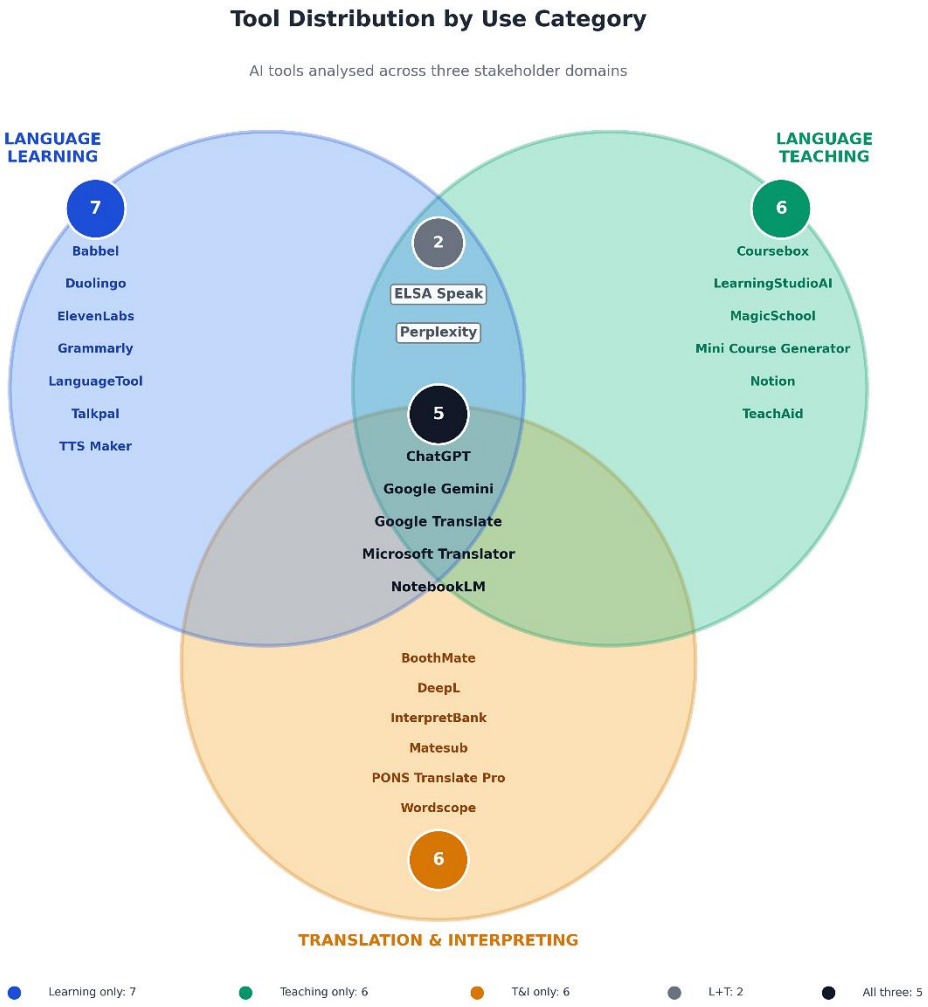

Figure 1. Tool Distribution by Use Category.

1.6 Impact and Dissemination

This comparative analysis is one of the foundational deliverables within the AI-NYELV project framework, with multiple pathways for impact and dissemination designed to extend the project's outreach beyond the consortium of partners.

- **Online software gallery** (*Activity A3.4*). The findings of this report will be integrated into an interactive online tool showcase, providing both relevant stakeholders and the general public with accessible, searchable information on AI tools for language education to foster research and informed decision-making aligned with individual and institutional needs.
- **Software presentation** (*Activity A3.3*). Insights from this comparative analysis will inform the development of training materials for educators, students, and general users, supporting

vocational development initiatives to build solid AI literacy and competence in integration of AI tools.

- **Partner institution implementation**. Each partner institution will have the opportunity to adapt findings for local settings, integrating recommendations into their language programmes, translator/interpreter training curricula, and broader institutional AI policies. This distributed implementation model helps project outcomes respond to the diverse linguistic, cultural, and educational environments represented across the consortium.

- **Stakeholder engagement**. Through programme dissemination channels and partner institution networks, the report and all associated resources will reach language professionals, learners, trainers, educational administrators, and policymakers, supporting considered evidence-rooted decisions regarding AI tool adoption in European higher education beyond project participants.

Academic publications: The project partnership also intends to pursue academic publication of key findings, contributing to the scholarly literature on AI in language education and translation/interpreting studies, while enabling broader engagement with the international research community.

2. Tools for Language Learning

This section examines and compares AI tools supporting language learners in vocabulary acquisition, grammar development, writing skills improvement, and speaking/pronunciation practice. *Table 3* and *Table 4* below provide an overview of the core analysed functional features and evaluation criteria for the considered tools, which are then developed in detail in the following subsections.

Tool	Core Functionality	Language Skills Covered	Assessment Tools Availability	Accessibility	Data Privacy
Duolingo	Autonomous gamified language learning	Reading, Listening, Speaking, Writing	Instant feedback, progress tracking	Web and mobile; partial access offline; freemium model	User data used for training
Babbel	Expert-designed language learning	Reading, Listening, Speaking, Writing	Instant feedback, progress tracking	Web and mobile; offline download; subscription required	Data may transfer outside the EEA with safeguards
Grammarly	Spelling correction and writing assistance	Writing	Progress tracking and reports	Web, desktop, mobile, and API; freemium model	Data encryption; user data used to train AI models
LanguageTool	Spelling correction and writing assistance	Writing	Feedback-based self-assessment	Web, desktop, mobile, and API; freemium model	Data encryption; user data used to train AI models
ChatGPT	Multimodal conversational chatbot	Reading, Writing, Listening, Speaking	No formal assessment	Web, desktop, mobile, and API; multimodal, freemium model	Free: data used to train AI models (opt-out available); Paid: data not used for training

Tool	Core Functionality	Language Skills Covered	Assessment Tools Availability	Accessibility	Data Privacy
Google Gemini	Multimodal conversational chatbot	Reading, Writing, Listening, Speaking	No formal assessment	Web, mobile, and API; multimodal; freemium model; Google ecosystem integration	Consumer: subject to Google privacy policies, Paid: data not used for training
Google Translate	Neural machine translation	Reading, Listening, Speaking	No formal assessment	Web, mobile, and API; multimodal; freemium model	Consumer: subject to Google privacy policies, Commercial API: user data not used to train AI models
Microsoft Translator	Neural machine translation	Reading, Listening, Speaking	No formal assessment	Web, mobile, and API; multimodal; freemium model	Temporary data processing only; subject to Microsoft privacy policies
Talkpal	Conversation practice	Speaking, Listening	Progress tracking and statistics	Web and mobile; freemium model	User data used to train AI models
ELSA Speak	Pronunciation and fluency practice	Speaking, Listening	Pronunciation analytics, instant feedback	Web, mobile, and API; freemium model	GDPR-compliant data processing
TTS Maker	Speech synthesis	Listening	No formal assessment	Web, mobile, and API; freemium model	User data used to train AI models
ElevenLabs	Speech synthesis and voice cloning	Listening	No formal assessment	Web, mobile, and API; freemium model	Data encryption; voice cloning requires user consent
NotebookLM	Knowledge analysis assistance	Reading, Listening	Quizzes and flashcards	Web and mobile; offline download; freemium model	User data not used to train AI models
Perplexity	Conversational search engine	Reading, Writing	No formal assessment	Web and mobile; freemium model	Data processed by third-party AI models

Table 3. Comparative Overview of Functional Features in AI Tools for Language Learning

Tool	Skill Development	Autonomy	Engagement	Reliability	AI Transparency
Duolingo	Vocabulary, grammar, speaking, listening, reading	Self-paced, structured learning paths	High (gamification)	Highly accurate	Opaque: processes not fully disclosed
Babbel	Vocabulary, grammar, speaking, listening, reading	Self-paced, structured learning paths	High (goal-oriented)	Highly accurate	Opaque: processes not fully disclosed
Grammarly	Writing, grammar	High (self-regulated use)	Moderate (instant feedback)	High for standard writing	Opaque: processes not fully disclosed
LanguageTool	Writing, grammar	High (self-regulated use)	Moderate (instant feedback)	High for major languages	Open source code
ChatGPT	Writing, grammar, vocabulary, conversational fluency	High (self-regulated use)	High (versatility), but over-reliance risk	Output may produce hallucinations, thus requires verification	Opaque: architecture, data, processes not fully disclosed; warns output may be inaccurate
Google Gemini	Writing, grammar, vocabulary, conversational fluency	High (self-regulated use)	High (versatility), but over-reliance risk	Output may produce hallucinations, thus requires verification	Opaque: architecture, data, processes not fully disclosed; warns output may be inaccurate
Google Translate	Vocabulary, translation, conversational fluency (*Practice* mode)	High (self-regulated use)	Variable: can drive engagement in communication; risks of MT dependence and passive processing	Appropriate for general comprehension; accuracy varies by language pair	Opaque: results provided without linguistic rationale; customisation controls offered

Tool	Skill Development	Autonomy	Engagement	Reliability	AI Transparency
Microsoft Translator	Vocabulary, translation	High (self-regulated use)	Variable: can drive engagement in communication; risks of MT dependence and passive processing	Appropriate for general comprehension; accuracy varies by language pair	Opaque: results provided without linguistic rationale; customisation controls offered
Talkpal	Speaking, conversational fluency, vocabulary, listening	Self-paced, personalised learning paths	High (role-plays)	High, but accuracy varies by language	Opaque: processes not fully disclosed
ELSA Speak	Speaking, listening, conversational fluency	High (self-regulated use)	High (role-plays, instant feedback)	High for standard accents	Feedback provides linguistic rationale
TTS Maker	Listening	High (self-regulated use)	Moderate (limited interactivity)	High, but accuracy varies by language	Limited (proprietary voice synthesis)
ElevenLabs	Listening	High (self-regulated use)	Moderate (limited interactivity)	High, but accuracy varies by language	Limited (proprietary voice synthesis)
NotebookLM	Reading, listening	High (self-regulated use)	Moderate: risks of AI over-reliance	Source grounding reduces hallucinations and enables verification	Google warns output may be inaccurate
Perplexity	Vocabulary, writing	High (self-regulated use)	Moderate: follow-up questions encourage deepening, but over-reliance risk	Source grounding reduces hallucinations and enables verification	Limited: source transparency, but synthesis processes not fully disclosed

Table 4. Comparative Overview of Evaluation Criteria in AI Tools for Language Learning.

2.1 Vocabulary and Grammar

Vocabulary acquisition and grammar development represent foundational components of language learning where AI tools have demonstrated particular effectiveness through diverse approaches — from gamified spaced repetition to conversational exploration.

Applications like *Duolingo* and *Babbel* employ adaptive algorithms which schedule vocabulary review based on individual learning and forgetting curves. *Duolingo*'s gamified approach embeds vocabulary within translation exercises and contextual sentences, reinforced through streaks, experience points, leaderboards, and reminders which motivate consistent practice. *Babbel*'s expert-designed courses also introduce vocabulary thematically, progressing from concrete to abstract terms within frameworks aligned with the Common European Framework of Reference for Languages (CEFR).

Although limited by reduced cultural content and subscription cost barriers, research[3] has found both platforms effective for vocabulary acquisition within practical conversational contexts.

Large language model (LLM)-based chatbots offer a different approach to vocabulary expansion: *ChatGPT* and *Google Gemini* enable learners to explore vocabulary through natural dialogue, requesting explanations, example sentences, collocations, register variations, and etymology on demand.

This flexibility supports autonomous exploration beyond pre-determined curricula, but lacks the systematic progression that structured learning platforms provide. While advanced learners may benefit from such a conversational discovery, beginners typically require the scaffolding of structured systems.

AI chatbots likewise offer learners a range of study support functions such as structured summaries of lecture notes and lesson materials highlighting key concepts and vocabulary, follow-up clarifications and additional examples outside classroom hours, personalised study guides, action plans, and revision checklists, and tailored feedback on assignments before submission.

Similarly, *Perplexity* serves as an AI conversational search engine which combines large language model capabilities with real-time web search, providing source-grounded responses to user queries as referenced answers.

For language learners, it acts as a research companion supporting language and culture exploration in authentic context across diverse sources, including news articles and academic sources. The citation feature enables learners to verify information and explore primary sources,

[3] Kessler, M., Loewen, S., and Gönülal, T. (2025) 'Mobile-assisted language learning with Babbel and Duolingo: comparing L2 learning gains and user experience', *Computer Assisted Language Learning*, 38(4), 690–714, https://doi.org/10.1080/09588221.2023.2215294.

while the conversational interface supports follow-up questions to progressively deepen their understanding.

Google Translate — and, similarly, *Microsoft Translator* — can also serve as a supplementary vocabulary acquisition tool, enabling learners to instantly look up unfamiliar words and phrases across hundreds of languages. By displaying alternative translations, definitions, and contextual examples, it helps learners understand semantic range and usage patterns.

However, it lacks systematic vocabulary progression and spaced repetition and does not provide explicit grammar instruction, rule explanations, or targeted exercises as found in dedicated language learning platforms — making it an immediate reference tool, rather than a structured learning environment.

Grammar treatment varies throughout tools: *Duolingo* introduces grammatical structures implicitly through pattern exposure, with explanatory sections available for some languages, whereas *Babbel* provides more explicit grammar instruction integrated within lessons, with dedicated review sections. LLMs excel at on-demand explanations, offering detailed analyses of specific structures, contrastive explanations (even across languages), and clarification of learner questions, functioning as always-available grammar references, rather than systematic curricula.

Tool selection for vocabulary and grammar development should, therefore, consider learner level, learning style preferences, and target language(s). Structured language learning platforms, such as *Duolingo* and *Babbel*, suit learners requiring systematic progression and external motivation. General-purpose conversational AI, like *ChatGPT* and *Gemini*, serves learners capable of self-directed study who can benefit from flexible exploration.

Finally, language coverage varies significantly: *Duolingo* and *Babbel* present a limited number of languages with deeper content, while *ChatGPT* and *Gemini* can support hundreds of languages, with variable quality for the less resourced.

2.2. Writing

Writing skill development in a foreign language presents a unique combination of challenges (linguistic accuracy, textual coherence, register appropriateness, rhetorical effectiveness, to name but a few). AI-enabled writing assistants and grammar checkers address these through on-the-fly feedback, though with relevant trade-offs regarding learner autonomy and skill internalisation.

Grammarly and *LanguageTool* are among the most sophisticated AI writing assistants, employing rule-based systems augmented by machine learning to identify spelling mistakes, grammatical errors, punctuation issues, and stylistic infelicities. Premium tiers extend beyond error correction to offer clarity improvements, engagement suggestions, and tone detection — features particularly valuable in educational or professional registers.

LanguageTool distinguishes itself through support for more than 30 languages (compared to *Grammarly*'s current primary English focus, with limited multilingual coverage) and custom rule configuration for language-specific error patterns.

ChatGPT and *Gemini* offer different writing assistance: in addition to correcting existing text, these tools engage in dialogue about writing, suggesting alternative phrasings, explaining why certain constructions may be preferred, and providing models for particular text types. Conversational AI tools are also capable of generating unlimited customisable exercises for practice and engage in cultural exploration with the user through natural dialogue.

Their explanatory capacity additionally supports the development of metalinguistic awareness, although it requires learners to formulate effective prompts and critically evaluate AI proposals. Moreover, despite impressive individual adaptation, these tools do not provide any formal structured curriculum or progress tracking specific to language learning needs.

For academic writing specifically, LLMs can assist with argumentation structure and paragraph organisation — features less developed in grammar-focused tools. However, academic integrity requires clear institutional guidelines regarding acceptable and accountable AI usage in assessed writing endeavours.

Automated writing evaluation raises concerns about over-reliance, as learners who uncritically accept AI corrections may not develop the independent editing skills necessary for contexts without AI support. Effective pedagogical integration requires teaching to evaluate AI output critically, understanding both capabilities and limitations of automated feedback. Furthermore, AI writing assistants privilege certain varieties and registers, typically standard written language in formal contexts, over regional dialects or specialised jargon. Learners targeting other varieties or creative expressions may find AI suggestions homogenising or inappropriate. Teachers should help learners recognise such biases, and develop judgement about when to accept or reject AI recommendations.

Therefore, tool selection depends on target language, writing context, and pedagogical or professional goals, as outlined in *Table 5* below:

Use Case	Recommended Tool(s)	Rationale
English writing assistance	Grammarly	Comprehensive feedback, extensive integrations across devices and software
Multilingual writing support	LanguageTool	Broader language coverage, personalisation features
Natural dialogue-based explanatory feedback	ChatGPT, Google Gemini	Unparalleled flexibility for customised language practice, metalinguistic capabilities

Table 5. Comparative Overview of Language AI Tools for Writing Correction and Assistance.

2.3. Conversational and Interactive Learning

Conversational competence to engage in real-time spoken interaction represents arguably the most challenging language skill to develop independently. AI conversational partners offer novel opportunities for practice outside classrooms, with always-available digital interlocutors tolerating errors without anxiety or social consequences.

However, significant limitations persist regarding conversational authenticity and feedback quality, as detailed in the following sub-sections.

2.3.1. Speaking and Pronunciation

Speaking skill advancement has traditionally been constrained by the availability of human interlocutors. AI speech recognition and synthesis platforms now enable independent practice, with early research[4] demonstrating their potential in supporting language acquisition and effectively adding to traditional language education for improved learning outcomes.

Talkpal represents a focused implementation of AI as a conversation practice partner, offering role-play scenarios spanning dozens of languages with real-time speech recognition and transcription, speech rate adjustment, and pronunciation feedback. Users can engage in simulated conversations (travel situations, ordering at restaurants, job interviews) with the AI partner adapting responses to learner proficiency.

ELSA Speak offers another specialised application of AI-supported pronunciation and speaking practice (currently focusing on English only), which analyses user speech at the phoneme level to detect mispronounced sounds, sentence stress, and intonation patterns. Its adaptive modalities and targeted corrective feedback facilitate self-regulated practice.

Babbel's AI conversational partner, *Babbel Speak*, and *Duolingo Max*'s incorporated *Roleplay* feature provide similar functionality, enabling application of learned vocabulary and structures in simulated dialogue to reinforce systematic learning progression.

General-purpose LLMs, like *ChatGPT* and *Gemini*, can equally serve as conversation partners through their voice modes. Considering their capability of adapting to virtually any subject or scenario, advanced learners can benefit from this advanced flexibility beyond text-based exchanges by prompting conversations in specific registers or about specialised topics.

[4] Torkhani, D. (2025) 'AI-enhanced language learning: The impact of Talkpal.AI on EFL undergraduate students' English-speaking skills', *American Journal of STEM Education*, 11, 69–82, https://doi.org/10.32674/232w1188; Grab, M. Ö. (2025) 'Integrated AI chatbot practice: A pathway to improved ESL speaking skills', *Social Sciences & Humanities Open*, 12, https://doi.org/10.1016/j.ssaho.2025.101933.

Speech recognition accuracy and pronunciation feedback vary significantly across languages and accents: tools typically perform best for standard varieties of high-resource languages. Hence, pronunciation scores should be interpreted as approximate indicators, rather than definitive assessments. Evaluation specificity also changes: some systems deliver binary correct/incorrect judgements, while others offer phoneme-level analysis identifying exact pronunciation difficulties. Thus, speaking practice with AI partners should complement — rather than replace altogether — human interaction, preparing learners for real communicative experiences.

2.3.2. Listening

Listening comprehension progress requires exposure to varied audio input at appropriate difficulty levels. AI tools can contribute through audio content generation and speech synthesis.

TTS Maker and *ElevenLabs* provide speech synthesis enabling learners to create listening materials from any text and hear target language pronunciation on demand. Voice cloning and emotional expression capabilities produce particularly natural output, suitable for creating listening materials or modelling pronunciation. However, synthetic speech cannot fully reproduce the coarticulation and pragmatic nuance of authentic human speech yet, warranting supplementation with authentic audio materials.

While not specifically designed for language learning, *NotebookLM*'s *Audio Overview* feature generates podcast-style discussions from uploaded documents, creating engaging conversations between AI hosts which summarise and explain source content.

Google Translate's recent *Practice* mode also introduced AI-generated conversational exercises with adaptive difficulty, representing an evolution towards more pedagogically oriented features, although still limited compared to dedicated learning-focused platforms (for instance, feedback on pronunciation accuracy remains restricted to successful intelligibility recognition, rather than detailed phonetic analysis).

Despite advances in synthesis naturalness, current AI conversation tools are still unable to replicate the entire complexity of intercultural communication, non-verbal cues, and genuine communicative pressure present in real-world contexts. Appropriate pedagogical practice should then combine AI-synthesised materials (for controlled input and pronunciation modelling) with authentic audio (for real-world listening comprehension readiness).

3. Tools for Language Teaching

This section examines AI tools supporting language educators in lesson planning, curriculum design, content creation, learner feedback, and assessment management. The analysis encompasses dedicated educational platforms alongside general-purpose AI assistants and MT utilities which can be re-purposed for pedagogical applications. Table 6 and Table 7 below provide an overview of the core analysed functional features and evaluation criteria for the considered tools, which are then developed in detail in the following subsections.

Tool	Core Functionality	Language Skills Covered	Assessment Tools Availability	Accessibility	Data Privacy
MagicSchool	Lesson generation, material design, and assessment creation	Reading, Writing, Listening	Test/quiz generator, rubric creator, questions, automated feedback	Web; LMS integration; freemium model	User data not used to train AI models; limited data retention for abuse monitoring
TeachAid	Lesson generation, material design, curriculum planning, and assessment creation	Reading, Writing, Listening	Test/quiz generator, rubric creator, questions, automated feedback	Web; freemium model	User data not used to train AI models
Coursebox	Online course generation	Reading, Listening, Writing (indirect: user-defined content)	Various question formats, automated grading and analytics	Web and mobile; offline download; freemium model; improved readability features	Data encryption
ChatGPT	Multimodal conversational chatbot	Reading, Writing, Listening, Speaking, Vocabulary, Grammar	No formal assessment	Web, desktop, mobile, and API; multimodal; freemium model	Free: data used to train AI models (opt-out available);

Tool	Core Functionality	Language Skills Covered	Assessment Tools Availability	Accessibility	Data Privacy
					Paid: data not used for training
Google Gemini	Multimodal conversational chatbot	Reading, Writing, Listening, Speaking, Vocabulary, Grammar	No formal assessment	Web, mobile, and API; multimodal; freemium model; Google ecosystem integration	Consumer: subject to Google privacy policies, Paid: data not used for training
NotebookLM	Knowledge analysis assistance	Reading, Listening	Quizzes and flashcards	Web and mobile; offline download; freemium model	User data not used to train AI models
Google Translate	Neural machine translation	Reading, Listening, Speaking, Translation, Vocabulary	No formal assessment	Web, mobile, and API; multimodal; freemium model	Consumer: subject to Google privacy policies, Commercial API: user data not used to train AI models
Microsoft Translator	Neural machine translation	Reading, Listening, Speaking, Translation, Vocabulary	No formal assessment	Web, mobile, and API; multimodal; freemium model	Temporary data processing only; subject to Microsoft privacy policies
Perplexity	Conversational search engine	Writing, Reading, Vocabulary	No formal assessment	Web and mobile; freemium model	Data processed by third-party AI models
Notion	Knowledge workspace	Reading, Writing, Listening (indirect: user-defined content)	No formal assessment	Web, desktop, mobile, and API; multimodal; freemium model	Data encryption

Tool	Core Functionality	Language Skills Covered	Assessment Tools Availability	Accessibility	Data Privacy
ELSA Speak	Speaking training with class monitoring for teachers	Speaking, Listening	Exercises, instant feedback	Web and mobile; subscription-based	GDPR-compliant data processing
Mini Course Generator	Online bite-sized course generation	Reading, Writing, Listening (indirect: user-defined content)	Various question formats, automated analytics	Web; LMS integration; freemium model	User data used for legitimate purposes only (Privacy Policy)
LearningStudioAI	Online course generation	Reading, Writing, Listening (indirect: user-defined content	Quizzes and flashcards	Web; LMS integration; subscription-based	Vague data storage and processing measures (Privacy Policy)

Table 6. Comparative Overview of Functional Features in AI Tools for Language Teaching

Tool	Skill Development	Autonomy	Engagement	Reliability	AI Transparency
MagicSchool	Lesson and curriculum design, requires self-directed development	High (self-regulated use)	Variable: facilitates learner-centred design to increase student engagement, but over-reliance risk for teachers	Output may produce hallucinations, thus requires verification	Opaque: generation processes not fully disclosed
TeachAid	Lesson and curriculum design, requires self-directed development	High (self-regulated use)	Variable: facilitates learner-centred design to increase student engagement, but over-reliance risk for teachers	Output may produce hallucinations, thus requires verification	Opaque: generation processes and architecture not fully disclosed
Coursebox	Curriculum design and lesson refinement, requires self-directed development	High (self-regulated use)	Variable: facilitates learner-centred design to increase student engagement, but over-reliance risk for teachers	Output may produce hallucinations, thus requires verification	Opaque: generation processes not fully disclosed
ChatGPT	Content design and critical verification, requires self-directed development	High (self-regulated use)	High (versatility), but over-reliance risk	Output may produce hallucinations, thus requires verification	Opaque: architecture, data, processes not fully disclosed; warns output may be inaccurate
Google Gemini	Content design and critical verification,	High (self-regulated use)	High (versatility), but over-reliance risk	Output may produce hallucinations, thus requires verification	Opaque: architecture, data, processes not fully

Tool	Skill Development	Autonomy	Engagement	Reliability	AI Transparency
	requires self-directed development				disclosed; warns output may be inaccurate
NotebookLM	Knowledge organisation and research-informed teaching, requires self-directed development	High (self-regulated use)	Variable: facilitates learner-centred design to increase student engagement, but risk of reduced engagement with original materials	Source grounding reduces hallucinations and enables verification	Google warns output may be inaccurate
Google Translate	Content adaptation (limited), communication	High (self-regulated use)	Variable: can drive engagement in communication, but limited pedagogical features	Appropriate for general comprehension; accuracy varies by language pair	Opaque: results provided without linguistic rationale; customisation controls offered
Microsoft Translator	Content adaptation (limited), communication	High (self-regulated use)	Variable: can drive engagement in communication, but limited pedagogical features	Appropriate for general comprehension; accuracy varies by language pair	Opaque: results provided without linguistic rationale; customisation controls offered
Perplexity	Content design and research-informed	High (self-regulated use)	Moderate: facilitates learner-centred design, but risk of reduced engagement with original materials	Source grounding reduces hallucinations	Limited: source transparency, but synthesis processes not fully disclosed

Tool	Skill Development	Autonomy	Engagement	Reliability	AI Transparency
	teaching, requires self-directed development			and enables verification	
Notion	Lesson and curriculum design, knowledge organisation, research-informed teaching	High (self-regulated use)	High: versatility and collaborative features promote active participation	AI-generated content may include hallucinations, thus requires verification	Opaque: AI generation processes not fully disclosed
ELSA Speak	Speaking instruction, personalised assessment	High (self-regulated use)	Variable: facilitates learner-centred design to increase student engagement	High for standard accents	Feedback provides linguistic rationale and analytics
Mini Course Generator	Lesson and content design	High (self-regulated use)	Variable: facilitates learner-centred design, but over-reliance risk	AI-generated content may include inaccuracies, thus requires verification	Limited: source transparency, but AI generation processes not fully disclosed
LearningStudioAI	Lesson and content design	High (self-regulated use)	High (versatility), but over-reliance risk	AI-generated content may include inaccuracies, thus requires verification	Limited: source transparency, but AI generation processes not fully disclosed

Table 7. Comparative Overview of Evaluation Criteria in AI Tools for Language Teaching

3.1. Curriculum Design and Lesson Planning

Effective curriculum development and content design are central to the educational experience. AI tools can support content creation and its adaptation to diverse learner populations, generate structured lesson sequences, and suggest pedagogical approaches. Still, the quality of AI-produced content significantly depends on educator input and subsequent refinement.

Focused tools like *TeachAid* and *MagicSchool* provide teacher-oriented resources specifically designed for curriculum planning, lesson preparation, and — as in the case of *Coursebox*, *Mini Course Generator*, and *LearningStudioAI* — generation of complete online courses from topic prompts. Following teachers' input on subject areas and learning objectives, these platforms can automatically generate lesson frameworks and instructional materials across dozens of languages (with varying performance levels) as well as practice exercises, assessment components, and even real-time activities for classroom interaction aligned with curriculum standards.

Unlike general-purpose systems, such educational AI assistants incorporate pedagogical and instructional design principles as well as specific educational terminology, and they generally support integrations with classroom and learning management systems to facilitate export of generated content.

ChatGPT, *Google Gemini*, and *NotebookLM* offer flexible curriculum design support through conversational interaction: teachers can request lesson outlines mapped to specific frameworks and differentiated activities for varied proficiency levels. Their multimodal capabilities also enable educators to upload existing materials and digital resources to request aligned lesson sequences.

The strength of these tools lies in their adaptability: instructors can iteratively refine generated content through natural dialogue, request specific approaches, or even ask for culturally relevant examples. Their extensive language coverage makes them particularly suitable for less common languages where dedicated platforms may be short of support. However, these tools lack the structured templates and education-specific guardrails of tailored platforms, requiring greater teacher expertise in prompt design and involvement in the evaluation of outputs.

Gemini's and *NotebookLM*'s integrations with the Google Workspace for Education suite enable educators to work within familiar environments while leveraging AI assistance. Their source-grounding capabilities reduce hallucination risks by anchoring generated analysis and output to uploaded documents, whereas *NotebookLM*'s *Audio Overview* feature can transform educational documents into engaging podcast-style discussions.

Notion is a standalone workspace platform, combining notes, documents, databases, and project management resources into a customisable environment. It can function as both a learning

management space and a knowledge construction tool, supporting course organisation, collaborative projects, and autonomous learning.

The platform's database functionality supports the creation of personalised vocabulary catalogues with AI-produced translations, example sentences, pronunciation audio, and grammatical annotations — organised by theme, frequency, proficiency level, or other criteria. Multimedia embedding capabilities support the integration of authentic materials like videos, podcasts, or articles into organised learning notebooks — while flashcard systems, spaced repetitions, and progress tracking can assist memorisation and assessment. Collaborative features allow group projects, peer review, shared glossary construction.

Finally, *Google Translate* and *Microsoft Translator* provide limited but useful supplementary support, primarily for multilingual classroom contexts. Translation accuracy varies by language, and teachers must verify translated materials before classroom use.

3.2. Educational Content Creation

Content creation encompasses the crafting of instructional materials, practice activities, and supplementary learning resources. This production can be considerably supported by AI tools from minimal input, while enabling personalisation and differentiation at scale and multilingually, supporting monolingual, bilingual or translanguaging pedagogies.

Coursebox, *MagicSchool*, and *TeachAid* are equipped with extensive libraries to create reading texts at specified difficulty levels, vocabulary reinforcement activities with contextual examples, grammar exercises, dialogue scripts for role-plays, and cultural information on demand or in real time, also during lesson delivery, to transform static resources into dynamic learning experiences.

Instructors can produce multiple versions of the same material adapted for different abilities and proficiency stages as needed, and generate worksheets, slide decks, quizzes, flashcards with text, images, or audio, or varied question format sets. Content can then be exported to common file formats for further customisation and distribution.

Chatbots like *ChatGPT* and *Gemini* provide highly flexible content creation capabilities, essentially limited only by teacher creativity in prompting. Their multimodality enables teachers to upload images or media and request related exercises, analyse authentic texts and generate comprehension questions, or create materials on current events through web search integration. These tools excel at interacting as custom tutors which respond to learner questions and at personalising content following students' interests, local contexts, or culturally relevant events.

However, as with any other AI tool, content accuracy demands verification — particularly for factual claims, cultural information, and linguistic examples in less resourced languages, although

the available educational versions of most of these tools provide additional safeguards and controls suitable for institutional contexts.

NotebookLM supports content creation through source transformation into pedagogical and learning resources — after uploading authentic materials, source-rooted summaries, study guides, question lists, flashcards, and quizzes are generated, ensuring alignment between content and assessment. *Audio* and *Video Overview* features convert uploaded and textual content into engaging AI-based podcasts and narrated presentations, while maintaining source material fidelity.

Perplexity also provides rapid, cited summaries of current events, cultural topics, and subject-specific content suitable for classroom materials. The tool's ability to access current web content makes it particularly valuable for creating relevant, engaging materials which connect language teaching to real-world contexts.

Google Translate and *Microsoft Translator* contribute to content creation primarily through translation and adaptation of existing materials, but do not currently generate original educational content or differentiated versions for varied proficiency levels.

3.3. Learner Assessment

AI tools can generate assessment items, provide automated feedback, analyse learner performance, and support teachers in designing evaluation frameworks.

Educational AI tools *TeachAid*, *MagicSchool*, *Coursebox*, *Mini Course Generator*, and *LearningStudioAI* offer comprehensive assessment capabilities through dedicated utilities for quiz design, checklist creation, and rubric development across varied testing instruments (multiple-choice questions, true/false items, fill-in-the-blank exercises, matching, ordering, etc.) aligned to the desired learning objectives and with customisable descriptors and scoring criteria.

These platforms support various formats and templates for assessing language skills: listening comprehension with audio clips, reading comprehension with questions, or writing tasks with model responses. AI-generated feedback identifies strengths and suggests areas for improvement, automated rewards incentivise learner achievement, while analytics track student participation and performance through completion rates, time per question, average scores, and competency mapping.

ChatGPT and *Gemini* enable flexible assessment generation through conversational interaction. Teachers can request questions, quizzes, and marking criteria for any topic or proficiency level (even for less common content or languages where item banks are unavailable), analyse learning objectives and generate aligned assessment items, and produce differentiated assessments for varied learner profiles. *Gemini* provides such capabilities with native Google Workspace integration. Still, teachers must evaluate generated items for validity, accuracy, and difficulty calibration.

NotebookLM generates assessments (quizzes, flashcards, and comprehension questions) grounded in uploaded source materials ensuring curriculum alignment, with explanations for correct and incorrect answers.

Google Translate and *Microsoft Translator* lack any standardised testing, progress tracking, or learning analytics capabilities to date — for assessment purposes, they only serve as reference tools which might be consulted during evaluation activities.

ELSA Speak's teacher resources contribute to feedback and monitoring in language skills, exemplifying how targeted AI can complement instruction in narrower areas more specifically than general educational platforms.

4. Tools for Translation and Interpreting

This section examines AI tools supporting translation and interpreting (T&I) professionals and students, from machine translation (MT) to specialised computer-assisted translation (CAT) and computer-assisted interpreting (CAI) software, including subtitling platforms. *Table 8* and *Table 9* below provide an overview of the core analysed functional features and evaluation criteria for the considered tools, which are then developed in detail in the following subsections.

Tool	Core Functionality	Language Skills Covered	Assessment Tools Availability	Accessibility	Data Privacy
DeepL	Neural machine translation	Translation, Writing, Reading (indirect)	No formal assessment or translation quality metrics	Web, desktop, mobile, and API; multimodal; freemium model	Data encryption; Free: data used to train AI models; Paid: data not used for training
Google Translate	Neural machine translation	Translation, Reading, Writing, Listening, Speaking (indirect)	No formal assessment (proficiency-based feedback in *Practice* mode only) or translation quality metrics (*MTQP* available in enterprise API only)	Web, mobile, and API; multimodal; freemium model	Consumer: subject to Google privacy policies, Commercial API: user data not used to train AI models
Microsoft Translator	Neural machine translation	Translation, Reading, Writing, Listening, Speaking (indirect)	No formal assessment or translation quality metrics	Web, mobile, and API; multimodal; freemium model	Temporary data processing only; subject to Microsoft privacy policies
PONS Translate Pro	Neural machine translation	Translation, Writing, Reading (indirect)	No formal assessment or translation quality metrics	Web, mobile, and API; paid, with free trial	User data not used to train AI models

Tool	Core Functionality	Language Skills Covered	Assessment Tools Availability	Accessibility	Data Privacy
ChatGPT	Multimodal conversational chatbot	Reading, Writing, Listening, Speaking, Translation, MTPE, Terminology management (indirect)	No formal assessment	Web, desktop, mobile, and API; multimodal, freemium model	Free: data used to train AI models (opt-out available); Paid: data not used for training
Google Gemini	Multimodal conversational chatbot	Reading, Writing, Listening, Speaking, Translation, MTPE, Terminology management (indirect)	No formal assessment	Web, mobile, and API; multimodal; freemium model	Consumer: subject to Google privacy policies, Paid: data not used for training
NotebookLM	Knowledge analysis assistance	Reading, Listening	Quizzes and flashcards for self-assessed preparation	Web and mobile; offline download; freemium model	User data not used to train AI models
Wordscope	Computer-assisted translation	Translation, MTPE, Writing (indirect)	Real-time QA checks, GPT-based revision and proofreading	Web; freemium model	Anonymisation features; Data not used for training
Matesub	Computer-assisted subtitling	Translation, Subtitling/Captioning, MTPE, Writing (indirect)	Real-time QA checks	Web; freemium model	Data encryption; Vague data storage and processing measures (Privacy Policy)
InterpretBank	Computer-assisted interpreting	Interpreting, Terminology management, Listening (indirect)	Flashcard-based self-assessment for glossary memorisation	Desktop and web; paid, with free trial	User data stored securely; data processed by third-party models for AI features

Tool	Core Functionality	Language Skills Covered	Assessment Tools Availability	Accessibility	Data Privacy
BoothMate	Computer-assisted interpreting	Interpreting, Terminology management	Flashcard-based self-assessment for glossary memorisation	Web; freemium model	Data processed by third-party models for AI features

Table 8. Comparative Overview of Functional Features in AI Tools for Translation and Interpreting

Tool	Skill Development	Autonomy	Engagement	Reliability	AI Transparency
DeepL	MTPE, vocabulary acquisition, written and oral communication in foreign languages	High (self-regulated use)	Drives engagement in communication and speeds up localisation workflows, but also poses risks of MT dependence and cognitive offloading	Variable accuracy mandates human supervision to meet professional standards	Opaque: results provided without linguistic rationale; customisation controls offered
Google Translate	MTPE, vocabulary acquisition, written and oral communication in foreign languages	High (self-regulated use)	Drives engagement in communication and speeds up localisation workflows, but also poses risks of MT dependence and cognitive offloading	Variable accuracy mandates human supervision to meet professional standards	Opaque: results provided without linguistic rationale; customisation controls offered
Microsoft Translator	MTPE, vocabulary acquisition, written and oral communication in foreign languages	High (self-regulated use)	Drives engagement in communication and speeds up localisation workflows, but also poses risks of MT dependence and cognitive offloading	Variable accuracy mandates human supervision to meet professional standards	Opaque: results provided without linguistic rationale; customisation controls offered
PONS Translate Pro	MTPE, vocabulary acquisition	High (self-regulated use)	Speeds up localisation workflows, but also poses risks of MT dependence and cognitive offloading	Variable accuracy mandates human supervision to meet professional standards	Moderate: alternative translations and contextual examples provided
ChatGPT	AI prompting, translation, MTPE, vocabulary acquisition, terminology management, text analysis	High (self-regulated use)	High (versatility), but over-reliance risk	Variable accuracy mandates human supervision to meet professional standards	Opaque: architecture, data, processes not fully disclosed; warns output may be inaccurate
Google Gemini	AI prompting, translation, MTPE, vocabulary	High (self-regulated use)	High (versatility), but over-reliance risk	Variable accuracy mandates human	Opaque: architecture, data, processes not fully

Tool	Skill Development	Autonomy	Engagement	Reliability	AI Transparency
	acquisition, terminology management, text analysis			supervision to meet professional standards	disclosed; warns output may be inaccurate
NotebookLM	AI prompting, text analysis, interpreting, listening	High (self-regulated use)	High (versatility), but risk of over-reliance and reduced engagement with original materials	Source grounding reduces hallucinations and enables verification	Google warns output may be inaccurate
Wordscope	Translation, MTPE, AI prompting	High (self-regulated use)	Enhances decision-making through multiple suggestions and speeds up localisation workflows, but also poses risks of MT dependence and cognitive offloading	Variable accuracy mandates human supervision to meet professional standards	Moderate: alternative translations and contextual examples provided through GPT integration; customisation controls offered
Matesub	MTPE, translation, subtitling	High (self-regulated use)	Speeds up localisation workflows, but also poses risks of MT dependence and cognitive offloading	Variable accuracy mandates human supervision to meet professional standards	Opaque: results provided without linguistic rationale; customisation controls offered
InterpretBank	Terminology management, interpreting	High (self-regulated use)	Supports interpreting and terminology work, but also poses risks of cognitive offloading	Variable accuracy mandates human supervision to meet professional standards	Moderate: limited AI features, but results provided without linguistic rationale

Tool	Skill Development	Autonomy	Engagement	Reliability	AI Transparency
BoothMate	Terminology management, interpreting	High (self-regulated use)	Supports interpreting and terminology work, but also poses risks of cognitive offloading	Variable accuracy mandates human supervision to meet professional standards	Moderate: limited AI features, but results provided without linguistic rationale

Table 9. Comparative Overview of Evaluation Criteria in AI Tools for Translation and Interpreting.

4.1. Machine Translation

Machine translation (MT) represents one of the earliest (and, hence, most mature) applications of AI in language, with contemporary neural and LLM-based systems achieving very high quality for a growing number of language pairs and domains.[5]

For translation professionals and students, understanding and mastering MT capabilities and limitations is instrumental for effective integration into professional workflows.

Google Translate, *Microsoft Translator*, and *DeepL* exemplify current best practices, combining neural machine translation (NMT) with large language model (LLM) post-processing. This dual architecture enables such systems to capture nuanced linguistic features (including idiomatic expressions, register variations, or contextual disambiguation) which challenged earlier rule-based and statistical engines.

The *DeepL* platform currently supports over 30 languages, with particularly strong performance for European language pairs, whereas *Microsoft Translator* and *Google Translate* offer a broader language coverage (more than 100 and 200 languages, respectively), with varying quality across language pairs and — for *Google Translate* — automated translation accuracy estimation scores (*Machine Translation Quality Prediction*, or MTQP) available only in enterprise tiers.

PONS Translate Pro integrates MT with editorially curated dictionary resources, providing linguistic context to support professional translation decision-making.

AI chatbots *ChatGPT* and *Google Gemini* offer translation capabilities embedded within conversational interfaces, enabling iterative refinement and explanation of translation choices.

A standalone, web-based interface named *ChatGPT Translate* was also introduced later on. The tool features a dual-pane layout similar to *Google Translate* and *DeepL*, while offering pre-set translation refinement options. When selecting these suggested prompts, users are redirected to the full *ChatGPT* conversational interface for further iterative adjustments. *ChatGPT Translate* currently lacks team-level integration infrastructure and collaboration capabilities, and the underlying model powering the tool has not been disclosed. Still, its launch highlights basic translation tasks as a frequent AI use case.

Nonetheless, while high-resource language pairs (such as English–Spanish, English–French, English–Chinese) typically achieve reliable results, low-resource pairs may produce outputs requiring significant post-editing.

[5] Kocmi, T. et al. (2025) 'Findings of the WMT25 General Machine Translation Shared Task: Time to Stop Evaluating on Easy Test Sets', *Proceedings of the Tenth Conference on Machine Translation*, 355–413, https://doi.org/10.18653/v1/2025.wmt-1.22.

General text also frequently outperforms specialised domains, with professional translation and localisation workflows increasingly adopting machine translation post-editing (MTPE) processes, where expert human translators review and refine MT output, rather than translating from scratch.

Professional standards mostly require human oversight of MT output, as final responsibility for translation quality in high-stakes settings cannot be delegated to AI systems. Confidentiality concerns may also arise when processing sensitive documents through cloud-based MT services.

Most MT tools offer additional features supporting professional workflows: customisable glossaries ensuring terminology consistency, document-level translation preserving formatting, register selection, and integration into translation management systems (TMSs) and computer-assisted translation (CAT) tools. Commercial API access normally ensures translated content data is not used for further model training.

4.1.1. Speech Translation

Real-time speech translation (converting spoken input in one language to spoken or written output in another) still represents one of the most technically demanding applications of language AI.

While recent advances have started to make it commercially viable (*Google Translate*'s *Speech Translation*, *Microsoft Translator*'s *Speech*, and *DeepL*'s *Voice* features introduced near instantaneous speech translation for online meetings and in-person conversations), professional simultaneous interpretation remains beyond current AI capabilities, as significant limitations — in accuracy and latency, as well as pragmatic competence and accountability — constrain applicability to date.

However, AI speech translation may expand language access in contexts previously unserved, where professional interpretation was not viable, and interpreter training programmes should familiarise students with such capabilities to enable informed judgement about when human interpretation remains critical.

4.2. Computer-Assisted Translation

Professional translation processes increasingly operate within technology-mediated workflows, where human expertise combines with computational resources and support. Computer-assisted translation (CAT) and terminology management environments support translation quality and consistency, while aiming at preserving human agency in translation decisions.

Contemporary CAT tools natively integrate multiple technologies: translation memory (storing previous translations for reuse), terminology databases (ensuring consistent terminology across projects), machine translation suggestions, and automated quality assurance checks.

As mentioned in the previous section, engines like *DeepL* offer direct integration with major CAT tools (such as *Trados* or *memoQ*), enabling translators to access MT drafts within familiar interfaces.

Wordscope is a web-based CAT tool, providing specialised support for translation with terminology databases drawing from authoritative sources, suggestions from multiple MT systems simultaneously, and native integration with a *ChatGPT*-based conversational dialogue box for supplemental AI assistance with rephrasing, explanations, alternatives, or proof-reading.

4.2.1. Audiovisual Translation

Audiovisual translation (AVT) encompasses subtitling, captioning, dubbing, and voice-over, and presents distinct challenges combining translation competence with multimodal awareness and technical constraints (e.g., conveying meaning within character limits, time constraints, or reading patterns, requiring strategic reformulation).

AI tools increasingly support and automate such technical aspects of AVT, whereas human expertise remains essential for quality assurance and cultural adaptation.

Matesub is a comprehensive AI-assisted subtitling platform, combining automatic speech recognition (ASR), machine translation (MT), text segmentation, and automated spotting for subtitle timing in an integrated interface. Users upload video content, the platform generates transcriptions, translates to target language(s), and produces timed subtitles and synchronised frames ready for review and refinement.

As with any other language AI technology, ASR and MT quality in AVT varies substantially across languages. These disparities may affect tool effectiveness and workflow, as time savings and efficiency gains from automation diminish if substantial manual correction is required.

For subtitle and caption creation, general-purpose AI tools *ChatGPT* and *Google Gemini* — with their ability to process multimodal input — can also assist with creative problem-solving and condensation strategies, including reformulating verbose dialogue to meet character-per-line constraints while preserving meaning and register, requesting alternative rendition options, or exploring equivalent idiomatic expressions in the target language(s).

The integration of MT into professional translation is fundamentally altering localisation work patterns, as translators reportedly[6] increasingly delegate initial translations to AI, focusing their expertise on post-editing, quality assurance, and complex processes beyond MT capability.

This shift entails evolving expert competencies: post-editing skills, MT output evaluation, and strategic judgement about when human translation from scratch remains preferable to AI editing.

[6] *Anthropic Economic Index: Insights from Claude 3.7 Sonnet* (2025), https://www.anthropic.com/news/anthropic-economic-index-insights-from-claude-sonnet-3-7.

Translation training programmes are, therefore, called to prepare students for hybrid workflows, rather than either fully manual or fully automated approaches exclusively.

4.3. Computer-Assisted Interpreting

Computer-assisted interpreting (CAI) tools support interpreters in preparation, terminology management, and real-time assistance, with the aim of 'augmenting' human interpreting capacity while preserving the interpreter's central cognitive and communicative role.

InterpretBank represents the most comprehensive and fully featured CAI solution to date, integrating glossary management, terminology extraction, AI-powered term retrieval, and real-time suggestions within a unified platform.

Its *Glossary* mode supports glossary construction from both source documents and AI-generated terminology lists, while the *Memory* module facilitates term memorisation through flashcard-based practice.

The *Digital Boothmate* feature provides live assistance during interpretation delivery, leveraging automatic speech recognition (ASR) to transcribe source speech and highlight pre-loaded terminology (with equivalent translations), figures, and acronyms as they appear.

This addresses a persistent challenge in simultaneous interpretation: managing specialised terminology, numbers, and other linguistic hurdles, while maintaining cognitive resources for the core interpreting task.

BoothMate (formerly known as *Interpreters' Help*) offers web-based glossary and interpreting workflow management, with AI-powered audio transcription supporting terminology extraction from recorded speeches and media files.

The platform's architecture facilitates collaborative glossary development and preparation support among interpreting teams, though it currently lacks real-time assistance features.

CAI tools hold promising pedagogical value for interpreter training programmes. For instance, *InterpretBank* introduced a speech generation feature to create customised training materials with natural-sounding voices through text-to-speech (TTS) technology and by incorporating terms from user glossaries to practice vocabulary.

However, the risk that AI-supported preparation may hamper the development of deep knowledge and competencies required for unassisted professional work must equally be addressed. Balanced curricula should, therefore, include both AI-assisted and 'traditional' preparation and delivery methods.

Moreover, in current tools, ASR accuracy varies by audio quality, speaker accent, and domain terminology, AI-generated glossaries require verification for accuracy and appropriateness, and real-time suggestions may require additional cognitive resources potentially competing with the primary interpreting processes.

Likewise, interpreters can leverage platforms like *ChatGPT*, *Google Gemini*, and *NotebookLM* for comprehensive assignment preparation and knowledge basis building, including background research, topic exploration, concept explanation, terminology extraction from source materials, glossary generation, practice speech generation, and even transcript analysis of record performances.

5. Conclusion

5.1. Findings

This comparative analysis has examined 26 AI-centred tools across three primary application domains: language learning, language teaching, and translation/interpreting.

Their functional capabilities, pedagogical value, and professional utility for stakeholders in higher education language programmes and vocational settings were evaluated. Findings reveal a rapidly evolving landscape, characterised by significant opportunities alongside continued challenges requiring careful consideration.

In institutional contexts, AI tools demonstrate considerable potential for **enhancing learner engagement** through diverse mechanisms. Gamification elements, present in platforms such as *Duolingo* or *Mini Course Generator*, with achievement tracking and progress visualisation, leverage behavioural psychology principles to sustain motivation over extended learning periods. Interactive features, including instantaneous pronunciation feedback (*Babbel*, *Duolingo*, *ELSA Speak*, *Talkpal*), conversational practice (*ChatGPT*, *Google Gemini*), customisable listening (*TTS Maker*, *ElevenLabs*), and adaptive difficulty adjustment (*Google Translate*), create adaptive and responsive learning environments.

However, engagement mechanisms vary considerably in terms of pedagogical soundness: while some applications foster active cognitive processing through meaningful interaction, others could facilitate superficial engagement or passive consumption of AI-produced content.

Table 10 below is a decision-oriented consultation aid underlining a pivotal insight — that beyond simple tool selection, user proficiency determines the most appropriate kind of interaction modalities with AI.

Learning outcome improvements remain mixed and context-dependent. Tools targeting specific skill development, like vocabulary acquisition (*Babbel*, *Duolingo*), pronunciation proficiency (*ELSA Speak*, *Talkpal*), or writing accuracy (*Grammarly*, *LanguageTool*), demonstrate advances when appropriately integrated into structured curricula (see *Figure 2* below). The immediate, personalised feedback these provide addresses a persistent limitation of 'conventional' instruction: the inability to offer individualised correction at scale.

For translation and interpreting students, terminology management tools (*InterpretBank*, *BoothMate*) and MTPE environments (*Wordscope*, *Matesub*) help develop competencies increasingly demanded by the labour market. Nevertheless, over-reliance on AI assistance providing answers without requiring effort risks eroding the acquisition of independent critical skills for complex work.

Proficiency	Approach	Representative Tools	Capabilities	Considerations
Elementary	Guided progression	Duolingo, Babbel, Grammarly, LanguageTool, ELSA Speak, Talkpal	Scaffolded path, systematic feedback, gamification sustaining motivation and engagement	Constrained cultural depth, requires judgement to verify AI output
Independent / Professional	Self-directed exploration	ChatGPT, Google Gemini, Perplexity, NotebookLM, Notion, Google Translate, Microsoft Translator, TTS Maker, ElevenLabs, MagicSchool, TeachAid, Coursebox, Mini Course Generator, LearningStudioAI, DeepL, PONS Translate Pro, Wordscope, Matesub, InterpretBank, BoothMate	Open-ended flexibility, task-oriented functionalities, workflow integration	Steep learning curve for specialised tools, over-reliance risks: requires self-regulation and critical assessment of AI output

Table 10. Orientation of AI Tools by Learner Proficiency and Autonomy.

AI integration is fundamentally **influencing pedagogical dynamics** and reshaping teacher-student relationships. Automation of routine tasks — educational content creation, lesson plan drafting, exercise generation, feedback provision (*TeachAid*, *MagicSchool*, *Coursebox*, *Mini Course Generator*, *LearningStudioAI*, as well as *ChatGPT*, *Google Gemini*, *NotebookLM*) — potentially frees educators to focus on higher-order pedagogical areas that AI cannot replicate.

However, such a redistribution of labour raises questions about professional identity, expertise validation, and the unique human components of linguistic and cultural education. Teacher 'AI literacy', that is the capacity to effectively prompt, verify, and scaffold AI usage, emerges as an essential competency to date, requiring systematic vocational development. The risk of *de-skilling*, with educators becoming dependent on AI for tasks they previously performed autonomously, warrants institutional attention.[7]

[7] Sarkar, A. et al. (2024) 'When Copilot Becomes Autopilot: Generative AI's Critical Risk to Knowledge Work and a Critical Solution', *Proceedings of the EuSpRIG 2024 Conference "Spreadsheet Productivity & Risks"*, 15–37, https://doi.org/10.48550/arXiv.2412.15030.

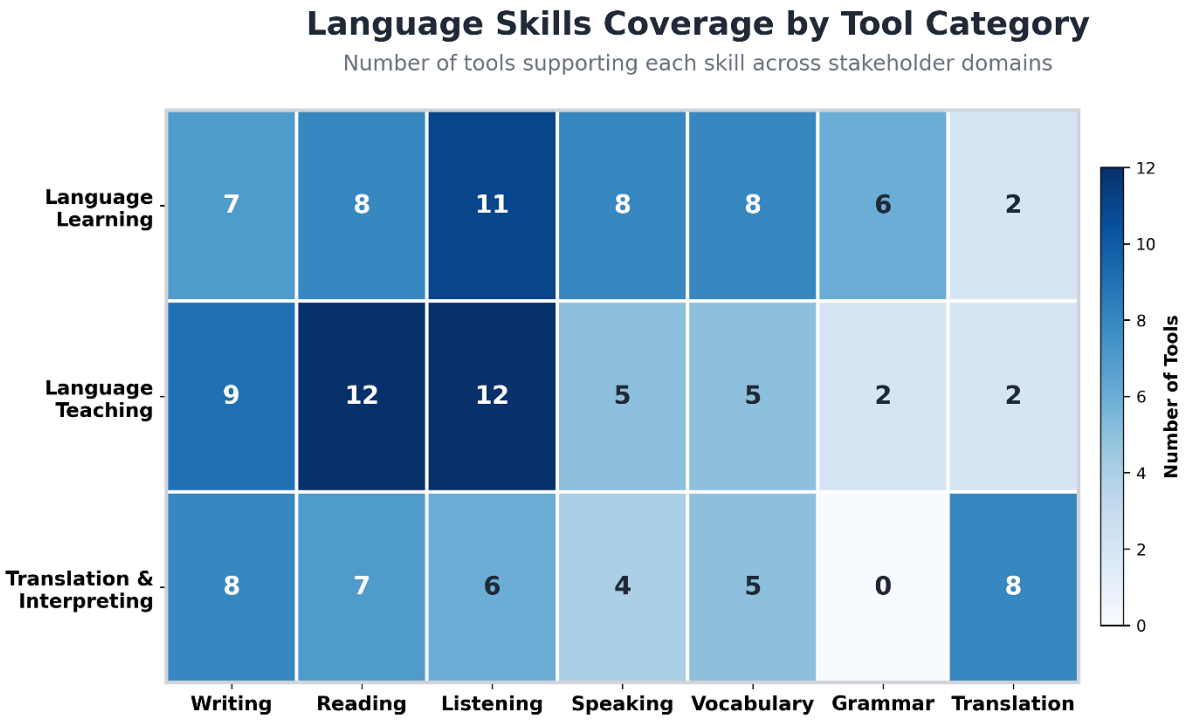

Figure 2. Tool Distribution by Language Skills.

The analysed tools demonstrate clear **professional functionality across language-related occupations**. Professional translators increasingly integrate MT systems (*DeepL*, *Google Translate*, *Microsoft Translator*, *PONS Translate Pro*), LLMs, and other applications of language AI (like speech-to-text and text-to-speech) into 'augmented' environments (*Wordscope*, *Matesub*, *InterpretBank*, *BoothMate*) and workflows, with documented productivity and accuracy gains.

Hence, for language professionals, the question appears to be no longer whether to engage with AI tools, but to what extent and how to do so effectively, while maintaining quality standards and professional ethics. Educational programmes must, therefore, prepare students not only for current professional practices, but for continued adaptation as AI capabilities evolve.

The traditional pillars of language knowledge, translation ability, and cultural expertise remain necessary, but further areas emerge as equally essential: AI literacy, as the capacity to successfully leverage AI tools through purposeful prompting, critical output evaluation, and awareness of system limitations, as well as suitability, that is the insightfulness to select appropriate multi-purpose resources for specific tasks and assess whether AI output meets functional requirements[8] — as visualised in *Table 11* below.

[8] Penet, JC, Moorkens, J., and Yamada, M. (eds.) (2026) *Teaching translation in the age of generative AI: New paradigm, new learning?*, Berlin, Language Science Press, https://doi.org/10.5281/zenodo.17580856.

Self-Learning (accessibility, flexibility, personalisation)	**Classroom** (collaboration, integration, assessment)	**Professional Practice** (reliability, transparency, productivity)
Duolingo, Babbel, Grammarly, LanguageTool, ELSA Speak, Talkpal, ChatGPT, Google Gemini, Perplexity, NotebookLM, Notion, Google Translate, TTS Maker, ElevenLabs, DeepL	MagicSchool, TeachAid, Coursebox, Mini Course Generator, LearningStudioAI, ELSA Speak, ChatGPT, Google Gemini, NotebookLM, Notion	ChatGPT, Google Gemini, Grammarly, LanguageTool, Google Translate, Microsoft Translator, DeepL, PONS Translate Pro, Wordscope, Matesub, InterpretBank, BoothMate, TTS Maker, ElevenLabs
Full availability and adaptability, requiring self-regulation and critical judgement	Integration and collaboration features in support of structured deployment, adoption entailing institutional and educator demands	Specialised functionality, task-specific scope limiting transferability across contexts

Table 11. Suitability of AI Tools by Integration Context.

5.2. Trends

The growing adoption of AI tools for everyday language-related tasks at the user level is also corroborated by findings from the *Copilot Usage Report*,[9] which lists language learning and translation among the primary topics of interaction with Microsoft's chatbots.

Learning personalisation through adaptive algorithms represents a defining trend across educational AI tools. Systems can adapt learning pathways and target instruction by adjusting content pacing, exercise difficulty, and focus areas based on individual learner performance and needs, rather than assumed levels. Machine learning models tailored to learner data enable increasingly sophisticated predictions about optimal learning sequences and intervention timing.

However, the pedagogical assumptions embedded in such algorithms, often prioritising measurable outcomes over holistic development, deserve critical examination to prevent narrowing learning experiences to quantifiable competencies while neglecting cultural understanding, linguistic nuance, and communicative spontaneity.

The integration of virtual, augmented, and mixed reality with AI assistants and language tools for **immersive and interactive educational environments** represents a significant — but still nascent —

[9] Costa-Gomes, B. et al. (2025) *It's About Time: The Copilot Usage Report 2025*, https://microsoft.ai/wp-content/uploads/2025/12/What_people_do_with_Copilot-8.pdf.

developmental trajectory. Current implementations focus primarily on contextualised vocabulary presentation and simulated conversational environments, with restrained adoption.

Future possibilities for sustained, immersive language learning could allow students to practice target language use in simulated real-world contexts, whereas for interpreter training immersive technologies could simulate realistic practice experiences in public service or conference settings.

Translation accuracy from both neural systems and LLMs continues improving, with quality approaching human level for certain language pairs and text types. The integration of AI language models into translation workflows enables more nuanced handling of context, register, and stylistic variation, while these advances complicate traditional translation pedagogy: if machines can produce acceptable translations, the value proposition of human translation training shifts towards quality assurance, cultural adaptation, critical problem-solving, and the judgement and accountability calls which remain beyond algorithmic capability.

Real-time custom feedback is one of the most pedagogically significant AI resources. Tools providing immediate grammar correction, writing improvement, and pronunciation assessment with explanations and suggestions address the feedback bottleneck that has historically constrained widespread, detailed language learning, including increasingly sophisticated feedback on dimensions like register selection, cultural sensitivity, and pragmatic appropriateness.

In translation and interpreting training, AI assessment applications promise systematic feedback on student work, though concerns persist about algorithmic evaluation of inherently context-dependent translation quality.

5.3. Considerations

The integration of AI tools into both language education and vocational contexts raises substantial ethical questions requiring ongoing critical reflection, institutional attention, and policy development. These considerations span equity, transparency, data and model governance, professional ethics, and academic integrity.[10]

AI language tools necessarily process user data (written productions, voice recordings, learning patterns, error profiles). This analysis reveals that while some tools offer robust compliance with **data protection** framework standards and explicit non-training commitments, others operate under broader privacy policies which may permit data use for improving proprietary commercial models. Institutional deployment, therefore, requires careful evaluation of data processing agreements, with the distinction between free and paid tiers — where the latter typically offer stronger

[10] European Commission: Directorate-General for Education, Youth, Sport and Culture (2026) *Guidelines on the ethical use of artificial intelligence and data in teaching and learning for educators*, Publications Office of the European Union, https://data.europa.eu/doi/10.2766/7967834.

confidentiality guarantees at higher costs — creating equity implications for resource-sensitive institutions.

AI tool performance, then, varies dramatically across languages, with high-resource languages (primarily English, followed by a handful of other major languages) receiving substantially better support than lower-resource languages. This **language coverage disparity** reflects training data availability and commercial incentives favouring dominant languages or misrepresenting diverse perspectives.

For language education programmes, tool selection must account for such gaps which may disadvantage certain learner communities. The concentration of AI development in restricted contexts raises concerns about **cultural bias** in content generation and pedagogical assumptions too, potentially privileging specific communicative norms.

Machine learning systems inevitably encode **biases** present in their training data. Socio-economic, ethnic, or gender stereotyping and assumptions may reproduce in generated content and pedagogical materials. Educators are called to approach and review algorithmic outputs critically for appropriateness and accuracy.

The 'freemium' model — with tiered access where advanced features, higher usage limits, and stronger privacy protections require paid subscriptions — prevails among AI tools, creating **cost barriers**. This structure risks exacerbating **digital gaps and educational inequalities**, with well-resourced institutions or affluent individual learners accessing superior tools. Institutional licensing addresses some equity concerns, but budget allocation may compete with other educational priorities.

Additionally, **accessibility support** (screen reader compatibility, alternative input methods, or other accommodations) for learners with disabilities varies considerably across tools, with some presenting significant blocks.

Due to their very architecture, AI tools mostly operate as 'black boxes' providing outputs without explaining underlying processes. For educational applications, this **opacity in transparency and explainability** poses pedagogical challenges: without such a reasoning visibility, users cannot understand failure modes and may accept AI judgements uncritically, rather than developing independent awareness.

Tools offering some transparency compensation (source attribution, alternative suggestions, confidence indicators) support more informed user engagement, though full explainability remains technically elusive for neural network-based systems.[11]

[11] Lindsey, J. et al. (2025) *Tracing the thoughts of a large language model*, https://www.anthropic.com/research/tracing-thoughts-language-model.

Educational institutions increasingly necessitate strong **policies governing AI** tool use in their programmes (although the rapid pace of AI development challenges the demands of such a framework), including tool lists balancing pedagogical value against privacy and security, vocational development ensuring educator AI mastery, infrastructure for institutional tool deployment, guidelines for appropriate AI use in assessed work, and procedures for evaluating emerging tools.

AI assistance in writing and translation tasks also questions conventional conceptions of **academic integrity**, especially in language education.

When polished text or accurate translations can be generated through AI, assessment design must evolve to evaluate learning processes and academic competencies that AI cannot replicate, from oral proficiency and real-time communication to critical analysis of AI outputs and linguistic problem-solving, rather than merely their products.

Since AI tools are not inherently safe or dangerous but their risk profile rather depends on how they are used, *Table 12* below maps interaction patterns to exposure levels.

Risk Dimension	Lower Risk	Higher Risk	Mitigation Strategies
Accuracy and Hallucinations	Source grounding or task specificity to constrain AI output	Open-ended generation	Awareness of AI limitations; verification protocols; cross-referencing
Data Privacy	Explicit commitment to non-training of AI models	Free general-purpose tools where user data may be retained to train AI models; opaque policies	Information processing agreements; clear guidance on data exposure
Linguistic and Cultural Biases	Human-curated content	Unreviewed AI-generated content in lower-resource languages or culturally sensitive contexts	Transparent bias acknowledgement; expert validation of AI output; source diversification
Integrity	Disclosed AI usage for preparation and assistance	Undisclosed AI usage for full-fledged production	Output analysis; redesigned assessments
Over-reliance and De-skilling	Draft revision; complementary support to independent work	Task delegation without engagement; independent effort replacement; uncritical acceptance of AI suggestions	Process assessment design; output quality reflection

Table 12. Risk Profiles in AI Tool Integration.

Likewise, for translation and interpreting programmes and practice, AI integration raises specific **professional ethical considerations**. Professional obligations regarding confidentiality and AI use disclosure, quality assurance responsibilities, and evolving industry standards require explicit training in adjacent competencies and critical evaluation of machine output.

Empirical research from AI products and safety laboratory Anthropic demonstrated that in the real world, educators focus on *augmentation* (human–AI collaboration) by task iteration[12] and university students use AI systems as academic companions,[13] offering evidence that AI can enhance rather than replace pedagogical work.

On the contrary, as already mentioned in *Section 4.2*, translators and interpreters show high *automation* rates with a directive behaviour, meaning that AI models are delegated to complete tasks with minimal human involvement afterwards.[14] Signals of augmentation can be found in revision, editing, and validation tasks, with higher levels of human oversight, learning, and task iteration.[15]

While data exposure considerations and transparency about AI use concern all stakeholder groups, it is worth noting that AI is already ever more embedded in digital tools used everyday across educational and professional settings. As general-purpose models become more prevalent, the pondered assessment of their appropriateness for specific purposes pivots to educators, professionals, and institutions employing them.

Learning critical AI literacy, preserving professional judgement, ensuring equitable and fair tool access, and developing sensitivity and awareness that AI may perpetuate systematic stereotypes and biases thus become essential requirements to navigate an evolving and challenging landscape.

5.4. Future Directions

The trajectory of AI in language education and professional services points towards increasingly integrated and pervasive applications, with developments meriting anticipation and adequate preparation. Forthcoming AI systems will likely offer even more granular adaptation to individual profiles, incorporating not only performance data or learning style preferences, but also longer-term paths and even affective states — though this entails addressing novel concerns in privacy, pedagogical validity, and the risk of algorithmic determinism constraining agency.

---

[12] Bent, D. et al. (2025) *Anthropic Education Report: How educators use Claude*, https://www.anthropic.com/news/anthropic-education-report-how-educators-use-claude.

[13] Handa, K. et al. (2025) *Anthropic Education Report: How university students use Claude*, https://www.anthropic.com/news/anthropic-education-report-how-university-students-use-claude.

[14] *Anthropic Economic Index: Insights from Claude 3.7 Sonnet* (2025), https://www.anthropic.com/news/anthropic-economic-index-insights-from-claude-sonnet-3-7.

[15] Appel, R. et al. (2026) *Anthropic Economic Index report: economic primitives*, https://www.anthropic.com/research/anthropic-economic-index-january-2026-report.

The boundaries between text, speech, image, and video processing are dissolving in frontier AI systems. Language education tools will increasingly handle authentic multimodal communication, prospectively integrating prosody and gesture alongside pure linguistic content, and enabling practice in realistic communicative contexts spanning multiple, holistic modalities.

While current AI tools are still limited in live support, incremental advances in speech recognition, low-latency translation, contextual understanding, and predictive terminology, suggest expanding AI assistance in real-time — with humans retaining responsibility for meaning, metalinguistic awareness, and cultural mediation.

Educational institutions will increasingly require dedicated AI infrastructure, including equally distributed connectivity, licensing agreements, integration with learning management systems, local deployment options for sensitive applications, and governance frameworks. The current landscape of individual tool subscriptions or ad-hoc adoption will probably evolve towards more systematic institutional approaches.

Language education programmes must hence prepare for vocational contexts where AI collaboration is not exceptional. Such a preparation extends beyond tool proficiency to encompass critical AI knowledge, quality assurance trustworthiness, technology misuse responsibility, ethical reasoning about AI, and the distinctly human skills (like creativity, cultural sensitivity, interpersonal communication, collaboration) beyond algorithmic reach.

The tools analysed in this report represent a snapshot of a rapidly evolving field and are intended to provide a foundation for informed decision-making, while acknowledging the dynamic nature of AI technology. While other specific platforms will emerge and evolve, the fundamental questions about the nature of language learning, the role of human expertise, the ethics of algorithmic mediation, and the future of language professions will persist and even intensify.

Critical engagement with these — guided by sound pedagogical values and informed by empirical evidence — remains essential, as AI increasingly influences how languages are learned, taught, and professionally practised.

Emerging regulatory provisions in Europe and worldwide may also increasingly shape AI tool development and deployment in educational contexts, as institutions could actively engage in policy discussions shaping the future landscape around shared values — including through collaborative projects like AI-NYELV, aimed at promoting and validating ethical AI integration and reflective engagement.

List of Tables

Figures

Key Abbreviations and Acronyms

AI	Artificial Intelligence
API	Application Programming Interface
ASR	Automatic Speech Recognition
AVT	Audiovisual Translation
CAI	Computer-Assisted Interpreting
CAT	Computer-Assisted Translation
CEFR	Common European Framework of Reference for Languages
EEA	European Economic Area
GDPR	General Data Protection Regulation
GPT	Generative Pre-trained Transformer
LLM	Large Language Model
LMS	Learning Management System
MT	Machine Translation
MTPE	Machine Translation Post-Editing
MTQP	Machine Translation Quality Prediction
NMT	Neural Machine Translation
QA	Quality Assurance
T&I	Translation & Interpreting
TMS	Translation Management System
TTS	Text to Speech
WP	Work Package