Check for updates

Received: 30 November 2025 | Accepted: 12 August 2026

DOI: 10.1111/flan.70082

ORIGINAL ARTICLE

ACTFL Language Connects WILEY

Co-existing and co-constructing with AI: Beliefs, practices, and professional readiness of online language coaches

Javier Muñoz-Basols[1] | **Tripp Strawbridge**[2]

[1]Faculty of Medieval and Modern Languages, University of Oxford, Oxford, UK

[2]Modern Languages and Literatures, Kenna Hall 112 D, College of Arts and Sciences, Santa Clara University, Santa Clara, California, USA

Correspondence
Javier Muñoz-Basols, Faculty of Medieval and Modern Languages, 41 Wellington Sq. University of Oxford, Oxford OX1 2JF, UK.
Email: javier.munoz-basols@mod-langs.ox.ac.uk

Abstract

This study examines how online Spanish language coaches, that is, trained native speakers who facilitate authentic conversational practice, perceive and use artificial intelligence (AI) in virtual language teaching. While previous research has predominantly focused on classroom teachers, we use a mixed-methods design that includes a questionnaire and open-ended responses from 96 coaches based in 15 Spanish-speaking countries. We examine their familiarity with AI, attitudes toward its use, and views on how it may complement human interaction in language learning. Coaches see clear benefits in using AI for lesson planning, personalization, and idea generation, but they also express concerns about data privacy and ethics, misinformation, reduced critical thinking, and the loss of emotional connection and cultural authenticity. Regression analyses show that familiarity with AI and comfort using it are the strongest predictors of perceived usefulness and willingness to integrate this technology into practice. The findings suggest that professional development should strengthen AI literacy while preserving the interpersonal and intercultural dimensions of language learning. Clear institutional guidance is also needed on AI

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
© 2026 The Author(s). *Foreign Language Annals* published by Wiley Periodicals LLC on behalf of ACTFL.

use and pedagogical approaches that integrate AI without diminishing meaningful human interaction.

KEYWORDS
Artificial Intelligence (AI), ethics in educational technology, language coaching, language education, personalized learning

1 | INTRODUCTION

The concept of *coaching*, common in fields such as sports, leadership, and psychology, is typically understood as a personalized, goal-oriented process designed to enhance performance, motivation, and self-efficacy. In recent years, this model has been adapted to language education, giving rise to *language coaching* (Cowden et al., 2022). Unlike traditional instruction, language coaching emphasizes learner autonomy, reflective practice, and non-directive facilitation (Thomson, 2013; Wielgolawski, 2011). Its distinctive feature is positioning the learner as an active agent supported by a coach who acts as "a guide or trainer who makes decisions about how the player or student performs" (Ceo-Francesco, 2015, p. 46). Coaching principles have been applied across diverse contexts: in medicine, to help international healthcare professionals overcome cultural and linguistic barriers (Cowden et al., 2022); and in mental health, where US military veterans with PTSD served as English coaches for learners abroad, enhancing both linguistic outcomes and veterans' well-being (Piette et al., 2024).

Within formal education, language coaching has increasingly complemented classroom instruction. Digital platforms such as LinguaMeeting, TalkAbroad, and Boomalang, widely adopted across US universities, connect students with trained native-speaking coaches who facilitate authentic (Abing, 2018) and naturalistic (Strawbridge, 2025) conversational practice in Spanish and other languages. These platforms are built on the principle that meaningful communicative interaction, supported by the flexible, supportive role of the coach, can enhance intercultural competence, reduce foreign language anxiety, and promote learner engagement (Marull & Kumar, 2020). This shift shows how coaching principles are being embedded into higher education curricula, with classroom instruction and real-world conversational practice increasingly overlapping.

The accelerating presence of artificial intelligence (AI) in education adds a new dimension to this field. Emerging technologies, including AI-driven avatars embedded in language-learning platforms such as Duolingo (Voss, 2024) or Praktika (Tulasi & Rao, 2023), are increasingly designed to simulate conversational interaction and provide automated guidance (Moorhouse & Wong, 2025; Pérez-Núñez, 2024). While scholarship has examined teachers' integration of AI in English language teaching and the affordances and risks of generative AI (Hınız, 2024; Kim et al., 2025), far less is known about how coaches themselves perceive such tools. Existing studies suggest that teachers often view AI with ambivalence: they recognize its potential for personalization and adaptive feedback but express concerns about ethical implications, learner dependency, and diminished creativity (Crompton et al., 2024; Hınız, 2024). Yet the perspectives of coaches, whose pedagogical identity is grounded in conversational practice and individualized learner support, remain largely absent from these discussions.

As language coaching becomes more firmly integrated into university curricula through digital platforms, understanding how coaches perceive AI becomes a practical imperative.

19449728, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/flan.70082 by NICE, National Institute for Health and Care Excellence, Wiley Online Library on [28/10/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Research that captures coaches' perspectives is therefore essential for informing the design of professional development programs, guiding institutional policy on AI adoption, and ensuring that technological innovation serves the pedagogical goals of language coaching. This mixed-methods study examines how online language coaches' familiarity and perceptions of AI shape their beliefs and willingness to integrate AI tools, their evaluation of AI's pedagogical usefulness, ethical implications, and long-term role in education, and the professional needs and challenges they identify for effective implementation. By centering coaches' perspectives, the study contributes new evidence on how AI is being incorporated into language learning.

In this study, AI is conceptualized as a broad socio-technical phenomenon that shapes educational contexts, professional identities, and expectations about teaching and learning, whereas "AI tools" refers to specific generative applications (e.g., ChatGPT) whose pedagogical implications are still emerging. This distinction mirrors the framing adopted by Moorhouse and Kohnke (2024), who treat AI as a systemic force influencing language teacher education while foregrounding generative AI tools as the immediate and uncertain sites of pedagogical change. This distinction guided both the questionnaire design and the analytical approach, enabling the study to capture how macro-level perceptions of AI intersect with situated professional decision-making.

2 | LITERATURE REVIEW

2.1 | Online language learning coaching: Key differences from traditional instruction

The concept of *language coaching* has its roots in broader traditions of self-directed learning and online coaching, where the role of the coach is defined less by transmission of knowledge than by the facilitation of learner autonomy (Zablotska, Olendr and Tsar, 2023). Rather than delivering a fixed curriculum, language coaches support learners in designing personalized programs, setting goals, and reflecting on progress. In this sense, coaching becomes a scaffolded process that enhances responsibility, agency, and metacognitive awareness in language learning. Within language education, coaching has become closely intertwined with virtual exchanges and telecollaboration (Strawbridge, 2021). However, while language coaching centers on guided interaction with trained facilitators to provide targeted practice and feedback, telecollaboration relies on reciprocal, task-based peer interaction between learners in different locations.

Ceo-Francesco (2015) documented how online conversation programs created "a standards-based approach to synchronous face-to-face interaction, integrating language and culture in a natural, authentic context" (p. 40). Slusarek (2024) illustrated this through a LinguaMeeting student whose course integrated 30-min coaching sessions with trained native speakers framed by the *Can-Do Statements* (NCSSFL & ACTFL, 2017); the student emphasized that synchronous interactions provided a more accurate measure of communicative competence than online testing alone.

Against this backdrop, a key question arises: what distinguishes language coaching from traditional language teaching? Chaika (2023) identifies five core principles of language coaching: facilitation (guiding rather than directly instructing), learner empowerment (transferring goal-setting and decision-making to students), personalization (adapting paths to individual needs and styles), goal-setting and self-regulation (fostering autonomy and strategic learning), and feedback and reflection (iterative performance assessment and refinement).

19449728, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/flan.70082 by NICE, National Institute for Health and Care Excellence, Wiley Online Library on [28/10/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Together, Chaika, (2023) principles operationalize Ceo-Francesco, (2015) earlier conceptualization of language coaching. Both position the coach as a facilitator who strategically shapes performance through authentic, culturally embedded interaction. Culture is not ancillary but central: linguistic material is embedded in the practices, products, and perspectives of the coach's country, and learners engage in authentic exchanges in a supportive environment where risk-taking is encouraged. Marull and Kumar (2020) showed that synchronous telecollaborative coaching in online Spanish courses helped learners apply skills in real-life situations through sustained interaction with native speakers; Guerrero-Rodriguez (2023) similarly found that interaction with Mexican peers helped heritage speakers move from linguistic insecurity to confidence. Spanish thus stands out as a central testing ground for language coaching in virtual environments, given its widespread adoption across US universities through institutionally embedded digital coaching platforms.

2.2 | Coach-centered language learning platforms

Over the past two decades, a growing number of digital platforms have partnered with universities to facilitate authentic conversational practice (Berninger-Schäfer, 2022; Jones, 2025). One of the earliest and most established initiatives is LinguaMeeting (www.linguameeting.com), founded in the aftermath of Hurricane Katrina in 2005. Similarly, TalkAbroad (established in 2010; www.talkabroad.com) and Boomalang (founded in 2014; www.boomalang.co) have become widely adopted platforms, collaborating with hundreds of institutions globally. These platforms reflect a pedagogical shift away from simulated classroom dialogues and textbook-driven practice, toward authentic, global interaction. Using these digital infrastructures, they extend the communicative turn in language teaching and support learner autonomy, intercultural competence, and flexible curriculum design. Their growing prominence reflects the emergence of hybrid learning models in which face-to-face instruction is complemented by virtual exchanges.

2.3 | Generative AI and language coaching

The emergence of Generative Artificial Intelligence (GenAI) has been widely described as a "game changer" in language education, marking a shift as profound as earlier milestones in Computer-Assisted Language Learning (CALL) (Chapelle, 2025). Unlike previous technologies that merely stored, retrieved, or corrected language, GenAI can generate original, contextually contingent, and grammatically accurate sentences in real-time interaction with users (Chapelle, 2025; Clementi & Jiménez López, 2025). This capability has left teachers both excited and anxious. On one hand, GenAI offers unprecedented opportunities for scaffolding input, feedback, and interaction. On the other, it raises existential questions about whether traditional classroom practices remain necessary when machines can emulate many of their outcomes.

Teachers' responses are shaped by pedagogical concerns and by perceptions of ethics, professionalism, and autonomy. Long et al. (2024) argue that teachers initially experience a *novelty effect* (Muñoz-Basols & Lvova, 2025) when engaging with AI, but their longitudinal study of workflow customization shows that familiarity strengthens perceptions of usefulness, highlighting the importance of sustained engagement and professional ownership (Tribes & Masid Blanco, 2025).

Recent scholarship has begun to map this area. Mills et al. (2025) demonstrated how AI can function as a pedagogical partner for foreign language writing when carefully scaffolded by

19449720, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/flan.70082 by NICE, National Institute for Health and Care Excellence, Wiley Online Library on [28/10/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

instructors. Tacelosky et al. (2025) found that GenAI tools sparked creative engagement in a Spanish class but revealed reliability gaps requiring instructor guidance, while Kessler et al. (2025) reported that L2 writers using GenAI for Spanish expressed ethical concerns about accuracy and authenticity. Liu and Reinders (2025) likewise found that AI chatbots produced mixed effects; while some students reported heightened confidence and engagement, others considered the interactions artificial or distracting. Li et al. (2025) further showed how world language teachers navigate complex negotiations between professional identity and ChatGPT adoption. Collectively, these studies highlight that AI integration in language education is shaped not only by technological affordances but by the beliefs, identities, and professional readiness of practitioners.

Wu and Yu (2023) and McGrath et al. (2025) further demonstrated that AI chatbots can enhance student performance in comprehension and writing tasks, but effectiveness depends heavily on teachers' ability to integrate the technology into lesson design, making professional readiness a central determinant of impact. The present study examines this dynamic from the underexplored perspective of language coaches.

A separate line of scholarship has begun to explore AI itself acting as a *language coach*. Crompton et al. (2024) noted that conversational AI, while not equivalent to natural human interaction (Voss & Waring, 2025), can function as an interactive partner in learners' development, blurring the line between human-led coaching and AI-mediated practice. Research indicates that AI can personalize instruction through adaptive feedback and scaffolding, creating new pathways for learner engagement (Kim et al., 2025). Since the launch of ChatGPT in late 2022, applied linguistics has moved from novelty toward critical reflection and sustained integration, with AI tools increasingly viewed as resources for pedagogical innovation (Chapelle, 2025).

Teachers increasingly use AI tools for lesson preparation, assessment, and feedback, but remain cautious about over-reliance, prompting calls for what Moorhouse and Wong (2025) describe as "professional-GenAI-competence": a multidimensional readiness encompassing technological skills, pedagogical integration, and awareness of risks to learner well-being.

These changes are particularly salient in Spanish language education. Platforms such as LinguaMeeting, TalkAbroad, and Boomalang, originally designed to deliver human-led coaching, now serve as experimental spaces for hybrid models where AI agents and human coaches co-exist (Manierre et al., 2025), enriching scaffolding and learner support without eroding the interpersonal and cultural dimensions central to language learning. This challenge is pedagogical as much as technical, requiring investigation into how language coaches perceive AI, that is, their beliefs, practices, and professional readiness.

3 | RESEARCH METHODS

3.1 | Research questions and design

The aim of this mixed-methods study was to examine the perspectives of online language coaches on AI in language education by addressing the following two dimensions:

RQ1. How do online language coaches' familiarity with and general perceptions of AI relate to their beliefs about its potential and their intention to integrate AI tools into virtual exchange–based language learning?

RQ2. How do online language coaches evaluate the pedagogical utility and long-term integration of AI tools in the context of language instruction?

19449728, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/flan.70082 by NICE, National Institute for Health and Care Excellence, Wiley Online Library on [28/10/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

3.2 | Research context and participants

This study involved 96 native Spanish speakers working as online language coaches for a major digital platform, 95 of whom were based in Spanish-speaking countries and one in the United States, with data collected in the spring of 2025. The platform was selected for its leading role in US Spanish language teaching, and partnered with the research team to explore how AI integration might inform coaches' professional development.

Participants (27 male, 69 female) represented a broad geographical distribution, including Argentina (6), Bolivia (3), Chile (3), Colombia (5), Dominican Republic (2), Ecuador (11), El Salvador (1), Guatemala (21), Honduras (8), Mexico (5), Nicaragua (5), Panama (3), Peru (11), Spain (8), United States (1), and Venezuela (3). Participants' age ranged from 18 to 43, with a mean and median age of 28. Participants' experience as language coaches (i.e., facilitating conversational practice through the digital platform) ranged from 1 month to 13 years, with an average of 3.5 years and a median of 3 years. It should be noted that some coaches may also have held formal teaching positions; however, the experience reported here refers specifically to their work as online language coaches. Participation in the study was voluntary, and all respondents received a modest honorarium. Informed consent was obtained from all participants, and ethical approval for the study was obtained from the platform.

3.3 | Procedures: Questionnaire and data collection

The questionnaire, distributed via Qualtrics in spring 2025, comprised 53 items, 33 close-ended and 20 open-ended, and was organized into four main sections preceded by a preliminary demographic section (age, gender, country, language repertoire, educational background, and coaching experience). The instrument was structured to move from coaches' general beliefs about AI as a broad concept (Section 1: familiarity, perceived societal impact) toward progressively more situated evaluations of AI in language teaching and coaching practice (Section 2–4: perceived usefulness, comfort, and actual use of AI tools). Developed as an exploratory instrument to capture language coaches' emerging beliefs, practices, and professional concerns regarding AI, its design is consistent with prior questionnaire-based research on educators' acceptance of AI in language education (Cabero-Almenara et al., 2024; Viberg et al., 2025) and studies examining language teachers' perceptions of generative AI tools (Belda-Medina & Kokošková, 2024; Butler & Jiang, 2025; Ma et al., 2024). Prior to full deployment, the instrument was piloted with a small group of online language coaches ($n = 5$), enabling refinement of item wording, scale anchors, and overall interpretability (the full bilingual Spanish–English questionnaire is available via IRIS [https://www.iris-database.org/details/75Yju-v1VGE]).

Section 0 (General profile) gathered demographic and professional background information (e.g., age, country, teaching experience). Section 1 (General beliefs about AI) included 6 Likert items on familiarity, perceived societal impact, frequency of use, and expectations about AI's future influence (e.g., "How familiar are you with the concept of Artificial Intelligence?"). Section 2 (General perceptions of AI in language teaching) comprised 7 items on integration interest, perceived pedagogical value, student perceptions, professional comfort, and future importance of AI (e.g., "How important do you think AI will be in the future of language teaching?"). Section 3 (Perceptions of AI's usefulness in language teaching) included 7 items on the usefulness, effectiveness, and trustworthiness of AI for instructional tasks such as lesson planning, feedback, assessment, and learner motivation (e.g., "How useful do you think AI is for

19449728, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/flan.70082 by NICE, National Institute for Health and Care Excellence, Wiley Online Library on [28/10/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

providing instant feedback to students?"). Section 4 (Attitudes and feelings about AI integration) contained 6 items on frequency of use, satisfaction, ethical considerations, training adequacy, and likelihood of recommending AI to colleagues (e.g., "How often do you use AI as a tool in your online sessions?"). Section 5 (Perception of the Spanish produced by ChatGPT), not included in the present analysis, contained 7 items on the grammatical accuracy, dialectal variation, coherence, and overall quality of AI-generated Spanish (e.g., "How would you rate the grammatical accuracy of the Spanish produced by AI chatbots?").

Together, these close-ended items provided the quantitative basis for the analyses, while the accompanying open-ended questions allowed participants to contextualize and elaborate on their responses. Closed-ended items yielded standardized measures for statistical analysis, for example, "Q2. How familiar are you with the concept of Artificial Intelligence (AI)?" or "Q23. How useful do you think AI is for creating engaging learning materials?" while open-ended questions elicited richer accounts of participants' views.To examine the diverse understandings Spanish language coaches may hold regarding ChatGPT and AI in education, the qualitative component was designed within a phenomenographic framework (Derakhshan & Ghiasvand, 2024). This approach, which emphasizes the range of ways individuals experience and interpret a shared phenomenon (Hajar, 2021; Stolz, 2020; as cited in Derakhshan & Ghiasvand, 2024), was particularly suited to investigating the perspectives of online coaches.

Open-ended items, included throughout Sections 1–4, elicited detailed reflections on coaches' beliefs, perceptions, and attitudes toward AI (e.g., *Q12. How do you think AI will shape the future of your field or profession?*; *Q43. How do you envision the role of AI in education, particularly in language teaching, over the next 10 years?*).

3.4 | Data analyses

3.4.1 | Mixed methods research

This mixed methods research design was guided by two complementary purposes: complementarity and triangulation. Complementarity is shown when the different components of the study, quantitative and qualitative, address different research questions, data, and analyses. In this study, RQ1, which examined how coaches' familiarity and general perceptions of AI relate to their beliefs about AI's potential and their intention to integrate AI tools, was addressed through quantitative methods (multiple linear regression on close-ended Likert items). These analyses complemented qualitative analyses of open-ended questionnaire responses, which served to answer RQ2 on coaches' evaluations of the pedagogical utility and long-term integration of AI tools. The design also included triangulation: close-ended Likert items and open-ended responses both probed coaches' views on AI in language teaching, allowing convergent or divergent patterns across data types to be examined.

3.4.2 | Quantitative analyses

To address RQ1, multiple linear regression was used to examine how coaches' familiarity and general perceptions of AI relate to their beliefs about the potential of AI tools and their intention to integrate AI into their language coaching. Six independent variables were considered, all of which were posed as 5-point Likert-type items (Table 1). Likert-type responses

19449728, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/flan.70082 by NICE, National Institute for Health and Care Excellence, Wiley Online Library on [28/10/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 1 Independent variables for linear regression models (familiarity and perception).

Variable name	Question on questionnaire
"Familiarity"	How familiar are you with the concept of AI?
"Society influence"	Do you perceive AI to be a positive or negative influence on society?
"General AI use"	How frequently do you interact with AI tools? (general, not specific to language instruction)
"Coach comfort"	How comfortable do you feel working with AI tools in the professional context?
"Student perception"	How do you think students perceive AI tools for language learning?
"Future importance of AI"	How important do you think AI tools will be to the future of language instruction?

(both dependent and independent) were treated as continuous variables, as they, despite not being numerical in nature, provided respondents with a range of responses that could be fairly considered as evenly distributed (e.g., "totally disagree," "somewhat disagree," "neither agree nor disagree," "somewhat agree," and "totally agree"). This decision was further justified by the normality of the data and the observation that the data met all other assumptions for the regression analysis.

The model distinguishes between general orientations toward AI and tool-specific experience, allowing us to examine whether support for AI in principle translates into willingness to integrate AI tools in practice.

Although only the six items listed in Table 1 were used to operationalize the independent variables in the regression models, the remaining close-ended questionnaire items served complementary purposes. These items were included to contextualize the participant sample (e.g., professional background and prior AI exposure), capture relevant covariates that may shape responses, and measure related constructs that support interpretation of the main variables (e.g., attitudes, concerns, and perceived instructional value of AI).

Under consideration for RQ1 was the influence of these six independent variables on three dependent variables (likewise measured as 5-point Likert-type items), relating to coaches' beliefs about the potential of AI, their intention to integrate AI tools into their language coaching, and their current use of AI tools in language coaching. These variables are listed in Table 2.

A separate multiple linear regression model was run to address the effect of the selection of independent variables on each dependent variable. In preparation for the regression models, two independent variables ("General AI use" and "Coach comfort") were observed to be highly correlated (Pearson's $r = .66$, $p < .001$) and to overlap conceptually as measures of experiential engagement with AI; they were therefore combined into a single variable, labeled "general AI use + comfort," to reduce redundancy and mitigate potential collinearity (Dormann et al., 2013; Field, 2018). A total of five independent variables were thus included in each regression model. The quantitative results are presented in Tables 3-5 (regression model summaries, ANOVA results, and regression coefficients, respectively).

19449728, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/flan.70082 by NICE, National Institute for Health and Care Excellence, Wiley Online Library on [28/10/2026]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 2 Dependent variables for linear regression models (beliefs and intention to integrate into language coaching).

Variable name	Question on questionnaire
"Integration interest"	How interested are you in integrating AI tools into your language instruction?
"AI improves quality"	Do you believe that AI tools can improve the quality of language instruction?
"Use AI in instruction"	How frequently do you currently use AI as a tool in your online language instruction?

TABLE 3 Regression model summaries.

	R	R^2	Adjusted R^2	SE of estimate
Integration interest	.72	.52	.49	0.66
AI improves quality	.68	.47	.44	0.57
Use of AI in instruction	.57	.33	.29	0.80

TABLE 4 ANOVAs run for regression models.

	df	*F*	*p*
Integration interest	5	19.27	<.001
AI improves quality	5	15.83	<.001
Use of AI in instruction	5	8.62	<.001

TABLE 5 Summary of regression model coefficients.

	Dependent variables					
	Integration interest		AI improves quality		Use of AI in instruction	
	B	*β*	*B*	*β*	B	*β*
Familiarity	−0.06 (0.09)	−.06	−0.12 (0.08)	−0.14	−0.04 (0.12)	−.04
Society influence	0.08 (0.09)	.07	0.17** (0.08)	.19	−0.1 (0.11)	−.09
Student perception	0.16 (0.11)	.13	0.02 (0.1)	.02	−0.19 (0.14)	−.14
Future importance of AI	0.18 (0.13)	.13	0.24** (0.11)	.21	0.09 (0.15)	.06
General AI use + comfort	0.27*** (0.05)	.56	0.2*** (0.04)	.5	0.32*** (0.07)	.66
Constant	0.32 (0.57)		1.38*** (0.49)		1.05 (0.68)	

Note: *$p < .10$, "Beta" = standardized coefficient; B coefficients are shown with standard errors in parentheses.; **$p < .05$; ***$p < .01$.

3.4.3 | Qualitative analyses

To address RQ2, the study incorporated a qualitative component based on participants' open-ended responses. These questions were designed to elicit in-depth perspectives on AI and were thematically organized in alignment with the overarching research objectives.

19449728, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/flan.70082 by NICE, National Institute for Health and Care Excellence, Wiley Online Library on [28/10/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Three broad domains were examined: (a) pedagogical usefulness of AI, participants were invited to reflect on the instructional value of AI; (b) ethical implications of using AI, questions in this category focused on potential concerns and risks; and (c) long-term integration of AI in language teaching, questions aimed to capture participants' views on the future role of AI.

To strengthen the reliability of the qualitative coding, an independent double-coding procedure was applied across all open-ended questionnaire items. Responses were systematically categorized into thematic units by two coders, and the outputs were compared in several stages. First, coding units were defined at the level of "mentions" of a given theme, allowing for a direct comparison of category frequencies. Second, observed agreement was calculated by summing the overlapping mentions across categories and dividing this value by the mean of the coders' total mentions. Third, chance agreement was estimated on the basis of the coders' marginal distributions. Finally, an approximate chance-corrected reliability coefficient (κ) was derived using the standard formula. Inter-rater reliability values for each open-ended item are available via IRIS [https://www.iris-database.org/details/75Yju-v1VGE]; most values exceeded the .80 threshold typically interpreted as "very good."

Following the independent coding, a reconciliation process was undertaken to resolve any remaining discrepancies between coders. The finalized categories for each open-ended questionnaire item are presented in Tables 6–8, together with descriptions of each category. Illustrative responses drawn directly from the data set (participant identifiers are denoted as "C__") are discussed in the narrative that follows. To ensure analytic relevance, only the three most representative variables per question are reported. Variables were retained if they contributed novel information and reached at least five mentions, a threshold equivalent to just over 5% of the 96 informants, thereby ensuring that included themes reflected patterns observed across multiple participants rather than isolated responses. In Tables 6–8, "n" represents the mean frequency between Rater 1 and Rater 2; averaging the two ratings minimizes coder bias and reflects a central tendency in category frequency.

4 | RESULTS

4.1 | Linear regression models to address RQ1

All regression models met assumptions for linearity, normality (Kolmogorov–Smirnov, all $p > .05$), independence (Durbin–Watson, all $p > .05$), multicollinearity (tolerance > 0.10, VIF < 10), and homoskedasticity, and were significant by ANOVA (Table 4). All R values were above .50, reaching .72 for the "integration interest" model. The independent variables explained a relatively high proportion of variation, particularly for "integration interest" (adjusted $R^2 = .49$) and "AI improves quality" (adjusted $R^2 = .44$) (Table 3).

Table 5 summarizes the regression coefficients. The variable "general AI use + comfort" had the largest and most consistent effect, significant at $p < .01$ across all three models. For the "AI improves quality" model, "society influence," "future importance of AI," and "general AI use + comfort" all showed a significant effect at $p < .05$.

Across all three regression models, the selection of independent variables displays a high, significant positive relationship with the dependent variables, in each case accounting for between 29% and 49% of the variation in the model (Table 3). Not all independent variables played an equal role, however. Coaches' combined comfort with and general use of AI tools

19449728, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/flan.70082 by NICE, National Institute for Health and Care Excellence, Wiley Online Library on [28/10/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

("general AI use + comfort") emerged as the most influential variable overall, and was the only variable to display a significant effect in all three regression models: greater general use and comfort was associated with stronger interest in integrating AI into instruction, stronger belief that AI can improve the quality of instruction, and higher current use of AI tools.

A notable pattern emerges in what was *not* significant. Coaches' interest in integrating AI tools was not significantly affected by their beliefs about AI's influence on society ("society influence"), their general familiarity with the concept of AI ("familiarity"), their perceptions of students' views ("student perception"), or their beliefs about the future importance of AI ("future importance of AI"). The last is especially noteworthy: believing that AI will matter for the field did not, in itself, sway coaches' intention to use AI tools. Conversely, even coaches who did not see AI tools as critical to the future of language instruction expressed interest in integrating them, provided they were generally comfortable using such tools ("general AI use + comfort"). The same pattern holds for coaches' current use of AI in their coaching, where again only "general AI use + comfort" was significant.

We can observe, then, a contrast between the variables influencing language coaches' *own* current and future interest in integrating AI tools in language coaching, on the one hand, and their beliefs regarding the potential of AI tools to improve language instruction writ large. For example, while coaches' positive perceptions of AI (including for language coaching, specifically) do not appear to lead them to be more interested in integrating these tools into their own coaching, such positive perceptions *are* correlated with coaches' belief that AI has the potential to improve the quality of language instruction: In the regression model for the dependent variable "AI improves quality," coaches' belief that AI holds a positive influence on society ("society influence") and their belief that AI tools are important to the future of language teaching ("future importance of AI") displayed a significant positive effect (in addition to the variable "general AI use + comfort").

4.2 | Thematic analysis to address RQ2

Coaches voiced a clear belief that AI tools can improve student learning and make coaching more efficient, while expressing concerns about the threat AI poses to coaches' future employment and to the quality of the learning experience. When asked about pedagogical usefulness, coaches second most frequently mentioned AI's ability to enable access to new, high-quality information for lesson plans and content delivery ("information and content"). Coaches do not choose conversation topics with students, so they are expected to be conversant in subject areas in which they have no prior training or personal experience. As one coach (C6) stated, "there are sessions where I'm asked to speak about certain topics that I have very little knowledge of, and by explaining this to the AI, it gives me a number of ideas for me to talk about during the session, so the AI can really help to make the content broader and less repetitive." Other "information and content" mentions referred to the delivery of content, and the capability of AI to help coaches "find new ideas and improve exercises" (C63), or to come across "some effective methods to make instruction more dynamic and fun" (C71). In the aggregate, coaches citing this type of potential for AI tools in instruction seemed to express a desire to push beyond the bounds of their own creativity when preparing for conversation sessions.

Seen from this perspective, coaches often viewed AI as a potential collaborative partner. As coach C83 stated, "[AI] gives us the opportunity to innovate with respect to new models of teaching/learning...it provides us with tools that can be interactive and not at all traditional when we go to teach."

19449728, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/flan.70082 by NICE, National Institute for Health and Care Excellence, Wiley Online Library on [28/10/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Coaches also highlighted the potential usefulness of AI in organizing lessons in a more efficient manner ("organization/lesson planning" in Table 6). In particular, coaches often mentioned the ability of AI to help with a lesson's structure and organization, for example, in order to "create a structure in accordance with the timing of each session" (C27) or to "optimize planning processes" (C74). This concern with organization and efficiency was frequently coupled with a desire to use AI in order to personalize lessons to a given student, language level, or university context. As stated by coach C2, "I feel that [AI] could help me to plan my sessions according to the university...so that students would feel more comfortable speaking about the topic." Coach C5 similarly wanted to use AI to gather "more information about the city where the students are located, in order to be able to empathize more with them." As indicated by the nature of coaches' comments, the ideal of personalized coaching is often at odds with the time constraints of a busy coaching schedule in which sessions must be held with students from any number of different universities. From the perspective of coaches, the tools provided by AI have the potential to save time in lesson planning, and in the time that is saved, those same tools provide the possibility of rapid lesson personalization as coaches jump from one student and university to another.

At the same time that coaches saw great potential in AI's ability to improve the organization and content of sessions, coaches also expressed a belief that AI is incapable of replacing a vital, human element to the language coaching experience. When asked if they thought that AI is comparable to human instructors in terms of providing cultural knowledge about language use, 70 of the 96 coach participants cited the inability of AI to replicate human connection, emotion, and experience (see Table 6, category "Need humans for emotional connection"). Similarly, when asked about the "potential risks of using AI as they relate to language teaching," coaches cited issues relating to "social skills and interaction" more than any other category (total 47 mentions; Table 7).

Coaches also expressed apprehension about data privacy ($n = 37.5$; Table 7) and the loss of critical thinking (cited by 32 coaches as a risk of AI in language teaching; Table 7), reflecting a broader societal unease that coaches brought into their professional reflections and a shared concern that students may become passive consumers of AI-generated content. Addressing the danger that AI poses on this front, coach C3 sums up the sentiment shared by many of the coach participants, and in so doing highlights an important distinction between the personalization of lessons facilitated by AI (a benefit to coaches), on one hand, and the personalization offered by a human coach (which cannot be replaced by AI), on the other:

> I would say that the main thing is the lack of human interaction and, as a consequence, the lack of experience. The personalization that AI offers is related to the implementation of a course adapted to the needs of a student. The personalization that a human instructor offers lies in [that instructor's] experiences, in the cultural context where the target language is spoken. This instructor has experienced the celebrations, they have tried the food, they've experienced an infinite number of moments that are valuable when they are shared with a student. Nothing will be able to change the light in the eyes of a student when [the coach] relates something that they loved and enjoyed, and the pair of attentive ears that are absorbing this information and hoping to visit that city, that country, or to try that food.

19449720, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/flan.70082 by NICE, National Institute for Health and Care Excellence, Wiley Online Library on [28/10/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 6 Thematic categories on the pedagogical usefulness of AI.

Q4. What are the primary benefits of AI?	
"Efficiency" ($n = 59$)	Mentions of the efficiency of AI, and its ability to help people to save time on various tasks. Included in this category are mentions of AI's ability to help with the general organization and management of information and tasks.
"Information & content" ($n = 41$)	Mentions of the ability of AI to enable access to new/better information, and/or improving the quality of information for lessons/work. This includes mentions of AI's ability to provide users with new or more creative ideas.
"Error reduction" ($n = 7.5$)	AI's ability to reduce errors (all contexts).
Q20. What are the potential benefits of AI for language teaching?	
"Information & content" ($n = 40$)	Any reference to AI's ability to provide ideas for lessons, new activities for language practice, etc.
"Grammar/feedback/correction" ($n = 30.5$)	AI's ability to help with grammar information/correction, including feedback for students (or, in some cases, for coaches themselves).
"Oral/aural" ($n = 21$)	Mentions of using AI for additional listening/speaking practice. This includes explicit mention of using chatbots for practice.
Q26. How do you think AI can support you in planning your coaching sessions?	
"Organization/lesson planning" ($n = 33$)	Mentions of AI's ability to help coaches to organize and plan their lessons (e.g., how to organize basic ideas into a coherent lesson, what order to present topics, etc.).
"Personalization" ($n = 32$)	Any reference to AI's ability to customize lessons for different students depending on a student's interests, learning style, proficiency, etc.
"Information & content" ($n = 66.5$)	Separate from the organization of lessons/materials, this category includes mentions of AI's ability to give coaches new ideas and/or materials to be incorporated into a lesson plan.
Q30. How do you think AI compares with human instructors in terms of providing cultural knowledge about language use?	
"Need humans for emotional connection" ($n = 70$)	The precise nature of coaches' responses here vary, but this category includes any reference to AI's inability to replicate human connection, emotion, and experience, and/or to the need to have human coaches that can share lived experiences grounded in the real world.
"Information & content" ($n = 42$)	Mentions of AI's ability to provide a greater quantity of information and/or more specific information on a given subject, compared to human instructors.
"Need human coaches to customize information/feedback" ($n = 9$)	Some coaches express the belief that human coaches are needed in order to interpret the information given by AI and adapt it to a given student or lesson.

19449720, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/flan.70082 by NICE, National Institute for Health and Care Excellence, Wiley Online Library on [28/10/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 7 Thematic categories on the ethical implications of AI.

Q5. What worries or risks do you associate with AI?	
"Data theft/privacy" ($n = 37.5$)	Any mention of the threat of to personal data, privacy concerns, and/or fraud.
"Academic misconduct/quality loss" ($n = 9.5$)	Any mention of the threat of academic misconduct, cheating, and/or the overall lowering of academic quality as a result of AI.
"Unreliable information" ($n = 23$)	The unreliability of information produced by AI.
Q21. What are the potential risks of using AI as they relate to language teaching?	
"Social skills & interaction" ($n = 47$)	Any mention of the loss of human interaction, including mentions of the value of interacting with other humans, learning from someone else's lived experience, etc. Also included in this category are mentions of the loss of the skills associated with human interaction (e.g., conversational skills, the development of empathy, etc.).
"Linguistic richness" ($n = 16$)	Reference to AI's inability to match the linguistic richness provided by human instructors/coaches.
"Critical thinking" ($n = 32$)	Mentions of the loss of critical thinking skills and/or "intellectual effort" (this is distinguished from mentions of the loss of interactional skills, part of the previous category, as "critical thinking" refers to intellectual—not social—development).

Coaches also highlighted the need for a human coach presence to identify, understand, and mediate students' emotions, in a way AI chatbots could not match. As coach C9 stated, "AI is not a human...[only] we have the capacity to perceive emotions and react to them." While AI might produce a simulacrum of emotionally supportive speech, coaches believed that human nature contributes most directly to the emotional connection forged in conversation sessions.

Many coaches expressed apprehension about their ability to maneuver a professional environment increasingly shaped by AI. The characteristic most frequently wanted in AI tools was that they be easy to use and access (category "easy to use"; Table 8). The most common anticipated challenge was "transition/training," encompassing both the need for professional training and the difficulty of adopting new AI tools without adequate prior training. Responses frequently mentioned worries over "not knowing how to use it" (C51), "not having much experience with using AI" (C59), or "receiving the necessary training" (C6). These echo the quantitative findings for RQ1, where comfort with and general use of AI tools was the only consistent predictor of coaches' interest in integration, belief in AI's value, and current use. Coaches also expressed concerns about overreliance on AI tools (23 mentions for Q37; Table 8) and increased reliance on AI (41 mentions for Q43; Table 8), as well as fears about job loss due to AI (18 mentions). Across the three thematic domains, several cross-cutting patterns emerge. Coaches most valued AI's capacity to provide content, enhance planning efficiency, and facilitate personalization (Table 6), while expressing prominent concerns about the loss of human interaction, critical thinking, and data privacy (Table 7). Regarding long-term integration (Table 8), coaches anticipated increased reliance on AI but voiced anxieties about job displacement, overreliance, and the need for training. These patterns are explored in the Discussion that follows.

19449728, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/flan.70082 by NICE, National Institute for Health and Care Excellence, Wiley Online Library on [28/10/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 8 Thematic categories on the long-term integration of AI in language teaching.

Q6. Do you think that AI can outperform humans on certain tasks? Which ones?	
"Data processing/information summary" ($n = 32.5$)	AI's ability to process large quantities of information and summarize it nearly instantaneously.
"Repetitive/automatic tasks" ($n = 18$)	Any reference to repetitive tasks, or tasks that may be automatized with AI.
"Writing" ($n = 15$)	Any references to AI's ability to write texts.
Q12. How do you think AI will shape the future of your field or profession?	
"Audiovisual content" ($n = 6$)	Creation of audiovisual content
"Access" ($n = 8.5$)	The ability of individuals to access information (on language learning/instruction) from anywhere, at any time.
"Will not replace coach" ($n = 7$)	Any explicit expression from coaches of the belief that AI will not replace human coaches.
Q36. What characteristics would you like AI tools to have in order to be of greater use in your coaching?	
"Easy to use" ($n = 25.5$)	Any reference to the importance of AI being easy to use and access.
"In-session support" ($n = 14$)	The need to have support while using AI tools, especially during language lessons/coaching sessions.
"Speed" ($n = 12.5$)	Similar to "easy to use," but these references are specific to the speed of being able to access and use AI tools.
Q37. What challenges do you think you would face upon integrating AI tools into your coaching?	
"Transition/training" ($n = 42.5$)	Any reference to the need for training in order to use AI tools, and/or references to the potential difficulty of making a transition to increase the use of AI tools.
"Overreliance" ($n = 23$)	Any reference to the potential for coaches to rely too much on AI tools, at a cost to the quality of instruction and/or the coaching experience.
"Loss of human connection" ($n = 18.5$)	Any reference to the general loss of human connection in the language learning process. This category also includes mentions of the loss of "naturalness" in conversations.
Q43. How do you imagine the role of AI in education, specifically in language teaching, over the next 10 years?	
"Job loss" ($n = 18$)	Any mention of the threat of AI to replace workers, or reduce their role.
"Increased reliance" ($n = 41$)	Mentions of the increasingly important role that AI will have in language teaching, but without necessarily citing a complete takeover or total loss of jobs, etc.
"Chatbots" ($n = 11.5$)	Mentions of the use of AI chatbots in language instruction/learning.

5 | DISCUSSION

The findings of this study suggest that online Spanish language coaches occupy a dual stance toward AI in language instruction. Quantitatively, coaches' interest in integrating AI tools, their belief that AI can improve instructional quality, and their current use of AI in coaching were

19449720, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/flan.70082 by NICE, National Institute for Health and Care Excellence, Wiley Online Library on [28/10/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

predicted most strongly (and, in all three models, exclusively) by their general comfort with and use of AI tools, rather than by abstract beliefs about AI's societal influence or its future importance to the field. Qualitatively, coaches consistently valued AI's capacity for content generation, lesson personalization, and planning efficiency, while voicing equally consistent concerns about the loss of human connection, the erosion of critical thinking, data privacy risks, and the prospect of overreliance. Overall, these results suggest that coaches are cautiously pragmatic adopters: open to AI as a complement to their work, but firm about preserving the interpersonal and intercultural dimensions they see as the core of coaching.

The finding that "general AI use + comfort" was the only consistent significant predictor across all three regression models aligns with Long et al. (2024), who likewise report that familiarity strengthens perceptions of the usefulness of AI tools. It also extends Cabero-Almenara et al. (2024) and Viberg et al. (2025), whose studies on educators' acceptance of AI emphasize prior exposure as a precondition for trust, into the under-researched context of online language coaching. Coaches appeared to grow more optimistic about AI's benefits as they grew more familiar with it; conversely, a degree of "fear of the unknown" may operate for coaches who, unfamiliar with AI tools and uncomfortable with their own ability to manage them, view adoption less favorably. This pattern highlights the importance of professional readiness and the availability of training to language coaches, a point that complements Wu and Yu (2023) and McGrath et al. (2025), who show that the effectiveness of AI chatbots in improving L2 outcomes depends heavily on instructors' ability to integrate such tools into lesson design, and resonates with Moorhouse and Wong's (2025) call for a multidimensional "professional-GenAI-competence" encompassing technological, pedagogical, and ethical readiness.

Equally informative is what the regression models did *not* show. Beliefs about AI's societal influence, general familiarity, perceptions of students' views, and beliefs about AI's future importance were not significant predictors of coaches' own intention to integrate AI tools. The non-significance of "future importance of AI" is particularly striking: believing that AI will matter for the field does not, in itself, translate into wanting to use it. This pattern reveals a gap between abstract endorsement and situated practice, similar to the distinction Belda-Medina and Kokošková (2024), Ma et al. (2024), and Butler and Jiang (2025) draw between language teachers' general acceptance of AI and their concrete classroom uptake. It also suggests that professional development premised on convincing coaches that AI matters is unlikely to shift practice; what shifts practice is hands-on, situated comfort.

Turning to the qualitative findings, coaches most often invoked AI's capacity to support content generation, lesson organization, and personalization (Table 6). This aligns with Crompton et al. (2024) and Kim et al. (2025), who describe AI's affordances for adaptive scaffolding and personalized feedback, and with Mills et al. (2025) on AI companions designed to support foreign language writing. The frequent mention of personalization in particular is apt, given that "personalization" is one of the core methodological principles of language coaching proposed by Chaika (2023); at its best, language coaching engenders a dynamic, less fixed learning environment where personalized content is central (Zablotska, Olendr and Tsar, 2023) and generic, mechanical drills are avoided (Ceo-Francesco, 2015). Coaches' depiction of AI as a "collaborative partner" that helps them push beyond their own creative bounds also extends prior work on educators' productive uses of generative AI (Hınız, 2024; Li et al., 2025) into the language-coaching context, which is characterized, as coaches themselves emphasized, by professional isolation across distant geographies.

Counterbalancing these affordances, coaches voiced prominent concerns about data privacy, unreliable information, loss of social interaction and empathy, and the erosion of

19449728, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/flan.70082 by NICE, National Institute for Health and Care Excellence, Wiley Online Library on [28/10/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

critical thinking (Table 7). The salience of "need humans for emotional connection" and "social skills & interaction" echoes Voss and Waring (2025) argument that conversational AI does not replicate naturalistic human interaction, and aligns with Ryan, (2025) critique of human-centered AI, which warns against framings that elide what is distinctively human in pedagogical exchange. The emotional skills coaches invoked, lowering anxiety, scaffolding communication, adapting to learners' needs, are core to coaching practice (Ceo-Francesco, 2015), and findings by Guerrero-Rodriguez (2023) on heritage speakers of Spanish suggest that sustained human interaction can help learners move from linguistic insecurity to confidence, an affective trajectory not easily replicated by AI; indeed, Liu and Reinders (2025) report that some students found chatbot interactions artificial and distracting. Concerns about overreliance and the loss of critical thinking link directly to the calls for critical AI literacy articulated by Ng et al. (2021), Sperling et al. (2024), Lu et al. (2025), and Tacelosky et al. (2025), and to Kessler et al. (2025) on the ethical dimensions of generative AI in non-English L2 contexts.

On the long-term horizon (Table 8), coaches anticipated increased reliance on AI while voicing anxieties about job displacement, overreliance, and the need for training. This mirrors the broader trajectory described by Moorhouse and Kohnke (2024) and Moorhouse and Wong (2025), in which language educators move from initial novelty toward critical reflection and integration, but with persistent unease about institutional readiness. The recurring "will not replace coach" theme, in particular, suggests that coaches see themselves as lasting mediators rather than roles AI could simply replace, a self-positioning consistent with Manierre et al. (2025) on hybrid models in which human coaches and AI agents co-exist.

Together, the quantitative and qualitative findings suggest that coaches do not simply adopt or reject AI, but manage its use selectively. They *co-exist* with AI by using it selectively for what it does well (content, organization, personalization) while keeping for themselves what they see as the human core of coaching: emotional connection, cultural authenticity, and critical guidance. In doing so, they *co-construct* an emerging professional practice in which AI's role is continually weighed against pedagogical and ethical concerns. This dual stance of co-existing and co-constructing suggests that the central question for the field is not whether coaches will integrate AI, but under what professional and institutional conditions that integration can preserve the human core of language education.

6 | LIMITATIONS AND FUTURE DIRECTIONS

Several limitations should be acknowledged. The cross-sectional design captures coaches' perceptions at a single point during the early stages of AI diffusion, and examines beliefs and readiness rather than session-based implementation. Future research should pursue longitudinal designs to trace coaches' "AI trajectories" as training initiatives and institutional policies are implemented, and experimental approaches to evaluate AI-supported tools' pedagogical impact on learner engagement and performance. Future work should also investigate co-construction in a more literal sense, analyzing how human coaches and AI agents collaboratively generate meaning, feedback, and motivation in real-time sessions, and conceptualize the ethics of "AI-human hybridity" (Ryan, 2025) in educational interactions.

19449728, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/flan.70082 by NICE, National Institute for Health and Care Excellence, Wiley Online Library on [28/10/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

7 | PEDAGOGICAL IMPLICATIONS

The findings carry several actionable implications for language educators, coaching platforms, and teacher education programs. First, coaching platforms should develop structured AI literacy training. Since comfort and familiarity with AI tools were the strongest predictors of adoption, training should prioritize hands-on experience in coaching-relevant scenarios such as lesson personalization, content generation, and scaffolding conversational topics, and should address not only technological proficiency but also pedagogical integration (Muñoz-Basols & Fuertes Gutiérrez, 2025; Muñoz-Basols et al., 2024; Muñoz-Basols, Fuertes Gutiérrez & Cerezo, 2025), ethical reasoning, intercultural awareness, and critical evaluation of AI-generated content (Lu et al., 2025; Ng et al., 2021; Sperling et al., 2024).

Second, language programs should adopt hybrid instructional models that position AI as a collaborative support tool rather than a substitute for human expertise. Coaches could use AI to prepare session content and generate tailored discussion prompts, while reserving synchronous time for the interpersonal exchange, cultural sharing, and emotional scaffolding where coaches emphasized the irreplaceability of human presence.

Third, platforms and institutions should establish transparent ethical guidelines on acceptable AI use in coaching, addressing coaches' expressed concerns about data privacy, misinformation, and over-reliance, as well as preserving the affective dimension of coaching interactions. These guidelines should be co-constructed with coaches themselves, drawing on their situated knowledge of what works in practice and what risks they encounter. Language coaching organizations must also articulate clear policies on data management and the co-construction of learning experiences between human and machine interlocutors.

Fourth, teacher education and coach training programs should incorporate reflective activities that help coaches articulate their own stance toward AI, following models of critical AI literacy in language education (Tacelosky et al., 2025). Activities might include collaborative workshops where coaches share strategies for using AI, discuss ethical dilemmas, and develop personalized AI integration plans for their coaching contexts.

Finally, given that many coaches work in isolation across different countries, platforms should create communities of practice where coaches can exchange AI-related ideas, concerns, and innovations, thereby transforming individual experimentation into collective professional growth. Through these implications, the study advocates for a model of reflective integration that enriches, rather than replaces, the human core of language education.

8 | CONCLUSIONS

This study has provided a first empirical examination of how online Spanish language coaches perceive and prepare for the integration of AI in their professional practice. The findings reveal a dual dynamic, *co-existing* and *co-constructing* with AI, in which coaches recognize the instrumental value of AI for lesson planning, personalization, and content generation, while consistently guarding the interpersonal, cultural, and emotional dimensions they consider central to effective coaching. Familiarity and comfort with AI tools emerged as the strongest predictors of adoption, pointing to the urgency of targeted professional development.

19449720, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/flan.70082 by NICE, National Institute for Health and Care Excellence, Wiley Online Library on [28/10/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

ACKNOWLEDGMENTS

The authors would like to express their gratitude to the Teacher Training Department at LinguaMeeting, especially Daniel Román and the platform Founder Elena Casillas, for their valuable support and collaboration throughout the development of this study.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

Research data are not shared.

ORCID

Javier Muñoz-Basols https://orcid.org/0000-0003-3856-3637
Tripp Strawbridge https://orcid.org/0000-0003-4992-0302

REFERENCES

Abing, J. L. (2018). Language coaching for speakers of Spanish as a heritage language: Novice-expert interactions via videoconference *(Doctoral dissertation, University of Texas at Austin*. ProQuest.

Belda-Medina, J., & Kokošková, V. (2024). ChatGPT for language learning: Assessing teacher candidates' skills and perceptions using the Technology Acceptance Model (TAM). *Innovation in Language Learning and Teaching*, *20*(3), 446–461. https://doi.org/10.1080/17501229.2024.2435900

Berninger-Schäfer, E. (2022). *Online-coaching*. Springer.

Butler, Y. G., & Jiang, S. (2025). How do pre-service language teachers perceive generative AIs' affordance?: A case of ChatGPT. *System*, *129*, Article 103606. https://doi.org/10.1016/j.system.2025.103606

Cabero-Almenara, J., Palacios-Rodríguez, A., Loaiza-Aguirre, M. I., & Rivas-Manzano, M. R. (2024). Acceptance of educational artificial intelligence by teachers and its relationship with some variables and pedagogical beliefs. *Education Sciences*, *14*(7), 740. https://doi.org/10.3390/educsci14070740

Ceo-Francesco, D. (2015). Engaging learners in culturally authentic virtual interactions. In A. J. Moeller (Ed.), *Learn languages, explore cultures, transform lives* (pP. 39–58). Central States Conference on the Teaching of Foreign Languages.

Chaika, O. (2023). Psycholinguistic factors in second language acquisition: Foreign language teaching via coaching. In X. Jiang (Ed.), Psycholinguistics—New advances and real-world application (pp. 1–22). https://doi.org/10.5772/intechopen.1003720

Chapelle, C. A. (2025). Generative AI as game changer: Implications for language education. *System*, *132*, 103672. https://doi.org/10.1016/j.system.2025.103672

Clementi, L., & Jiménez López, M. D. (2025). Análisis de textos generados por ChatGPT para la enseñanza del español: complejidad y adecuación a niveles de competencia. *Journal of Spanish Language Teaching*, *13*(1), 36–63. https://doi.org/10.1080/23247797.2025.2534203

Cowden, J. D., Martinez, F. J., Dickmeyer, J. J., & Bratcher, D. (2022). Culture and language coaching for bilingual residents: the first 10 years of the CHiCoS model. *Teaching and learning in medicine*, *35*(5), 589–600. https://doi.org/10.1080/10401334.2022.2092113

Crompton, H., Edmett, A., Ichaporia, N., & Burke, D. (2024). AI and English language teaching: Affordances and challenges. *British Journal of Educational Technology*, *55*(6), 2503–2529. https://doi.org/10.1111/bjet.13460

Derakhshan, A., & Ghiasvand, F. (2024). Is ChatGPT an evil or an angel for second language education and research? A phenomenographic study of research-active EFL teachers' perceptions. *International Journal of Applied Linguistics*, *34*(4), 1246–1264. https://doi.org/10.1111/ijal.12561

Dormann, C. F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., Marquéz, J. R. G., Gruber, B., Lafourcade, B., Leitão, P. J., Münkemüller, T., McClean, C., Osborne, P. E., Reineking, B., Schröder, B., Skidmore, A. K., Zurell, D., & Lautenbach, S. (2013). Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, *36*(1), 27–46. https://doi.org/10.1111/j.1600-0587.2012.07348.x

19449720, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/flan.70082 by NICE, National Institute for Health and Care Excellence, Wiley Online Library on [28/10/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Field, A. (2018). *Discovering statistics using IBM SPSS statistics* (5th ed.). SAGE Publications.

Guerrero-Rodriguez, P. (2023). Telecollaboration: from linguistic insecurity to linguistic confidence in Spanish as a heritage language. *Computer Assisted Language Learning*, *38*(5–6), 1141–1166. https://doi.org/10.1080/09588221.2023.2269405

Hajar, A. (2021). Theoretical foundations of phenomenography: A critical review. *Higher Education Research & Development*, *40*(7), 1421–1436. https://doi.org/10.1080/07294360.2020.1833844

Hınız, G. (2024). A year of generative AI in English language teaching and learning - A case study. *Journal of Research on Technology in Education*, *58*(2), 291–311. https://doi.org/10.1080/15391523.2024.2404132

Jones, L. (2025). WhatsApp and Zoom for communication activities: speaking proficiency and learner perceptions. *Journal of Spanish Language Teaching*, *12*(1), 33–57. https://doi.org/10.1080/23247797.2025.2457210

Kessler, M., Rostrán Valle, J. M., Çekmegeli, K., & Farrell, S. (2025). Generative AI for learning languages other than English: L2 writers' current uses and perceptions of ethics. *Foreign Language Annals*, *58*(3), 508–531. https://doi.org/10.1111/flan.70005

Kim, J., Yu, S., Lee, S. S., & Detrick, R. (2025). Students' prompt patterns and its effects in AI-assisted academic writing: Focusing on students' level of AI literacy. *Journal of Research on Technology in Education*, *58*(3), 638–655. https://doi.org/10.1080/15391523.2025.2456043

Li, M., Belpoliti, F., Taha, G., & Zhang, M. (2025). World language teachers' reported practices of integrating ChatGPT into instruction: A positioning analysis. *Foreign Language Annals*, *58*(4), 995–1017. https://doi.org/10.1111/flan.70041

Liu, M., & Reinders, H. (2025). Do AI chatbots impact motivation? Insights from a preliminary longitudinal study. *System*, *128*, 103544. https://doi.org/10.1016/j.system.2024.103544

Long, T., Gero, K. I., & Chilton, L. B. (2024). Not just novelty: A longitudinal study on utility and customization of an AI workflow. In *Proceedings of the Designing Interactive Systems Conference* (pp. 780–792). ACM. https://doi.org/10.1145/3643834.3661587

Lu, X., Zaitseva, A., & Troyan, F. J. (2025). AI literacy: A core practice in world language education. *Foreign Language Annals*, *59*(1), 9–33. https://doi.org/10.1111/flan.70037

Ma, Q., Crosthwaite, P., Sun, D., & Zou, D. (2024). Exploring ChatGPT literacy in language education: A global perspective and comprehensive approach. *Computers and Education: Artificial Intelligence*, *7*, 100278. https://doi.org/10.1016/j.caeai.2024.100278

Manierre, M., Propst, L., Cohen, A., & Rogers, J. (2025). Coexisting with ChatGPT: Evaluating a tool for AI-based paper revision. *Computers and Composition*, *76*, 102923. https://doi.org/10.1016/j.compcom.2025.102923

Marull, C., & Kumar, S. (2020). Authentic language learning through telecollaboration in online courses. *TechTrends*, *64*(4), 628–635. https://doi.org/10.1007/s11528-020-00488-2

McGrath, C., Farazouli, A., & Cerratto-Pargman, T. (2025). Generative AI chatbots in higher education: A review of an emerging research area. *Higher Education*, *89*(6), 1533–1549. https://doi.org/10.1007/s10734-024-01288-w

Mills, N., Hok, H., Dressen, A., & Veillas, Q. (2025). The design and evaluation of an interactive AI companion for foreign language writing. *Foreign Language Annals*, *58*(1), 40–69. https://doi.org/10.1111/flan.12790

Moorhouse, B. L., & Kohnke, L. (2024). The effects of generative AI on initial language teacher education: The perceptions of teacher educators. *System*, *122*, Article 103290. https://doi.org/10.1016/j.system.2024.103290

Moorhouse, B. L., & Wong, K. M. (2025). Generative artificial intelligence and language teaching. In *Cambridge elements in language teaching*. Cambridge University Press.

Muñoz-Basols, J., Fuertes Gutiérrez, M., & Cerezo, L. (Eds.). (2024). La enseñanza del español mediada por tecnología: de la justicia social a la Inteligencia Artificial (IA*)*. Routledge. https://doi.org/10.4324/9781003146391

Muñoz-Basols, J., Fuertes Gutiérrez, M., & Cerezo, L. Eds. (2025). *Technology-mediated language teaching: From social justice to artificial intelligence*. Multilingual Matters. https://doi.org/10.21832/9781800419889

Muñoz-Basols, J., & Fuertes Gutiérrez, M. (2025). Opportunities for artificial intelligence (AI) in language teaching and learning. In J. Muñoz-Basols, M. Fuertes Gutiérrez, & L. Cerezo (Eds.), *Technology-mediated language teaching: From social justice to artificial intelligence* (pp. 297–315). Multilingual Matters. https://doi.org/10.21832/9781800419889-017

19449720, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/flan.70082 by NICE, National Institute for Health and Care Excellence, Wiley Online Library on [28/10/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Muñoz-Basols, J., & Lvova, Y. (2025). El efecto novedad en la integración de la IA generativa: perspectivas de docentes y aprendientes en la preparación de exámenes de certificación DELE. *Revista Internacional de Lenguas Extranjeras/International Journal of Foreign Languages*, *23*, 121–160. https://doi.org/10.17345/rile23.4247

NCSSFL & ACTFL. (2017). *NCSSFL-ACTFL Can-Do statements*. ACTFL. https://www.actfl.org/resources/ncssfl-actfl-can-do-statements

Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S. (2021). Conceptualizing AI literacy: An exploratory review. *Computers and Education: Artificial Intelligence*, *2*, Article 100041. https://doi.org/10.1016/j.caeai.2021.100041

Pérez-Núñez, A. (2024). ChatGPT in Spanish language instruction: exploring AI-driven task generation and its implications for teaching practices. *Journal of Spanish Language Teaching*, *11*(1), 61–82. https://doi.org/10.1080/23247797.2024.2366053

Piette, J. D., Marinec, N., Chen, J., Yon, S., Maly, M., & Pfeiffer, P. N. (2024). Pilot study of an English language coaching opportunity for volunteering among US military veterans with depression, anxiety, and/or post-traumatic stress disorder. *Journal of Affective Disorders Reports*, *16*, 100761. https://doi.org/10.1016/j.jadr.2024.100761

Ryan, M. (2025). We're only human after all: A critique of human-centred AI. *AI & Society*, *40*(3), 1303–1319. https://doi.org/10.1007/s00146-024-01976-2

Slusarek, M. (2024). *World language teachers' beliefs and practices in asynchronous courses at two-year colleges* (Doctoral dissertation, University of Iowa). https://doi.org/10.25820/etd.007536

Sperling, K., Stenberg, C. J., McGrath, C., Åkerfeldt, A., Heintz, F., & Stenliden, L. (2024). In search of artificial intelligence (AI) literacy in teacher education: A scoping review. *Computers and Education Open*, *6*, 100169. https://doi.org/10.1016/j.caeo.2024.100169

Stolz, S. A. (2020). Phenomenology and phenomenography in educational research: A critique. *Educational Philosophy and Theory*, *52*(10), 1077–1096. https://doi.org/10.1080/00131857.2020.1724088

Strawbridge, T. (2021). Modern language: Interaction in conversational NS-NNS video SCMC eTandem exchanges. *Language Learning & Technology*, *25*(2), 94–110. http://hdl.handle.net/10125/73435

Strawbridge, T. (2025). Conversations abroad: The effect of interlocutor type, relationship quality, and L2 proficiency on interaction in naturalistic settings. *The Modern Language Journal*, *109*(2), 407–427. https://doi.org/10.1111/modl.12997

Tacelosky, K., Sue Kasun, G., Shapiro, B. R., Liao, Y.-C., & Harris, K. (2025). Exploring critical AI literacy in language education: A case study. *Foreign Language Annals*, *58*(4), 966–994. https://doi.org/10.1111/flan.70029

Thomson, B. (2013). *Non-directive coaching: Attitudes, approaches and applications*. Critical Publishing.

Tribes, M., & Masid Blanco, O. (2025). Perspectivas docentes del uso de ChatGPT en la enseñanza del español: aplicaciones, desafíos e implicaciones. *Journal of Spanish Language Teaching*, *12*(1), 17–32. https://doi.org/10.1080/23247797.2025.2504266

Tulasi, L., & Rao, C. S. (2023). Integration of AI-technologies into ELT: A brief study. *Journal for Research Scholars and Professionals of English Language Teaching*, *7*(38), 1–7. https://doi.org/10.54850/jrspelt.7.38.003

Viberg, O., Cukurova, M., Feldman-Maggor, Y., Alexandron, G., Shirai, S., Kanemune, S., Wasson, B., Tømte, C., Spikol, D., Milrad, M., Coelho, R., & Kizilcec, R. F. (2025). What explains teachers' trust in AI in education across six countries? *International Journal of Artificial Intelligence in Education*, *35*(3), 1288–1316. https://doi.org/10.1007/s40593-024-00433-x

Voss, E. (2024). Language assessment and artificial intelligence. In A. Kunnan (Ed.), *The concise companion to language assessment* (pp. 112–125). John Wiley & Sons.

Voss, E., & Waring, H. Z. (2025). When ChatGPT can't chat: The quest for naturalness. *TESOL Quarterly*, *59*(2), 1064–1075. https://doi.org/10.1002/tesq.3374

Wielgolawski, C. (2011). On the road to self-directed learning: A language coaching case study. In B. Morrison (Ed.), *Independent language learning: Building on experience, seeking new perspectives* (pp. 119–135). Hong Kong University Press.

Wu, R., & Yu, Z. (2023). Do AI chatbots improve students' learning outcomes? Evidence from a meta-analysis. *British Journal of Educational Technology*, *55*(1), 10–33. https://doi.org/10.1111/bjet.13334

19449720, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/flan.70082 by NICE, National Institute for Health and Care Excellence, Wiley Online Library on [28/10/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Zablotska, L., Olendr, T., & Tsar, I. (2023). Influence of language coaching techniques on undergraduate students' motivation in the process of foreign languages teaching at non-linguistic specialties. *Arab World English Journal*, *14*(1), 394–408. https://doi.org/10.24093/awej/vol14no1.25

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Muñoz-Basols, J., & Strawbridge, T. (2026). Co-existing and co-constructing with AI: Beliefs, practices, and professional readiness of online language coaches. *Foreign Language Annals*, 1–22. https://doi.org/10.1111/flan.70082

19449720, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1111/flan.70082 by NICE, National Institute for Health and Care Excellence, Wiley Online Library on [28/10/2025]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License