Time	AIRiAL 2025 Presentations Day 1: Friday, September 26, 2025
8:00 - 9:00	Registration
9:00 - 9:15	Welcome and Opening Remarks ***Erik Voss**, Teachers College, Columbia University*
9:20 - 10:40	**Session 1**
Paper 1: 9:20- 9:40	**Emotional AI in Language Education: A Systematic Review** *Erik Voss, Teachers College, Columbia University* *Dan Eskin, Teachers College, Columbia University* *Kedi Mo, Teachers College, Columbia University* *Shamini Shetye, Teachers College, Columbia University* *Xiaoya Wang, Teachers College, Columbia University* *Yilin Zhang, Teachers College, Columbia University*
Paper 2: 9:40 - 10:00	**“Hey there! I’m here to help you”: Investigating the Use of an Interactive AI Character to Support Affective Engagement in a Scenario-Based Language Assessment** *Soo Hyoung Joo, Teachers College, Columbia University*
Paper 3: 10:00 - 10:20	**Anxiety-Responsive AI for L2 Speaking Practice: Reducing Affective Barriers in Oral Production** *Yusuf Cengiz, Boğaziçi University, Türkiye* *Rabia Büşra Çemç, Boğaziçi University, Türkiye*
Paper 4: 10:20 - 10:40	**AI and Social/Emotional Perceptions of L2 Speech: Inconsistencies, Contradictions, and (Un)reasonable Evaluations with GPT Audio Prompting** *Kevin Hirschi, University of Texas, San Antonio* *Okim Kang, Northern Arizona University* *Yuna Bae, Northern Arizona University*

10:50 - 11:50	**Plenary Speaker**
	Julia Hirschberg, *Columbia University* **<u>Title:</u>** **Generating Empathetic Dialogue for English Language Learning** (Followed by Friday Group Picture)
12:00 - 1:00	Lunch Break
1:00 - 2:00	**Poster & Technology Demonstration Session** (See presenters below)
2:00 - 3:20	**Session 2**
Paper 1: 2:00 - 2:20	**Toward Emotionally Intelligent AI Feedback:** **Insights from a Longitudinal Study of Learner-AI Interaction** *Inyoung Na, Iowa State University* *Mahdi Duris, Iowa State University* *Volker Hegelheimer, Iowa State University*
Paper 2: 2:20 - 2:40	**Enhancing L2 Spanish Sentence-Level Production and Communicative Practice with an AI Chatbot: Students' Perspective on ChatGPT-Enhanced Learning Activities** *Pablo Robles-García, University of Toronto, Mississauga, Canada* *Ji-young Shin, University of Toronto, Mississauga, Canada* *Claudia Sánchez-Gutiérrez, University of California, Davis*
Paper 3: 2:40 - 3:00	**The Secret Life of International Graduate Students' Use of ChatGPT:** **Identity and (Dis)empowerment Impacts (Cancelled)** *Seyyedeh Mobina Hosseini, University of Buffalo*
Paper 4: 3:00 - 3:20	**Bridging the Empathy Gap: A Comparative Analysis of Emotional Intelligence in Human and AI-Generated Educational Feedback** Johanathan Woodworth, Mount Saint Vincent University, Canada Danny Tan, Seneca Polytechnic, Canada
3:20 - 3:50	Coffee Break

4:00 - 5:00	**Session 3**
Paper 1: 4:00 - 4:20	**Can AI Augment Human Roles in ELT Instruction? Exploring EFL Educators' Perspectives on Language as a Social and Cultural Practice** *Jamola Urunbaeva, Boston College* *Kamola Muradkasimova, Uzbek State World Languages University, Uzbekistan*
Paper 2: 4:20 - 4:40	**Generative AI-Integrated Task-Based Instruction to Enhance L2 English for Newcomers and Refugees: A Mixed-Methods Approach** *Ji-young Shin, University of Toronto, Mississauga, Canada* *Liz Coulson, University of Toronto, Mississauga, Canada*
Paper 3: 4:40 - 5:00	**Social AI in Online Classrooms: A Corpus-Based Study of AI-Generated Teaching Styles** *Tony Berber Sardinha, Pontifical Catholic University of São Paolo, Brazil* *Marilisa Shimazumi, Cultura Inglesa College, Brazil*
5:00 - 5:15	Announcements
5:30 - 7:00	Opening Reception (Wine & Cheese in Everett Lounge)

AIRiAL 2025 Poster and Technology Demonstration

Friday, September 26, 2025
Time: 1:00 - 2:00 pm

POSTERS

Is AI a Competent Guide? Learner and Teacher Perceptions on AI's Role in Developing Intercultural Communicative Competence
Zeynep Saka, Syracuse University

Exploring the Integration of Affective Computing into the Italian Scenario-Based Assessment
Giulia Peri, University for Foreigners of Siena, Italy
Sabrina Machetti, University for Foreigners of Siena, Italy

AI Reformulation of Policy Documents: Linguistic and Discursive Implications for Educational Administration
Fernanda Peixoto Coelho, Pontifícia Universidade Católica de São Paulo, Brazil
Tatiana Schmitz de Almeida Lopes, Pontifícia Universidade Católica de São Paulo, Brazil
André Luis Andrade Lansac, Pontifícia Universidade Católica de São Paulo, Brazil

Guided by AI: Language Acquisition and AI Literacy Through Scaffolded Integration
Krisztina Domjan, American University

Generative AI Feedback to Foster Growth Mindsets in Highly Sensitive Students
Aubrey Michelle Sahouria, Center for Applied Linguistics

The Social and Intercultural Mediation: Building Bridges in Global Communication
Fernanda Peixoto Coelho, Pontifical Catholic University of São Paolo, Brazil
Tatiana Schmitz de Almeida Lopes, Pontifical Catholic University of São Paolo, Brazil
Adilson Gomes, Pontifical Catholic University of São Paolo, Brazil
André Luis Andrade Lansac, Pontifical Catholic University of São Paolo, Brazil
Maurício Ayres Cunha, Metropolitan University of Santos, Brazil

BoodleBox-Assisted PBL for Scaffolding Collaborative Inquiry
Haiyan Li, Purdue University

Synthetic Authenticity: Performativity and Enregisterement in AI-Generated Language from Bot Conversations
Daniel Murcia, The Pennsylvania State University

Co-Creating Clarity: AI as a Collaborative Partner in Navigating Emotional Complexities of Educational Migration
Adekunmi Olatunji, University of Hawai'i at Mānoa

TECHNOLOGY DEMONSTRATIONS

Context-Aware Intelligent Learning System with Multi-Agent AI Architecture
Venkat Podugu, Q3 Learners

Pangea Chat: AI-Enhanced Peer-to-Peer Conversation
Will Jordan-Cooley, Pangea Chat

Time	AIRiAL 2025 Presentations Day 2: Saturday, September 27, 2025
8:00 - 9:00	Registration
9:00 - 9:10	Opening Remarks
9:10 - 10:30	**Session 4**
Paper 1: 9:10 - 9:30 *Paper 2:* 9:30 - 9:50 *Paper 3:* 9:50 - 10:10	**Evaluating a Generative AI-Based System for Assessing Interactional Competence: System Performance and User Perceptions** *Inyoung Na, Iowa State University* **Tutor, Text Expert, or Diagnostician? Investigating Learner Positioning of an AI-Based Chatbot in Dynamic Assessment of L2 Reading** *Ari Huhta, University of Jyväskylä, Finland* *Matthew Poehner, The Pennsylvania State University* *Dmitri Leontjev, University of Jyväskylä, Finland* *Luke Harding, Lancaster University, United Kingdom* *Vera Vesala, University of Jyväskylä, Finland* **Maintaining Ethical and Secure Data Practices in Machine Scoring of the AAPPL Test** *Scott Gravina, Language Testing International*
10:10-10:40	Break
10:40-11:40	**Plenary Speaker**
	***Àgata Lapedriza**, Northeastern University* **Title:** **Emotion-Aware AI: Methods, Challenges, and Opportunities** (Followed by Saturday Group Picture)

11:40-12:40	Lunch Break
12:40 - 2:00	**Session 5**
Paper 1: 12:40 - 1:00	**Understanding Emotion Learning Dynamics and Attention Mechanisms in Multimodal Transformers** *Bingjie Wang, University of Rochester* *Chenxi Shi, Teachers College, Columbia University*
Paper 2: 1:00 - 1:20	**High EQ or not?** **AI and human evaluation of social media subtext in Chinese pragmatic contexts** *Chuhao Wu, Clemson University* *Zimeng Shao, The Pennsylvania State University*
Paper 3: 1:20 - 1:40	**Testing the Reliability of Generative AI Systems in Emotion Analysis: A Human–Machine Comparison** *Şebnem Kurt, Purdue University* *Mahdi Duris, Iowa State University* *Kimberly Becker, Midland University*
Paper 4: 1:40 - 2:00	**Emotional Safety as Interaction: Designing Relational Learning with Generative AI** *Chenxi Shi, Teachers College, Columbia University* *Jing Hao, Teachers College, Columbia University*
2:00 - 2:30	Break
2:30 - 3:50	**Session 6**
Paper 1: 2:30 - 2:50	**Identifying Authorship in Student Writing: Insights from Teacher Judgments and Computational Analysis** *Şebnem Kurt, Purdue University* *Shuhui Yin, Iowa State University* *Danilo Calle Londolo, Iowa State University* *Nergis Daniş, Iowa State University* *Hwee Jean Lim, Iowa State University* *Inyoung Na, Iowa State University* *Carol Chapelle, Iowa State University*
Paper 2: 2:50 - 3:10	**AI-DA of L2 Writing: Designing Cognitively and Emotionally Responsive Mediation to Diagnose Learner Writing Development** *Xiaozheng Dai, The Pennsylvania University* *Xiaofei Lu, The Pennsylvania State University* *Matthew Poehner, The Pennsylvania State University* *Jingyuan Zhang, The Pennsylvania State University* *Lu Yu, University of Melbourne, Australia*
Paper 3: 3:10 - 3:30	

Paper 4: 3:30 - 3:50	**Can AI Write like an Undergraduate Student?** **A Multidimensional Analysis Looking at Disciplinary Variation** *Larissa Goulart, Montclair State University* *Wesley Acorinti, Northern Arizona University* *Yejin Jung, University of Utah* *Marine Matte, Instituto Federal de Educação, Ciência e Tecnologia Sul-rio-grandense, Brazil* **Between Irony and Indictment:** **Exploring the Discursive Dimensions of Social AI Interpretations of Politicized COVID-19 Imagery** *Yara Maria de Toledo Dias Romeiro, Pontifical Catholic University of São Paolo, Brazil* *Tony Berber Sardinha, Pontifical Catholic University of São Paolo, Brazil*
3:50 - 4:15	Closing Remarks & Best Student Paper Award Announcement
6:30 - 9:30	Networking Dinner *The Ellington on Broadway - The Strauss Room* *2745 Broadway, New York, NY 10025*

Conference Abstracts

Paper Abstracts: Friday

Session 1
Friday, September 26, 2025 9:20 - 10:40 am

Social & Emotional AI Scoping Review
Erik Voss, Teachers College, Columbia University
Dan Eskin, Teachers College, Columbia University
Kedi Mo, Teachers College, Columbia University
Shamini Shetye, Teachers College, Columbia University
Xiaoya Wang, Teachers College, Columbia University
Yilin Zhang, Teachers College, Columbia University

As artificial intelligence (AI) technologies continue to evolve, there is growing debate about the extent to which computer systems should mimic human emotions or be able to change a human's emotional state. As a growing field of research, emotional AI refers to systems that can identify, classify, and respond to affective states such as happiness, frustration, or anxiety (Ho et al, 2023) with the goal of creating systems that can monitor, interpret, and react to users' emotional cues (Camuñas, et al., 2025; Wu, et al, 2025). However, an overview of emotional AI research in language education is not well documented. This review aims to describe the characteristics of existing publications, emotional theories/models and emotional states and outcomes or impacts (if any) are reported in empirical studies of emotional AI in language education. This systematic review was conducted following the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines (Page et al., 2021). It was conducted to help contextualize this rapidly growing body of research for attendees at the AIRiAL 2025 conference as a means of describing the landscape of this year's conference theme, the future of emotionally intelligent machines.

"Hey there! I'm here to help you": Investigating the Use of an Interactive AI Character to Support Affective Engagement in a Scenario-Based Language Assessment
Soo Hyoung Joo, Teachers College, Columbia University

Recognizing the need to measure and promote 21st-century competencies, such as critical thinking and collaborative problem solving (OECD, 2015), recent research has emphasized the importance of assessing situated second/foreign language (S/FL) proficiency: the ability to marshal a range of resources, including emotional engagement, confidence, and effort, to participate in the socio-cognitive practices of disciplinary communities (Purpura, 2016, 2017, 2021; Purpura & Oh, 2024). As a response to this need, Scenario-Based Language Assessments (SBLAs) offer a promising approach by simulating goal-driven, socially contextualized tasks that elicit authentic language use (O'Reilly & Sabatini, 2013; Purpura, 2021). Simulated characters play a pivotal role in SBLAs, guiding the narrative and providing the emotional support needed for goal-oriented engagement. However, these characters have traditionally been represented as static images with voice-overs that limit emotional responsiveness (Purpura et al., 2025). With advances in multimodal generative AI, it is now possible to create interactive characters that respond in real-time.

This study investigates how an interactive AI character embedded in a STEM-focused SBLA can support the affective dimension, defined in the Learning-Oriented Language Assessment (LOLA) framework as learners' emotional perceptions (e.g., interest, confidence) and behavioral

dispositions (e.g., engagement, persistence) that moderate performance (Purpura & Turner, 2018; Purpura, 2024). Created using HeyGen, the character scaffolds CEFR B1-level learners through a Coulomb's Law simulation using prompts designed to model curiosity and offer emotional support. Transcript analysis identified emotional support strategies, revealing how interactive AI can enhance learners' engagement, confidence, and willingness to persist in cognitively demanding tasks.

Anxiety-Responsive AI for L2 Speaking Practice: Reducing Affective Barriers in Oral Production

Yusuf Cengiz, Boğaziçi University, Türkiye
Rabia Büşra Çemç, Boğaziçi University, Türkiye

Among the second language (L2) skills, speaking is usually regarded as the one that invokes anxiety the most for L2 learners (Luo, 2014). Speaking anxiety continues to hinder spoken language development among L2 learners, often resulting in reduced language proficiency and performance (Fathi et al., 2024; Teimouri et al., 2019; Zhang, 2019). Effective strategies to deal with speaking anxiety include providing learning opportunities (Chen, 2024) and feedback opportunities (Shadiev et al., 2023) for L2 learners.

Generative artificial intelligence (AI) tools have offered a paradigm shift in education (O'Dea, 2024). Owing to their generative capabilities, these tools can offer real-time analysis of the spoken language and deliver immediate feedback (Zou et al., 2023). Moreover, they can provide opportunities to practice speaking autonomously (Fathi et al., 2024). Such timely feedback and practice opportunities can decrease the levels of anxiety and enhance speaking accuracy (Rad et al., 2024; Shadiev et al., 2023).

To this end, this study investigates the potential of ChatGPT in acting as a speaking practice tool to decrease speaking anxiety for B1-level L2 learners at a higher education institution in Turkey. ChatGPT has been prompted by the researchers to respond to participants' (N=15) anxiety by slowing its responses, providing positive reinforcement, and simplifying follow-up questions. Speaking anxiety levels of the participants are measured through a pre-posttest design, and semi-structured interviews are conducted to gain in-depth insight into the participants' experiences. The study discusses the pedagogical implications of employing ChatGPT both as a speaking practice tool and as an anxiety-responsive intervention.

AI and Social/Emotional Perceptions of L2 Speech: Inconsistencies, Contradictions, and (Un)reasonable Evaluations with GPT Audio Prompting

Kevin Hirschi, University of Texas, San Antonio
Okim Kang, Northern Arizona University
Yuna Bae, Northern Arizona University

AI is both lauded for increased accuracy and capabilities (Radford et al., 2023), and criticized for reproducing human bias (Bommasani et al., 2021). These biases are potentially compounded when applying AI to L2 accented speech (Bae & Kang, 2024; Hirschi & Kang, 2024). To explore the alignment between AI and social bias, 24 audio recordings of L2 speakers with various proficiency levels and accentedness were evaluated by 161 listeners on 10 social and emotional dimensions. The same recordings were subjected to OpenAI's GPT Preview-4o-Audio model, which is trained in both speech recording and textual data, using an enhanced rating prompt with reasoning. The results indicate almost no agreement between humans and AI in ratings of graceful/awkward, intimate/distant, humorous/serious, modern/old-fashioned, emotional/unemotional, and casual/formal ($r<.38$). However, there were negative correlations between human and AI ratings of relatable/unrelatable ($r =.-.42$, $p=.04$), effective/ineffective ($r =-.61$, $p<.01$), and pleasant/unpleasant ($r =-.65$, $p<.01$). AI only replicated human perception of plain/expressive to a moderate degree ($r =.53$, $p<.01$). The GPT audio model was also asked to

reason while evaluating, which revealed that the model relied on textual features for determining a rating value with few notable exceptions, including explicit mentions of 'intonation' when describing rating reasoning for plain/expressive, and 'tone' or 'voice' with pleasant/unpleasant. The findings suggest L2 speech may inhibit accurate sentiment analysis by AI and highlight the potential for models to disaggregate textual and acoustic dimensions in ways that diverge from human perception. Implications and a roadmap for future research will be discussed.

Session 2
Friday, September 26, 2025 2:00 - 3:20 pm

Toward Emotionally Intelligent AI Feedback: Insights from a Longitudinal Study of Learner-AI Interaction

Inyoung Na, Iowa State University
Mahdi Duris, Iowa State University
Volker Hegelheimer, Iowa State University

Although GenAI, as an Automated Writing Evaluation (AWE) feedback tool, has been rapidly adopted for writing support, most studies focus on immediate benefits, offering little insight into how learners engage with GenAI over time. Yet sustained interaction, including shifts in trust, confidence, and feedback use, is critical for understanding how such tools can meaningfully support writing development (Ranalli, 2021). Closer examination of engagement patterns is essential for understanding their pedagogical potential (Stevenson & Phakiti, 2019).

This longitudinal study investigated learner interaction with ChatGPT-4o during an eight-week period, including lower-level ESL learners (n = 5), higher-level ESL learners (n = 4), and native English-speaking undergraduates (n = 4). Data sources included screen recordings, ChatGPT-learner conversation logs, reflective learning logs, and semi-structured interviews.

Findings revealed that prompting ChatGPT with a social persona (e.g., "act as my professor") consistently led to learners' perceptions of more targeted, useful, and human-like feedback. Learners initially treated ChatGPT as a mechanical tool but increasingly viewed it as a collaborative social agent (e.g., teacher, peer). Concurrently, learners' identities shifted from passive users to active collaborators and critical learners, regulating their trust, confidence, and strategies for integrating AI feedback. Concerns about impersonal feedback also emerged, revealing tension between accepting help and preserving authorship.

These insights highlight the need to design GenAI-based AWE systems that attend not only to linguistic support but also to social framing, learner agency, and the evolving emotional dynamics of human-AI interaction. Implications for learner training and instructional design are discussed.

Enhancing L2 Spanish Sentence-Level Production and Communicative Practice with an AI Chatbot: Students' Perspective on ChatGPT-Enhanced Learning Activities

Pablo Robles-García, University of Toronto, Mississauga, Canada
Ji-young Shin, University of Toronto, Mississauga, Canada
Claudia Sánchez-Gutiérrez, University of California, Davis

Recent studies have reported positive roles of generative AI-powered feedback in enhancing L2 English writing (Guo & Wang, 2023; Han & Li, 2024; Steiss et al., 2024), while little is known regarding L2 Spanish. This study reports on the creation and implementation of two types of ChatGPT-enhanced activities used in an intermediate Spanish language course: (1) chat-based role-plays and (2) sentence writing tasks with corrective feedback. In addition to documenting

the design and application of these activities, this presentation aims to explore students' overall impressions of using ChatGPT for language learning, as well as their specific experiences during the two course-based activities mentioned above. Results from 14 student semi-structured interviews revealed that most students found ChatGPT helpful, valuing its convenience, instant feedback, and adaptability for improving conversational skills, vocabulary, and grammar, considering it an excellent supplement and a valuable resource for fostering student agency and self-paced learning. In role-play interactions, students found ChatGPT effective at directing and facilitating quality practice, providing a consistent, non-judgmental environment with accurate feedback and personalized scenarios. However, they noted its lack of emotional depth and spontaneity compared to peer interactions, which foster richer conversations and deeper connections. Finally, the corrective feedback provided was clear, organized, and tailored to their needs. Students appreciated its promptness, which facilitated quick corrections, and noted that structured responses enhanced understanding and encouraged self-reflection. Nevertheless, they raised concerns about occasional failures to identify errors, sometimes flagging correct information as incorrect or providing inaccurate feedback, leading to confusion and frustration.

The Secret Life of International Graduate Students' Use of ChatGPT: Identity and (Dis)empowerment Impacts (Cancelled)

Seyyedeh Mobina Hosseini, University of Buffalo

Technologies like AI introduce new possible constellations of identity, agency, and social practices. Williams (2023) stresses the need for AI literacy at all educational levels so that students become "technosocial change agents" or people who use powerful technologies for personal and societal empowerment and addressing complex problems. Moreover, the introduction of ChatGPT into higher education has seemingly eased academic burdens, particularly for international students, making their workload more manageable. However, this tool acts as a double-edged sword: while it enhances learning and provides valuable academic support, it also carries certain drawbacks. Although the prevalent view is that AI-powered technologies positively impact students' learning by assisting them, their influence on students' identity and self-perception remains underexplored. To contribute to fill this gap, this study seeks to understand how ChatGPT plays a role in the academic experiences of international graduate students and explores its impact on their identity and self-perception. Therefore, interviews were conducted with 15 international PhD students across the United States. The findings identified five key themes regarding how ChatGPT facilitates and supports students' academic endeavors and four overarching themes concerning its impact on students' identities and potential disempowerment effects. Emerging from this analysis, an AI-Powered Fugitive Learning Model is proposed, reflecting the way students' ChatGPT practices lead to ambivalence and how their use of AI tools is veiled in secrecy. The findings highlight the need to reframe academic policies and pedagogical practices to reduce stigma around AI use.

Bridging the Empathy Gap: A Comparative Analysis of Emotional Intelligence in Human and AI-Generated Educational Feedback

Johanathan Woodworth, Mount Saint Vincent University, Canada
Danny Tan, Seneca Polytechnic, Canada

As GenAI systems increasingly supplement or replace traditional teacher feedback in educational settings, important questions arise regarding their ability to deliver emotionally intelligent responses that support student growth and well-being. This study investigates the "empathy gap" between human teacher feedback and GenAI-generated feedback on student assignments. Employing a mixed-methods approach, we collected and analyzed paired feedback samples (human teacher and GenAI) on identical student work across various disciplines. Feedback was evaluated using a multi-dimensional framework encompassing technical accuracy, constructive guidance, motivational language, empathetic understanding,

and personalization. Preliminary findings reveal notable differences in the integration of emotional intelligence within feedback from human teachers versus GenAI systems. While GenAI excelled in technical accuracy and consistency, human teachers demonstrated a deeper contextual understanding of students' emotional needs and learning trajectories. Moreover, student responses indicated that perceived empathy in feedback significantly influenced their engagement and willingness to act on the feedback provided. Our findings indicate that, despite recent advances in natural language processing, current GenAI systems still struggle to replicate the nuanced emotional intelligence characteristic of effective teacher feedback. Specifically, GenAI feedback often appears formulaic, excessively positive, and lacks the directness and authenticity that students value in human responses. We discuss implications for developing more emotionally intelligent AI feedback systems and propose a hybrid model that leverages the complementary strengths of both human and AI feedback. This research contributes to the evolving discourse on Social AI by highlighting the specific dimensions of emotional intelligence that remain challenging for AI systems to authentically manifest in educational contexts.

Session 3
Friday, September 26, 2025 4:00 - 5:00 pm

Can AI Augment Human Roles in ELT Instruction? Exploring EFL Educators' Perspectives on Language as a Social and Cultural Practice

Jamola Arunbaeva, Boston College
Kamola Muradkasimova, Uzbek State World Languages University, Uzbekistan

Artificial Intelligence (AI) has been a heated debate in language education for the past few years. Although many teachers have been using AI in their lesson planning and teaching English as a Foreign Language (EFL) in Uzbekistan, their practices, opinions, and concerns about ways they imagine AI as a support for routine tasks and personalized learning have not been known. Ethical questions that have arisen in the past few years about human roles in English Language Teaching (ELT), considering that the language is a social and cultural phenomenon, which needs instruction across various cultural contexts. This study examines attitudes and critical AI literacy of 48 EFL teachers, 6 curriculum developers, and 3 department leaders at the Uzbek State World Languages University through an online survey and semi-structured interviews. The research findings mirror that EFL teachers' perspectives towards AI tools vary, from cautious to optimistic. Their concerns show that AI frequently misses human interaction and cultural and sociocultural spheres of language. Curriculum developers see the potential of AI in augmenting but not replacing the teachers' roles in ELT classrooms, and AI both supports and hinders the teaching of language as a social and cultural practice. The university management envisions educator-AI partnerships in AI integration to promote personalized, social, and cultural nuances of language education, which have been overlooked. The study recommends listening to EFL educators' voices in humanizing AI-driven instruction and doing cross-cultural comparisons and longitudinal studies in teaching EFL.

Generative AI-Integrated Task-Based Instruction to Enhance L2 English for Newcomers and Refugees: A Mixed-Methods Approach

Ji-young Shin, University of Toronto, Mississauga, Canada
Liz Coulson, University of Toronto, Mississauga, Canada

While GenAI's affordances in L2 instruction have been widely studied, less is known about its integration into specific instruction models or its use with non-traditional learners. (Jeon et al, 2024). This mixed-methods, quasi-experimental study explores ChatGPT's impact on teaching English to refugees of varying proficiency levels and their perceptions.

Data were collected from 38 newcomers and refugees enrolled in a seven-week English program in Canada, which taught expressions and language functions for everyday conversations within TBLT. Participants' proficiency ranged from high-beginner to low-advanced, with Arabic being the most common L1. With random assignment, the experimental group participated in ChatGPT-integrated pre-, during-, and post-tasks, while the control group completed tasks with assistance from volunteering tutors. Participants took elicited imitation (EI) tests before and after the program.

A multilevel random intercept model was used to assess pre-post differences between groups, where forty EI item scores (level 1) were clustered within each individual test taker/participant (level 2). The best-fitting model indicated that pre-post test score changes across individuals and groups were significant, even when controlling for pre-existing L2 English proficiency. However, the group differences were not significant, which means that AI chatbot-incorporated activities were as effective as interactions with human tutors. Follow-up interviews with participants also demonstrated the positive impact of AI chatbot-incorporated activities for their engagement, motivation, and attitude, while pronouncing increased accessibility. The study details AI-enhanced TBLT and supports its potential for L2 English learners in non-school settings such as refugees, making implications for diversity, equity, and inclusion.

Social AI in Online Classrooms: A Corpus-Based Study of AI-Generated Teaching Styles

Tony Berber Sardinha, Pontifical Catholic University of São Paolo, Brazil
Marilisa Shimazumi, Cultura Inglesa College, Brazil

This paper looks at the linguistic and discursive performance of an LLM (GPT o3) in enacting the social and affective functions of teaching in online higher-education classes. Whereas most social AI research relies on decontextualized scenarios with no human interlocutors, this study relies on authentic human interactions drawn from real-world classrooms. We collected 50 human-taught online classes from two use-of-English tracks and five disciplinary subjects and divided them into segments to comply with LLM context-window limits. We selected segments based on teacher turn density and determined 56 pedagogic, affective, demographic, and sociocultural teaching styles based on the previous literature (Csomay & Crawford, 2024; Grasha, 1996)and our personal experience. We prompted the LLM with a summary of each teacher-turn for each segment, instructing it to simulate contextually appropriate teacher responses based on each style, class goals and subject matter. This resulted in a 28.6-million-word corpus comprising 20,300 human-taught and AI-generated college sessions in English. We used Lexical Multi-Dimensional Analysis (Berber Sardinha & Fitzsimmons-Doolan, 2025) to detect the underlying discourses, resulting in five dimensions: 1) Macro-structuring versus Performative stagecraft; 2) L1-scaffolded guidance versus provocative challenge; 3) Procedural control versus Rapport-driven engagement; 4) Authoritarian control versus Reflective aesthetic engagement; and 5) Informal facilitation versus Formal deductive exposition. In the paper presentation, we will detail the similarities and differences between the human-led and AI-led classes for the teaching styles and classes along each dimension. The study provides an empirically grounded framework for evaluating the adequacy of machine-mediated social interaction through corpus-based discourse evidence.

Paper Abstracts: Saturday

Session 4
Saturday, September 27, 2025 9:10 - 10:30 am

Evaluating a Generative AI-Based System for Assessing Interactional Competence: System Performance and User Perceptions
Inyoung Na, Iowa State University

A central challenge in large-scale speaking assessment is eliciting interactional competence (IC) while minimizing interlocutor variability. Although Spoken Dialogue Systems, used as speech partners instead of humans, have improved score dependability in paired tasks (Ockey & Chukharev-Hudilainen, 2021), their rule-based designs limit discourse authenticity and responsiveness. This study introduces a GenAI-based system for a university English Placement Test to assess IC in paired discussions.

To inform system development, three large language models (LLMs)—Claude 3.5, GPT-4o, and Gemini 2.0 Flash—were compared for their capacity to sustain stance-based dialogue, generate communication breakdowns, and elicit ratable IC features. Claude 3.5 was selected as the AI speech partner, and GPT-4o as the examiner. To support targeted IC elicitation, the system used scripted tools to introduce miscommunication at predetermined turns. L2-accented and emotionally inflected speech was incorporated for the AI partner to simulate realistic interaction conditions. Performances from assessment experts (n = 2) and test takers (n = 4) were analyzed through discourse transcripts coded for 12 IC features. Semi-structured interviews explored perceptions of the system's conversational quality and assessment potential.

Results showed successful elicitation of clarification responses, topic development, and stance recognition. However, participants' comprehension checks and self-correction were rarely observed, likely due to the LLMs' inferencing capabilities. Participants perceived the system as fair and effective, though limitations in topic elaboration were noted.

This study offers insights into GenAI's role in high-stakes speaking assessment and demonstrates methods for selecting and implementing GenAI tools in applied linguistics. Implications for AI-mediated assessment design will be discussed.

Tutor, Text Expert, or Diagnostician? Investigating Learner Positioning of an AI-Based Chatbot in Dynamic Assessment of L2 Reading
Ari Huhta, University of Jyväskylä, Finland
Matthew Poehner, The Pennsylvania State University
Dmitri Leontjev, University of Jyväskylä, Finland
Luke Harding, Lancaster University, United Kingdom
Vera Vesala, University of Jyväskylä, Finland

Research on dynamic assessment (DA) has recently started to explore the use of AI-based chatbots either as mediators (e.g., Jeon, 2023) or as additional resources within a dynamic assessment procedure (e.g., Dai et al., 2025; Yu & Poehner, 2025). In this presentation, we focus on the second use – chatbots as additional resources. Specifically, we explore how learners conceptualise the role of an AI-based chatbot used within a dynamic assessment

process through discourse analysis of learner-chatbot interactions. This study forms part of a four-year ongoing project (DD-LANG), which focuses on exploring synergies between the dynamic assessment and diagnostic assessment frameworks. We analyse data drawn from ca. 200 Finnish upper-secondary school learners who engaged in a DA of L2 reading ability. Within the assessment procedure, an AI-based chatbot was integrated as a resource complementing computerised mediation (support provided when learner performance on tasks breaks down; see Poehner & Lantolf, 2023). Data from learner-chatbot interactions are analysed using discourse analytic approaches (including positioning theory and sociocultural discourse analysis) to identify the positions learners ascribe to themselves and to the chatbot, and how these change during the dynamic assessment. We argue that investigating learner-chatbot interactions can provide insights into the specific role that learners understand the chatbot to be playing, potentially impacting on engagement with the dynamic assessment procedure more generally. Finally, we discuss the benefits of this novel approach to using AI as a resource in mediating learner performance and how AI might best be incorporated into future computerised DA approaches.

Maintaining Ethical and Secure Data Practices in Machine Scoring of the AAPPL Test

Scott Gravina, Language Testing International

As machine scoring systems become increasingly integrated into language assessment, ensuring data privacy and ethical accountability is more urgent than ever (Xi, 2023), especially when tests serve young learners. This presentation focuses on the ACTFL Assessment of Performance toward Proficiency in Languages (AAPPL) and explores how ACTFL and Language Testing International (LTI) apply rigorous data security protocols across both human- and machine-scored environments. Anchored in the ILTA Code of Ethics (ILTA, forthcoming), particularly the Principle of Technological Responsibility, our approach underscores transparency, data stewardship, and responsible innovation.

LTI's longstanding commitment to data protection includes COPPA compliance, use of secure data servers, and maintaining strict controls over access and processing. As we expand machine scoring capabilities (already operational for the AAPPL Presentational Writing (PW) mode and in research for the Interpretive Listening and Reading (ILS) mode) these same principles guide our development. Our presentation outlines the ethical and technical protocols embedded in our machine scoring pipeline: from using diverse, representative datasets, and transparent system training to regular monitoring and oversight. We also describe how we build all machine scoring models in-house to ensure test-taker data never leaves our secure environment.

In doing so, we demonstrate how high-stakes, large-scale assessments like the AAPPL can implement a machine scoring system while upholding ethical imperatives, maintaining trust, and protecting test taker data from breaches and misuse. Our work illustrates a model for ethical technological integration that aligns with both legal requirements and the evolving language assessment field.

Session 5
Saturday, September 27, 2025 12:40 - 2:00 pm

Understanding Emotion Learning Dynamics and Attention Mechanisms in Multimodal Transformers

Bingjie Wang, University of Rochester
Chenxi Shi, Teachers College, Columbia University

Understanding how emotion-aware multimodal models learn to represent and attend to human affect is critical for building AI systems that can genuinely perceive and respond to user feelings. In this work, we investigate the emotion learning dynamics and internal attention mechanisms of a state-of-the-art multimodal Transformer trained on synchronized text, audio, and visual emotion corpora. First, we introduce a suite of layer-wise probing tasks that measure how and where discrete affective categories (e.g., joy, sadness, anger) emerge in the model's hidden representations over the course of training. Second, we develop an Attention Attribution Framework that quantifies each head's contribution to emotion detection, using concentration and weight-distribution metrics to identify "emotion heads" that specialize in processing vocal prosody, facial expressions, or sentiment-laden words. Through experiments on four public multimodal emotion datasets, we demonstrate that (1) affective information is first encoded in early layers via acoustic and visual cues, (2) mid-level layers integrate modalities into unified emotion embeddings, and (3) a small subset of high-concentration attention heads in late layers drives final classification. Our analysis not only sheds light on the inner workings of multimodal emotion models but also suggests targeted pruning and interpretability strategies to enhance empathy and explainability in social AI.

High EQ or not? AI and human evaluation of social media subtext in Chinese pragmatic contexts

Chuhao Wu, Clemson University
Zimeng Shao, The Pennsylvania State University

Emotional intelligence (EQ) is widely recognized as a key factor in successful interpersonal communication, including in digital environments (Benke et al., 2022; Coronado-Maldonado & Benítez-Márquez, 2023). However, on Chinese social media, the notion of "high EQ" has taken on new meanings, often associated with sarcasm, irony, or performative politeness. For example, on one of China's most popular social networking platforms, Xiaohongshu (also known as RedNote), users frequently use the hashtag #高情商 (#HighEQ) to describe awkward or socially challenging situations and to solicit tactful responses. Yet, rather than offering pragmatic solutions, the comment sections often feature a mix of sarcastic, humorous, or even passive-aggressive replies.

This phenomenon provides a unique context for evaluating how well large language models (LLMs) align with human interpretations of emotional nuance and pragmatic appropriateness. While previous studies show that LLMs perform well on semantic tasks, their sensitivity to pragmatic implicature, emotional tone, and cultural specificity remains an open question (Ma et al., 2025; Sravanthi et al., 2023). This study aims to compare how LLMs and human annotators evaluate user comments from Xiaohongshu across three dimensions: (1) appropriateness for the "High EQ" label, (2) degree of aggressiveness, and (3) likelihood of being upvoted. Using a curated corpus of #HighEQ posts, both human and model-based evaluations are collected and analyzed to identify patterns of agreement or divergence. The results contribute to our understanding of human-AI (mis)alignment in interpreted pragmatic strategies, humor, sarcasm, and community idiosyncrasy on socially and culturally situated social media contexts.

Testing the Reliability of Generative AI Systems in Emotion Analysis: A Human–Machine Comparison

Şebnem Kurt, Iowa State University
Mahdi Duris, Iowa State University
Kimberly Becker, Midland University

As the integration of generative AI (GenAI) systems in linguistic analysis increases, it is critical to continue testing these systems' ability to provide reliable and consistent results compatible with the gold standard, which is human judgments. This study is a follow-up to Kurt & Duris (2025), which compared manual annotations and GenAI evaluations of emotional expression in U.S. tweets on abortion. The analysis, based on a dataset of 100 tweets collected in Boston, Massachusetts, between September 2022 and January 2023, applied Martin and White's (2005) Appraisal Framework, focusing on the AFFECT subcategory to classify emotions as positive, negative, or neutral. The current study incorporates DeepSeek's R1 Chain-of-Thought (CoT) processing to analyze the Twitter dataset using Martin & White (2005) to compare a) manual human vs. automated AI (ChatGPT 4o and DeepSeek R1) analysis of emotions, and b) compare the consistency of the two chatbots with the human annotations by calculating precision, recall, and F1 scores for both AI models to assess which system demonstrates greater reliability in detecting human emotion.

While previous research has primarily compared human emotion judgments with GPT's classifications, our study moves beyond output comparison by examining DeepSeek's intermediate reasoning steps, which can surface latent discourse features (e.g., sarcasm, pragmatic inference) that standard LLMs typically obscure. Preliminary findings suggest notable variances in categorization among human coders, GPT, and DeepSeek, with indications for CoT having a more comparable approach to human judgments.

Emotional Safety as Interaction: Designing Relational Learning with Generative AI

Chenxi Shi, Teachers College, Columbia University
Jing Hao, Teachers College, Columbia University

As generative AI becomes increasingly present in education, its influence extends beyond cognitive assistance into the affective conditions of learning. This paper focuses on emotional safety—the experience of being able to think, express, and revise without fear of judgment—as an interactional construct shaped through human–AI dialogue.

Drawing from sociocultural learning theory, affective computing, and relational pedagogy, we propose a conceptual framework for understanding emotional safety as co-produced through how generative AI systems recognize hesitation, interpret uncertainty, and respond with structured, non-evaluative feedback. Rather than focusing on affect recognition alone, we analyze how prompt design, turn-taking patterns, and linguistic tone mediate learners' willingness to take cognitive risks or tolerate ambiguity.

To illustrate this framework, we conduct exploratory case analysis of AI-mediated learning interactions—such as writing assistants and tutoring bots—examining how their conversational features may support or inhibit emotional safety in both learner-facing and teacher-support contexts.

By reframing emotional responsiveness as a relational process, this paper offers a lens for examining how AI systems shape the emotional architecture of educational engagement. Our approach combines theoretical analysis with case-based illustration to examine emotional safety not as a background condition, but as a dynamic property constructed—and at times constrained—through generative AI interaction.

Session 6
Saturday, September 27, 2025 2:30 - 3:50 pm

Identifying Authorship in Student Writing: Insights from Teacher Judgments and Computational Analysis

Şebnem Kurt, Iowa State University
Shuhui Yin, Iowa State University
Danilo Calle Londolo, Iowa State University
Danis Nergis, Iowa State University
Hwee Jean Lim, Iowa State University
Inyoung Na, Iowa State University
Carol Chapelle, Iowa State University

The increasing use of AI tools like ChatGPT in educational settings presents new challenges for identifying the authenticity of student writing. Previous research shows that teachers struggle to reliably distinguish between AI- and human-generated texts (e.g., Waltzer et al., 2024), often relying on surface-level features such as errors. This study explored how experienced university instructors judged the authorship of undergraduate summary essays and identified linguistic features that distinguish human- versus AI-generated writing. Using a mixed-methods design, we collected think-aloud data from three experienced teachers as they evaluated 60 anonymized essays (30 human-written and 30 AI-generated) and conducted a focus group interview to explore their decision-making processes. Additionally, a key feature corpus analysis (Egbert & Biber, 2023) compared linguistic features of the two sets of texts. Findings revealed that teachers could often correctly identify authorship, and their success depended on recognizing personal voice, rhetorical choices, and alignment with assignment guidelines. Teachers noted that human-written texts tended to use personal pronouns, rhetorical questions, and accessible vocabulary, while AI-generated texts relied on nominalizations, complex noun phrases, and formal structures. The key feature analysis supported these observations, identifying linguistic patterns that distinguished the two types of writing. These findings suggest that there are identifiable linguistic characteristics of AI-generated writing that teachers can recognize. By revealing analytical strategies and insights into AI writing features, this study may contribute to more effective practices for promoting academic integrity and accurately assessing student writing skills.

AI-DA of L2 Writing: Designing Cognitively and Emotionally Responsive Mediation to Diagnose Learner Writing Development

Xiaozheng Dai, The Pennsylvania University
Xiaofei Lu, The Pennsylvania State University
Matthew Poehner, The Pennsylvania State University
Jingyuan Zhang, The Pennsylvania State University
Lu Yu, University of Melbourne, Australia

In Dynamic Assessment (DA) diagnostic insights into fully-formed and emerging learner abilities are obtained through interaction during which different forms of support, or mediation, are offered to learners when difficulties arise. Efforts to expand the scale of DA face the challenge of replacing a human assessor/mediator with a computerized procedure while retaining maximum flexibility to mirror dialogic interaction. We report data from a pilot study of an ongoing project that develops an online DA system that integrates Generative AI to assess and mediate L2 learners' argumentative writing. After submitting an essay, learners engage in a mediation session where ChatGPT-4o, informed by rubric-based scores from a fine-tuned GPT model and system prompts detailing predefined mediation steps, guides them through self-reflection of essay elements, and provides revision suggestions.

We present two issues observed in our L2 AI-DA system that are central to its refinement into a more emotionally intelligent pedagogical tool. While adhering to an implicit-to-explicit sequencing of mediation, the system currently lacks responsiveness to learner moves that characterizes person-to-person interaction. Moreover, the system employs abstract and metalinguistic language by ChatGPT rather than attuning language to the level and needs of the individual. These issues may limit the quality of diagnostic information obtained through the procedure while also adversely impacting learners' experiences with the system and motivation to continue to engage with it. Reporting selected reflections from pilot study participants on emotional responses to DA, including satisfaction and confusion, we then elicit feedback from the audience for the model's continued development.

Can AI Write like an Undergraduate Student? A Multidimensional Analysis Looking at Disciplinary Variation

Larissa Goulart, Montclair State University
Wesley Acorinti, Northern Arizona University
Yejin Jung, University of Utah
Marine Matte, Instituto Federal de Educação, Ciência e Tecnologia Sul-rio-grandense, Brazil

The emergence of generative AI tools like ChatGPT has raised important discussions about the nature of student writing and the role of AI in educational settings. As AI-generated evolves, instructors must assess how closely these tools can approximate authentic student writing across different academic disciplines. Prior research (e.g., Goulart et al., 2025; Berber-Sardinha, 2024) suggests that while large language models (LLMs) excel at replicating formal academic genres, they tend to struggle with assignments that demand personal voice or emotional nuance—qualities often emphasized in the humanities and social sciences.

This study investigates the extent to which ChatGPT-4o can simulate undergraduate student writing across four disciplinary domains: Arts and Humanities, Physical Sciences, Life Sciences, and Social Sciences. Using a matched corpus of student-authored and AI-generated texts based on identical prompts, a multidimensional analysis (Biber, 1988) was conducted to examine key linguistic dimensions associated with formality, personal involvement, and dependent reference.

Findings reveal significant discrepancies between student and AI writing in Dimensions 1 (involved vs. informational production), 2 (narrative vs. non-narrative concerns), and 3 (situationally dependent vs. elaborated reference). Notably, divergences are more pronounced in the Arts and Humanities and Social Sciences—disciplines that typically require a higher degree of emotional engagement and contextual sensitivity—while technical disciplines tend to show greater alignment with formal academic writing.

These results have important implications for future applications of AI for writing instruction, particularly in understanding how features of involvement and disciplinary nuances impact AI's ability to authentically simulate human communication in academic writing.

Between Irony and Indictment: Exploring the Discursive Dimensions of Social AI Interpretations of Politicized COVID-19 Imagery

Yara Maria de Toledo Dias Romeiro, Pontifical Catholic University of São Paolo, Brazil
Tony Berber Sardinha, Pontifical Catholic University of São Paolo, Brazil

This study explores how Social AI interprets politically and emotionally charged visual content circulated on Brazilian Twitter during the pandemic. Focusing on tweets containing the term genocida, 970 unique images were selected using a shell-based pipeline to filter high-loading images from a previous Lexical Multidimensional Analysis of a larger tweet-image corpus. To

assess how AI systems construct meaning from these visuals, a prompt-based methodology was employed using GPT-4. Each image was described using three prompts: one objective ("Describe the image objectively"), one subjective ("Describe the image focusing on its emotional and political impact."), and one subjective with the corresponding tweet ("Describe the image focusing on its emotional and political impact based on the accompanying tweet"). This process yielded 2,910 descriptions, forming a new corpus for lexical analysis. Descriptions were annotated and lemma frequencies were analyzed through Lexical Multidimensional Analysis (Berber Sardinha & Fitzsimmons-Doolan, 2025). The resulting dimensions were interpreted through semantic clustering of high-loading variables in each pole. This revealed latent oppositions in how the model construed visual content when prompted from different interpretative angles. By combining fine-grained prompt engineering, automated corpus construction, and dimensional interpretation, this research provides a replicable framework for probing AI-generated discourse under ideologically loaded conditions. It contributes to the field of Social AI by evaluating how current LLMs represent affect, ideology, and context when tasked with interpreting complex, real-world visual input. The study also contributes to the field of Social AI by assessing how language models simulate human discursive engagement with affective and political meaning.

Poster and Technology Demonstration Abstracts

Friday, September 26, 2025 1:00 - 2:00 pm

Posters

Is AI a Competent Guide? Learner and Teacher Perceptions on AI's Role in Developing Intercultural Communicative Competence

Zeynep Saka, Syracuse University

This study explores the role of artificial intelligence (AI) as a potential guide for developing Intercultural Communicative Competence (ICC) in English as a Foreign Language (EFL) contexts. Using a mixed-methods design, data was collected from 135 participants - 89 university students and 46 language instructors - across nine state universities in Turkey. Perceived ICC was measured through a 25-item Likert-scale questionnaire adapted from Shirazi and Shafiee (2017), while actual ICC performance was assessed using two sets of Discourse Completion Tasks (DCTs): the first completed independently, and the second with the option to consult any AI tool. Responses were analyzed to evaluate the influence of AI use on culturally appropriate language production, particularly operationalized in Speech Acts and Politeness Theory.

Participants also answered open-ended questions about AI's usefulness in supporting ICC. Preliminary statistical and thematic analyses revealed discrepancies between perceived and demonstrated ICC. While many participants appreciated AI's structural support and practicality, they expressed concerns about its limited cultural sensitivity and emotional awareness.

Findings provide early insight into how AI is perceived and used in socially and emotionally complex communicative tasks, and underscore the need for ethically designed tools that support - rather than replace - human judgment in language education.

Exploring the Integration of Affective Computing into the Italian Scenario-Based Assessment

Giulia Peri, University for Foreigners of Siena, Italy
Sabrina Machetti, University for Foreigners of Siena, Italy

The Italian Scenario-Based Assessment (I-SBA) is an online language test measuring situated Italian language proficiency (Peri et al., forthcoming). Grounded in the Learning-Oriented Assessment (LOA) framework (Purpura & Turner, 2018), integrates instructional and socio-cognitive dimensions into the assessment experience.

This study reflects on the potential integration of affective computing within the I-SBA. Given that emotional intelligence is fundamental to effective social interaction (Pantic & Rothkrantz, 2003) and the affective dimension has been increasingly recognized as a pivotal element influencing language acquisition and proficiency (Turner & Purpura, 2016), affective computing could enhance language assessment by providing insights into learners' confidence, stress, and engagement through voice, facial expressions, and keystroke analysis (Voss, 2024). It may support both the speaking component and the instructional chats embedded in the I-SBA, providing also personalized, emotionally responsive feedback (Guo et al., 2022). Given theoretical priorities and practical constraints, the study also considers which affective dimension may be most relevant and feasible to integrate first, depending on available tools for Italian or the need for new development.

In the Italian context, where the integration of affective computing and language assessment

studies is still unexplored, this proposal represents a pioneering step. This poster presents methodological considerations for a future pilot.both resources, checking whether this AI is capable of capturing the most recurring speeches.

AI Reformulation of Policy Documents: Linguistic and Discursive Implications for Educational Administration

Fernanda Peixoto Coelho, Pontifícia Universidade Católica de São Paulo, Brazil
Tatiana Schmitz de Almeida Lopes, Pontifícia Universidade Católica de São Paulo, Brazil
André Luis Andrade Lansac, Pontifícia Universidade Católica de São Paulo, Brazil

Official education policy documents constitute complex records of performance indicators, legal statutes, guideline provisions, and compliance requirements that education administrators must consult to support decision making. Their style, however, can hinder comprehension and timely implementation. This study examines the feasibility of using generative AI to re-author such documents for improved readability while preserving institutional meaning. A parallel corpus was compiled from internal management and policy texts produced by a state-wide technological education authority in southeastern Brazil and an equal-sized AI-generated rewrite set produced by GPT under fidelity constraints. Lexical Multi-Dimensional Analysis (LMDA; Berber Sardinha & Fitzsimmons-Doolan, 2025) was applied to the human- and AI-generated subcorpora to identify latent discursive dimensions. Discourses are understood as socially situated representations conveyed through language use through which meaning is assigned to material and social phenomena (Berber Sardinha & Fitzsimmons-Doolan, 2025; Hajer, 1993). LMDA extracted dimensions consisting of correlated lexemes, which were interpreted qualitatively. Readability metrics (e.g., Flesch-Kincaid grade level) were also computed. Preliminary findings suggest that generative AI alters key features of governance discourse, indicating the need for caution in its use, especially where fidelity to institutional voice and regulatory intent is essential. Dimensions and statistical analyses will be explored in the poster session.

Guided by AI: Language Acquisition and AI Literacy Through Scaffolded Integration

Krisztina Domjan, American University

Academic English instructors and their students face increasing complexity when introducing generative Artificial Intelligence (AI) chatbots to support language acquisition, for international learners navigating unfamiliar academic norms. This poster presents a qualitative case study involving 29 first-year international students in an Academic English course who engaged with AI through three scaffolded assignments: a cultural identity journal, a migration-themed infographic, and a military video analysis. Drawing from chat transcripts, student submissions, and AI disclosure statements, the poster shows how AI can function as a tutor, structuring aid, or reflection partner when intentionally integrated into instructional design. Thematic analysis reveals four patterns of AI interaction: (1) AI as a confidence booster, (2) AI used as an organizational tool, (3) misunderstandings between AI as a mentor vs. a shortcut, and (4) resistance or ethical gaps in AI use. These are interpreted through the lens of the SAMR model (Modification), ISTE Empowered Learner Standard (1.1), and the Acceptable AI Use Scale (AAIUS). When students received well-structured prompts and transparent guidance, AI-supported reflection, and planning, concept mastery occurred. A visual framework, the AI Engagement Spectrum, is introduced to represent the spectrum of student mindsets, ethical choices, and metacognitive engagement when using AI.

Generative AI Feedback to Foster Growth Mindsets in Highly Sensitive Students

Aubrey Michelle Sahouria, Center for Applied Linguistics

This work in progress explores the utility of generative AI feedback for developing growth

mindset behaviors in highly sensitive students. The disposition of highly sensitive people (HSPs) is characterized by differences in processing sensory information, often exhibited as heightened environmental sensitivity and emotional reactivity. Written feedback in academic environments has been demonstrated to elicit negative emotions in students upon reception, and HSPs with lowered excitation thresholds may have heightened emotional reactions to feedback, impeding the adoption of a growth mindset. Feedback requested and received from generative AI, however, has been reported to engender positive emotions and reduce anxiety during student writing revision. This research attempts to answer how and to what extent receiving AI-generated compared to human feedback may foster growth mindsets in HSPs. Research was conducted in a transdisciplinary framework to investigate the interactions of HSP traits with feedback modality for writing and the development of growth mindsets. Synthesis of literature suggests HSPs may find the feedback process more challenging and taxing than non-HSPs, that both human and AI feedback can simultaneously facilitate and hinder HSP growth mindset development, and generative AI may serve as a valuable tool for HSPs to gradually build resilience in processing critical feedback.

The Social and Intercultural Mediation: Building Bridges in Global Communication

Fernanda Peixoto Coelho, Pontifical Catholic University of São Paolo, Brazil
Tatiana Schmitz de Almeida Lopes, Pontifical Catholic University of São Paolo, Brazil
Adilson Gomes, Pontifical Catholic University of São Paolo, Brazil
André Luis Andrade Lansac, Pontifical Catholic University of São Paolo, Brazil
Maurício Ayres Cunha, Metropolitan University of Santos, Brazil

The advancement of Artificial Intelligence (AI) in intercultural communication raises questions about its ability to handle emotions across different languages and cultures. This study investigates the potential of emotionally intelligent AI systems in translation and intercultural mediation, with an emphasis on metaphors as cognitive and cultural structures (Berber Sardinha, 2009, 2011). Grounded in Applied Linguistics (Pennycook, 2001; Kramsch, 2009) and Discourse Analysis (Baker, P., 2022; Fairclough, 2003), the research analyzes how models such as ChatGPT, Google Bard, and Microsoft Translator translate emotionally charged expressions and metaphors in Portuguese, English, and Spanish. Using methods from Corpus Linguistics (Berber Sardinha, 2004; Berber Sardinha & Fitzsimmons-Doolan, 2025), the results of machine translations are compared to a corpus of human translations to assess emotional fidelity and cultural sensitivity. The study found that, despite their effectiveness in literal translation, AI models struggle with irony, sarcasm, and culturally specific metaphors, revealing limitations related to algorithmic bias and contextual understanding (Bender et al., 2021). The research highlights the ethical challenges of emotional mediation by AI and advocates for greater transparency, data diversity, and human oversight, contributing to the development of a socially responsive AI (Nass & Moon, 2000), sensitive to the complexity of global communication.

BoodleBox-Assisted PBL for Scaffolding Collaborative Inquiry

Haiyan Li, Purdue University

This study examines the integration of AI into project-based learning (PBL) to enhance collaborative research and presentation outcomes in undergraduate education. Conducted in Purdue University's SCLA 111: Language and Cultural Exchange II, the intervention employed BoodleBox, an AI-augmented platform, to scaffold team-based inquiry into systemic U.S. sociocultural issues (e.g., educational equity, labor-leisure paradigms). Undergraduate teams (n=32) utilized BoodleBox's multi-agent architecture over 12 weeks, synthesizing theoretical frameworks (e.g., Hall's Iceberg Model) with mixed empirical data (surveys, interviews) while addressing challenges in collaborative engagement and research organization.

Data from surveys, presentation posters, discussion and reflective journals revealed that AI scaffolding enhanced students' ability to frame research questions, gather/summarize sources, and streamline collaborative workflows, yielding deeper inquiry and polished outputs. Preliminary results highlighted improved teamwork dynamics, heightened research engagement, and more effective presentations, with participants reporting increased motivation. However, efficacy varied across learner profiles. These findings demonstrate how AI-mediated PBL can transform collaborative inquiry in undergraduate education, particularly by structuring complex research processes and amplifying team productivity. The innovation lies in AI's dual role as a collaborative facilitator and research organizer, suggesting a scalable model for inquiry-driven learning. Future research should examine long-term impacts and compare AI tool efficacy across instructional modalities.

Synthetic Authenticity: Performativity and Enregisterement in AI-Generated Language from Bot Conversations

Daniel Murcia, The Pennsylvania State University

This paper centers on the analysis of an AI-generated conversation created using Google's GEMMA Language Model: Notebook LM, an advanced generative AI tool, allows the synthesis of ideas and generates dialogue based on multiple source texts. The conversation was crafted as a synthetic interaction between virtual entities, designed to represent key arguments and insights from three academic articles spanning distinct fields: computer science, philosophy, and pharmacology. The first article, Schneider (2021), addresses AI's role in language technologies, particularly its impact on language diversity and linguistic hierarchies. The second article, Coffin (2021), delves into the philosophical notion of the "machinic unconscious," exploring how environmental and technological systems shape subconscious influences. Finally, the seminal paper by Lowry et al. (1951) was chosen for its historical significance and as the most cited paper in the Web of Science Index, with 305,148 citations. Together, these articles provide an interdisciplinary lens through which to examine AI's impact on language, thought, and scientific processes. Guided by the theoretical frameworks of Agha (2007) and Silverstein (2023, 1976), this paper investigates how AI bots enact socially recognizable identities and adapt linguistic strategies to perform specific "experiential" and "emotional" roles. For more grounded theoretical support, Wolfram's (2023) insights into the mechanics of neural networks for GenAI offer a technical understanding of how language models structure and replicate human communication. Agha's concept of enregisterment and Silverstein's notions of contextual shifters frame the analysis of how bots simulate conversational practices that align with human rituals: "The Podcast".

Co-Creating Clarity: AI as a Collaborative Partner in Navigating Emotional Complexities of Educational Migration

Adekunmi Olatunji, University of Hawai'i at Mānoa

In this autoethnographic study, I examine the intricate emotional and social challenges

experienced by a bi-cultural, multilingual educational migrant and transnational English language teacher navigating the sociocultural and geopolitical context of Honolulu, Hawai'i. Central to this exploration is the iterative use of Generative Artificial Intelligence (Gen AI) chatbots (Google Gemini and Microsoft Co-Pilot) as co-constructive partners in emotional processing and scholarly inquiry. The study details how Gen AI chatbots facilitated the articulation of vague emotions and acted as a research collaborator, accelerating the exploration and application of theoretical frameworks (e.g., posthumanism, intersectionality) through the collaborative synthesis of literature and the iterative development of conceptual connections to extract meaning from periods of emotional discomfort. Analyzing personal narratives and journal reflections through the lens of Black Feminist Posthumanist New Materialism, the research investigates the interplay of identity, power, technology, embodiment, and environment in shaping my experience as an international doctoral student and the role of Gen AI within this context. I demonstrate the co-creation of a coherent narrative structure, emphasizing how I, as researcher, and Gen AI worked together to challenge assumptions, refine arguments, and ensure analytical rigor. I also critically reflect on the ethical considerations of employing Gen AI in deeply personal explorations. I argue for the transformative potential of reframing AI as a collaborative partner in well-being, education, and research, enhancing a capacity to understand the multifaceted nature of human emotion and the nuanced interplay between human and artificial intelligence.

Technology Demonstrations

Context-Aware Intelligent Learning System with Multi-Agent AI Architecture

Venkat Podugu, Q3 Learners

Traditional educational platforms often treat all learners the same, leading to frustration and untapped potential. Our Ciela framework reimagines education through Context-Aware, Intelligent, Emotionally-Tuned, Learning Science-Driven, and Agentic & Adaptive personalized learning experiences.

Our demo showcases a collaborative multi-agent AI architecture featuring four specialized agents:
- A Personalization Agent that builds evolving student profiles
- A Recommendation Agent that suggests optimal learning paths
- An Explanation Agent that adapts content delivery to individual learning styles
- A Feedback Agent that provides emotionally tuned, real-time guidance

The system demonstrates intelligent adaptation and visual progress tracking. Unlike static educational tools, Ciela evolves with every interaction, creating truly personalized learning journeys.

During the demo, attendees will witness:
- Adaptive content generation
- Intelligent learning path recommendations
- Dynamic progress visualization
- A "what-if" scenario feature showing instant adaptations based on changing student preferences and learning contexts

Q3Learners™ empowers every learner to feel seen, supported, and successful through context-aware, emotionally intelligent learning experiences.

Pangea Chat: AI-Enhanced Peer-to-Peer Conversation

Will Jordan-Cooley, Pangea Chat

Pangea Chat is an intelligent language learning environment which facilitates “learning by doing” within the context of peer-to-peer instant messaging conversation. Pangea Chat integrates human and artificial intelligence (AI) to immediately augment student communicative abilities while fostering learning and supporting educator practices. By building game, learning and assessment mechanics into AI-powered language tools, Pangea Chat helps students produce and understand their target language while supporting engagement and measuring understanding. Within the open-ended possibility space of natural language, Pangea Chat instantiates core principles of ideal educational games: freedom of choice; integrated content and gameplay; adaptive feedback and instruction; and contextually-embedded assessment. To mitigate growing concerns of security and data control, with both social media and AI, Pangea Chat is built on a decentralized and open-source protocol to allow institutions the option to self-host while building a connected network for language and intercultural exchange.