



TSLL

Technology for Second Language Learning

2023

20th Annual Technology for Second Language Learning Conference
October 19-21, 2023
Iowa State University

Table of Contents

<u>Welcome</u>	01
<u>Conference Sponsors</u>	02
<u>Conference Organizing Team</u>	03
<u>Abstract Reviewers</u>	07
<u>Conference Information</u>	08
<u>Session Guidelines</u>	09
<u>Schedules</u>	12
<u>Plenary Sessions</u>	24
<u>Abstracts</u>	31

Welcome



Dear TSL2023 Participants,

Welcome to the 20th annual Technology for Second Language Learning Conference at Iowa State University. The conference theme, “Advancing Technologies – Expanding Research,” signifies the multiple strands of research and practice that are exploring the potentials, uses, and implications of AI technologies for language teaching, learning, assessment, and research. AI in applied linguistics is not new, but the widespread accessibility of generative AI built on large language models has opened unprecedented possibilities for our work in applied linguistics. The TSL conference this year creates a forum for discussion of these possibilities and presentation of the new avenues for research as well as teaching and assessment practices prompted by the new technological landscape we occupy today. The conference program was built from the contributions of dozens of professionals from around the world who recognize the imperative to learn how new AI technologies change assessment of spoken and written language, language teaching and learning, corpus research on language use, and teacher education. Changes range from the technical and material to the philosophical and ideological. There is much to discover, discuss, and learn, so we started planning TSL2023 by inviting forward-looking scholars who have been tackling these issues for years.

We are fortunate to have six of these scholars as invited keynote speakers <https://apling.engl.iastate.edu/conferences/technology-for-second-language-learning-conference/tsll-2023/plenary-speakers/> whose work on the use of AI in applied linguistics and education is charting directions for the expanding community that recognizes the importance of AI in language education.

In addition, we are eager to welcome many of the professionals who responded to our call for proposals to present and discuss their research. From over 100 submissions, the review process resulted in 50 papers, four posters, and three roundtables to be presented beginning at 16:00h Central Daylight Time on Thursday, October 19 in Ames, Iowa. All sessions, including the periods for social interaction will be available to participants attending locally and those who are joining remotely. Whatever your mode of participation, we invite you to take a virtual peek at the Iowa State University campus where the in-person sessions will be held. We invite you to a virtual visit campus visit <https://www.iastate.edu/life-at-iowa-state/campus-and-ames> You can also visit the site of the program in Applied Linguistics & Technology <https://apling.engl.iastate.edu/> as well as the one recounting the history of the first digital computer at Iowa State University <https://jva.cs.iastate.edu/> We are grateful to the Department of English and the program in Applied Linguistics & Technology for making TSL2023 possible through their sponsorship. We thank the conference organizing committee and applied linguistics faculty for their enthusiastic support of this event. We thank all participants whose contributions have created an excellent program and rich opportunity for all to learn. On behalf of the entire organizing committee, we wish you all a stimulating conference experience and an enjoyable visit to Iowa State University!

Sincerely,

Carol A. Chapelle, Gulbahar Beckett, and Jim Ranalli

Sponsors

Department of English



**Applied Linguistics & Technology
Iowa State University**



Organizing Committee



Dr. Carol A Chapelle
Iowa State University



Dr. Gulbahar Beckett
Iowa State University



Dr. Jim Ranalli
Iowa State University



Mahdi Duris
Iowa State University

Organizing Committee



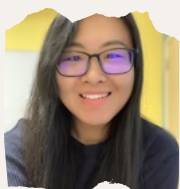
Jeanne Beck
Iowa State University



Fatemeh Bordbarjavid
Iowa State University



Noémie Sollier
Iowa State University



Hwee Jean (Cindy) Lim
Iowa State University

Organizing Committee



Widya R Kusumaningrum
Iowa State University



Inyoung Na
Iowa State University



Sara Afifi
Iowa State University

Organizing Committee



Duong Nguyen
Iowa State University



Andrias Susanto
Iowa State University



Shuhui Yin
Iowa State University

Abstract Reviewers

Dr. Elena Cotos

Dr. Fatemeh Bordbarjavidi

Dr. Gulbahar Beckett

Dr. Jim Ranalli

Dr. John Levis

Dr. Kim Becker

Dr. Kristin Terrill

Dr. Taichi Yamashita

Dr. Sarah Huffman

Dr. Sinem Sonsaat-Hegelheimer

Agata Guskaroska

Ali Garib

Altay Ozkul

Amy Walton

Ananda Astrini Muhammad

April Tan

Ella Alhudithi

Emily Dux Speltz

Fatima Aseeri

Febriana Lestari

Haeun Kim

Hitoshi Nizawa

Hwee Jean (Cindy) Lim

Inyoung Na

Jayme Wilken

Jeanne Beck

Junghun Yang

Mahdi Duris

Mutleb Alnafisah

Noémie Sollier

Sara Afifi

Reza Neiriz

Sarah Davis

Shangyu Jiang

Shireen Baghestani

Shuhui Yin

Stacy Amling

Thomas Elliott

Xixin Qiu

Wren Bouwman

Zoë Zawadzki

Thank You

Conference Information

Break and Social Time

The Break/Social time is designed for in-person and remote attendees. This time will be reserved for exchanging and enjoying some academic conversations. Please refer to the day's schedule for specific room locations.

Dinner and Social Time

All in-person presenters will be invited to complimentary lunch and dinner during the conference. Please refer to the schedule for locations.



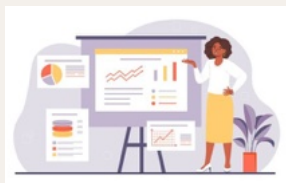
Wifi information

- The wifi is available for all participants: IASTATE-Guest
- No password is required. However, you will be asked to sign in to use the wifi. Please follow the pop-up instruction.
- Click "Accept" and you can connect to the Internet

Online participation

All online registered participants will receive an email with Zoom instructions and passcodes to access the conference. Please be sure to use the TSSL 2023 Detailed Schedule page to find the presentations you wish to attend. It is easily accessible here: <https://bit.ly/tssl23schedule>

Session Guidelines



Individual paper

- **In-person**

Paper presentations delivered in person are 20 minutes with five minutes for Q&A, followed by a five-minute transition. We will provide a laptop with audio and visual connections. Presenters may also choose to come with their laptop or similar device with the ability to connect via HDMI. (We will have a dongle ready if your computer needs an adapter).

- **Remote**

Paper presentations presented remotely are 20 minutes with five minutes for Q&A, followed by a five-minute transition. Specific Zoom links and passcodes will be sent before the conference starts.

Session Guidelines



Roundtable Discussion

- **In-person**

Roundtable presentations are 20 minutes. We will provide a laptop with audio and visual connections.

Presenters may also choose to come with their laptop or similar device with the ability to connect via HDMI. (We will have a dongle ready if your computer needs an adapter).

- **Remote**

Roundtable presentations are 20 minutes. A **hybrid option** is available and the remote presenter should be in their designated Zoom room five minutes before the start of their presentation. Specific Zoom links and passcodes will be sent before the conference starts.

Session Guidelines



Poster

- **In-person**

Posters for 50 minutes. Posters should be no larger than 36" H x 48"W. An easel support will be provided.

- **Remote**

Poster presentations are 50 minutes. The presenters should have a PowerPoint deck with their main poster as slide 1 and all the main parts of their poster in the subsequent slides. Specific Zoom links and passcodes will be sent before the conference starts.

SCHEDULE

Thursday, October 19, 2023

Time (CDT)	Events	Ross
3:45 p.m.	Welcome: Dr Volker Hegelheimer Department of English, Iowa State University	212
4:00–5:00 p.m.	Plenary: Peter Crosthwaite <i>The University of Queensland, Australia</i> <i>Generative AI and the end of corpus-assisted data-driven learning? Not so fast!</i>	212
Posters		
5:00 – 6:00 p.m.	Kate Knill, et al., <i>Look Who’s Speaking</i>	403 & 412
5:00 – 6:00 p.m.	Shan Chen, et al. <i>ChatGPT in Language Learning: A Content Analysis of Bilibili Videos</i>	403 & 412
5:00 – 6:00 p.m.	Agata Guskaroska <i>Exploring Acceptance of ASR Tools for Pronunciation Learning: ChatGPT for qualitative data analysis</i>	403 & 412
5:00 – 6:00 p.m.	Shangyu Jiang <i>The Discourse Styles of ChatGPT: A Corpus-Based Study</i>	403 & 412
5:00 – 6:00 p.m.	Dinner/Social	212
Individual Paper		
6:00 – 6:30 p.m.	Kevin Hirschi, et al. <i>Data-Driven Learning for pronunciation: A study of prominence and lexical stress in the EAP context</i>	212
6:00 – 6:30 p.m.	Kun Sun. <i>Reshaping College English Teaching and Learning in China: An AI Perspective</i>	412

SCHEDULE

Thursday, October 19, 2023

Time (CDT)	Events	Ross
6:30 – 7:00 p.m.	Hakan Cangir <i>Detecting Writing Quality in L2 English: At the Crossroads of Learner Corpus and NLP</i>	212
6:30 – 7:00 p.m.	Joseph Collentine et al., <i>Promoting Learner–Computer Interaction with Extractive Question–Answer Technologies</i>	412
7:00 – 7:30 p.m.	Danjie Su et al., <i>Evaluating Pragmatic Competence of Artificial Intelligence with the Lens Concept: ChatGPT–4 for Chinese L2 Teaching</i>	212
7:00 – 7:30 p.m.	Solbee Kim, et al., <i>Gearing up for the ChatGPT Era: Technological Trajectory in Second Language Writing Education</i>	412
7:30 – 8:00 p.m.	Marus Mkrtchyan et al., <i>The Use of AI in Teaching Translation</i>	212
7:30 – 8:00 p.m.	Laela H Nurbatra et al., <i>Virtual Tourism in Applying Students’ Speaking Skills: A Case Study in Indonesian EFL Classroom</i>	412

SCHEDULE

Friday, October 20, 2023 (Morning Session)

Time (CDT)	Events	Ross
7:50 – 8:00 a.m.	Opening: Dr. Carol A. Chapelle	212
8:00 – 9:00 a.m.	Plenary: Mike Sharples The Open University, UK <i>Expanding Pedagogy: New Ways of Teaching, Learning and Assessment with AI</i>	212
9:00 – 9:30 a.m.	Break/Social	212
Individual Paper		
9:30 – 10:00 a.m.	Joanna Baumgart, et al., <i>An Empirical Analysis of the use of AI-based Text Tools in Academic Writing: Attitudes, Usefulness and Interactions</i>	212
9:30 – 10:00 a.m.	Serge Bibauw <i>Which interactivity matters in TSLL? Agency, engagement and negotiation in conversational AI</i>	412
10:00 – 10.30 a.m.	Massimo Innamorati <i>Efficient Automated Essay Scoring using Feature Extraction from Transformer-based Large Language Models</i>	212
10:00 – 10.30 a.m.	Aya Owada et al., <i>Design based research: ChatGPT-mediated speaking practice</i>	412

SCHEDULE

Friday, October 20, 2023 (Morning Session)

Time (CDT)	Events	Ross
10:30 – 11:00 a.m.	Thomas Stringer et al., <i>Empowering EFL Learners and Teachers: Harnessing AI for Writing Instruction</i>	212
10:30 – 11:00 a.m.	Ulf Schuetze <i>Speaking to an AI chatbot as a foreign language learner: A pedagogical perspective</i>	412
11:00 – 11:30 a.m.	Roman Lesnov et al., <i>Exploring the Pedagogical Potential of VR- Based Language Learning Applications: Immerse vs. ImmerseMe</i>	212
11:00 – 11:30 a.m.	Okim Kang et al., <i>Technology-Based L2 Intelligibility Feedback: Balancing Global and Local Pronunciation Feedback</i>	412
11:30 a.m.– 12:00 p.m.	Ziqian Zhou et al., <i>ChatGPT-3.5 as an Automatic Scoring System and Proofreader: Strategies for Grading and Revising Essays</i>	212
11:30 a.m.– 12:00 p.m.	Shilin Gao and Mark Gales <i>Detecting Aberrant Responses in L2 Spoken English Automated Assessment</i>	412
12:00 – 1:00 p.m.	Lunch/Social	212

SCHEDULE

Friday, October 20, 2023 (Afternoon Session)

Time (CDT)	Events	Ross
1:00 - 2:00 p.m.	Plenary: Mark Warschauer University of California, Irvine The Open University, UK <i>Researching Generative AI in Writing</i>	212
	Individual Paper	
2:00 - 2:30 p.m.	Jining Han et al., <i>Using ChatGPT for EFL writing feedback: Teachers' perspectives</i>	212
2:00 - 2:30 p.m.	Haiping Wang et al., <i>Building validity arguments for the use of measures of utterance fluency in Automatic Speech Evaluation</i>	412
2:30 - 3:00 p.m.	Sara Afifi, et al., <i>Student Engagement with Teacher and Automated Written Corrective Feedback on L2 Writing: A Multiple Case Study</i>	212
2:30 - 3:00 p.m.	Yichen Jia <i>Exploring the affordance of ChatGPT as a Rating Tool in assessing interpreting accuracy</i>	412
3:00 - 3:30 p.m.	Esmael Ali Salimi, et al., <i>Exploring Meaning-packing Use in EFL Learners' Writing Over Genre-based Pedagogy in Augmented Virtual Reality</i>	212

SCHEDULE

Friday, October 20, 2023 (Afternoon Session)

Time (CDT)	Events	Ross
3:00 - 3:30 p.m.	Haeun Kim, et al. <i>ChatGPT for writing evaluation: Examining the accuracy and reliability of AI-generated scores compared to human raters</i>	412
3:30 - 4:00 p.m.	Sibel Söğüt <i>AI-powered Tools as Agents of Change in Pre-service English Language Teacher Training</i>	212
3:30 - 4:00 p.m.	Hamidreza Moeiniasl <i>Investigating the Potential of a Computerized Dynamic Assessment With a New Testing Format in Assessing Critical Reading Skills Among International College Students</i>	412
4:00 - 5:00 p.m.	Plenary: Evgeny Chukharev Iowa State University <i>Large language models in hybrid natural-language processing applications for language learning and assessment</i>	212
5:00 - 6:00 p.m.	Dinner/Social	212
6:00 - 7:00 p.m.	Plenary: Xiaoming Xi Hong Kong Examinations and Assessment Authority <i>Adaptive language learning in the new age of generative AI</i>	212

SCHEDULE

Friday, October 20, 2023 (Afternoon Session)

Time (CDT)	Events	Ross
Individual Paper		
7:00 – 7:30 p.m.	Zhang Wenxin et al., <i>Analyzing the Potential of generative AI for Text Generation in Reading Assessment</i>	212
7:00 – 7:30 p.m.	Kimberly Vinall et al., <i>How Students Use Machine Translation & Why: Insights from a Computer Tracking Study</i>	412
7:30 – 8:00 p.m.	Jean Chun, et al., <i>Comparing ChatGPT-Generated and Human-Designed Multiple-Choice Test Items</i>	212
7:30 – 8:00 p.m.	Sibusiso Cliff Ndlangamandla <i>Using Digital Story Telling to integrate technology when Learning Academic Language and Literacy in English</i>	412

SCHEDULE

Saturday, October 21, 2023 (Morning Session)

Time (CDT)	Events	Ross
7:25 a.m.	Opening: Carol A. Chapelle	212
Individual Paper		
7:30 – 8:00 a.m.	Joel Meniado <i>How Does ChatGPT Enhance Language Teaching and Learning? A Rapid Review of the Literature</i>	212
7:30 – 8:00 a.m.	Individual Paper: Lazgin Barany, et al. <i>EFL Teachers' Awareness and Opinions on the Use of Technology as 21st Century Language Education Skill</i>	412
8:00 – 8:30 a.m.	Ziqi Chen, et al., <i>Generative AI technology and language learning: language learners' responses to ChatGPT videos in social media</i>	212
8:00 – 8:30 a.m.	Betul Kinik <i>AI-powered technologies and English language teacher education</i>	412
8:30 – 9:00 a.m.	Liwei Hsu <i>How Does Artificial Intelligence Affect EFL Learners' Cognitive Load and Learning Anxiety? A Neuroscientific Perspective</i>	212
8:30 – 9:00 a.m.	Ilaria Compagnoni <i>When iVR meets AI: practices and challenges for language educators</i>	412
9:00 – 9:30 a.m.	Break/ Social	

SCHEDULE

Saturday, October 21, 2023 (Morning Session)

Time (CDT)	Events	Ross
Individual Paper		
9:30 – 10:00 a.m.	Wiwik Mardiana <i>Prospective Teachers' Competence in Creating Multimodal Chatbot-Based English Learning Media</i>	212
9:30 – 10:00 a.m.	Alison Porter et al., <i>Digital Professional Development: Can it improve language teacher sense of competence, autonomy and relatedness?</i>	412
Roundtable		
10:00 – 10:30 a.m.	Ebtehal Asiri et al., <i>Effectiveness of Technology Enhanced Project-Based Language Learning (TEPBL) Approach In L2 Pragmatics Enhancement</i>	212
Individual Paper		
10:00 – 10:30 a.m.	Alba Paz-López et al., <i>Preservice Teachers' Insights on AI in Language Education</i>	412
Plenary:		
10:30 – 11:30 a.m.	Michael Thomas Liverpool John Moores University <i>Critical Project-Based Learning and Social Justice: Implications for Digital Citizenship</i>	212

SCHEDULE

Saturday, October 21, 2023 (Morning Session)

Time (CDT)	Events	Ross
11:30 - 12:00 p.m.	Roundtable: Jeanne Beck, et al., <i>PBLL @ ISU: A roundtable on PBLL projects, works in progress, and AI connections</i>	212
11:30 - 12:00 p.m.	Individual Paper: Harshitha H <i>Primary Level Learners' Perceptions of Bilingual Voicebots in Indian ESL Classrooms: A Pilot Study</i>	412
12:00 - 1:00 p.m.	Lunch/Social	212

SCHEDULE

Saturday, October 21, 2023 (Afternoon Session)

Time (CDT)	Events	Ross
1:00 – 2:00 p.m.	Plenary: Abram Anders Iowa State University <i>Generative AI: Experiments in Writing and Learning</i>	212
2:00 – 2:30 p.m.	Roundtable: Robert Godwin-Jones, et al., <i>Integrating AI tools into instructed SLA</i>	212
Individual Paper		
2:00 – 2:30 p.m.	Charlene Polio <i>Using an AI program to assess accuracy in second language writing research</i>	412
2:30 – 3:00 p.m.	Svetlana Koltovskaia, et al., <i>ESL graduate students' use of ChatGPT for text revision: Insights into behavior, cognition, and emotion</i>	212
2:30 – 3:00 p.m.	J. Elliott Casal <i>Generative AI, Research Writing, and Ethics in Applied Linguistics: How do we Identify Human-Produced Writing</i>	412
3:00 – 3:30 p.m.	Gatot Prasetyo et al., <i>Exploring Benefits and Challenges of Quizlet's Q-chat in Indonesian Language Instruction for Military Personnel</i>	212
3:00 – 3:30 p.m.	Widya Kusumaningrum, et al., <i>Exploring the Impact of ChatGPT on English Language Learners' Email Writing Process</i>	412

SCHEDULE

Saturday, October 21, 2023 (Afternoon Session)

Time (CDT)	Events	Ross
3:30 - 4:00 p.m.	Thor Sawin <i>A five-ecosystems approach for applying AI to the language classroom</i>	212
3:30 - 4:00 p.m.	Shuyuan Tu <i>ChatGPT as an AI-powered writing assistant: How reliable it is in editing L2 writing?</i>	412
4:00 - 4:30 p.m.	Break/ Social	
Individual Paper		
4:30 - 5:00 p.m.	Carol Chapelle, et al., <i>Exploring Open AI in Language Teacher Professional Development</i>	212
4:30 - 5:00 p.m.	Andrias Tri Susanto <i>ChatGPT vs. Human IELTS Tutor: A Comparative Study of IELTS Task 2 Writing Sample</i>	412
5:00 - 5:30 p.m.	Hao Xu <i>EFL school teacher's AI assessment literacy: An exploratory study</i>	212

Plenary Sessions

GENERATIVE AI AND THE END OF CORPUS-ASSISTED DATA-DRIVEN LEARNING? NOT SO FAST!

Peter Crosthwaite, *The University of Queensland*

This talk explores the potential advantages of corpora over generative artificial intelligence (GenAI) in understanding language patterns and usage, while also acknowledging the potential of GenAI to address some of the main shortcomings of corpus-based data-driven learning (DDL). One of the main advantages of corpora is that we know exactly the domain of texts from which the corpus data is derived, something that we cannot track from current large language models underlying applications like ChatGPT. We know the texts that make up large general corpora such as BNC2014 and BAWE, and can even extract full texts from these corpora if needed. Corpora also allow for more nuanced analysis of language patterns, including the statistics behind multi-word units and collocations, which can be difficult for GenAI to handle. However, it is important to note that GenAI has its own strengths in advancing our understanding of language-in-use. For example, GenAI's ability to generate results from almost any register, domain or even language can greatly widen the scope of DDL beyond its current focus on tertiary academic English language.



Additionally, the size and speed at which current large language models like ChatGPT can be queried is unprecedented with even the best available corpus tools. I argue that both corpora and GenAI have valuable roles to play in advancing our understanding of language-in-use. By combining these approaches, language learners can gain a more comprehensive understanding of how language works in different contexts than is currently possible using only a single approach.

Plenary Sessions

EXPANDING PEDAGOGY: NEW WAYS OF TEACHING, LEARNING AND ASSESSMENT WITH AI

Mike Sharples, *The Open University*

As we come to understand the affordances and limitations of **generative AI**, it is time to flip the narrative away from “How will AI impact education?” to “What are new and effective ways of teaching and learning enabled by AI?”. In this presentation I will explore how AI can support **innovative pedagogy**. Roles for generative AI include: Possibility Engine (AI generates alternative ways of expressing an idea), Socratic Opponent (to develop an argument), Collaboration Coach (to assist group learning), Exploratorium (to investigate and interpret data), Personal Tutor and Dynamic Assessor.

I propose that future research into generative AI for education should be based on a new science of learning with AI – to include understanding cognitive and social processes of **AI-assisted learning**, exploring future roles for **AI in education**, developing generative AI that explains its reasoning, and promoting **ethical education systems**.



Plenary Sessions

RESEARCHING GENERATIVE AI IN WRITING

Mark Warschauer,

University of California, IrvineThe Open University

The development and diffusion of generative AI is ushering in the greatest disruption to writing practices in modern history. This presentation delves into the significance of generative AI within the framework of literacy theory and research. It provides an overview of five recent and ongoing investigations conducted by the UC Irvine Digital Learning Lab on generative AI, comprising experimental analyses on the quality of its scoring and feedback, classroom research on its impact in writing courses, and a case study on a second language learner's authoring techniques. The presentation concludes with discussion of a research agenda for this burgeoning field, as well as an introduction to an innovative online tool in development at UCI that promises to support both classroom pedagogy and research.



Plenary Sessions

LARGE LANGUAGE MODELS IN HYBRID NATURAL-LANGUAGE PROCESSING APPLICATIONS FOR LANGUAGE LEARNING AND ASSESSMENT

Evgeny Chukharev, *Iowa State University*

In this presentation, I will explore the affordances of large language models (LLMs), such as GPT, for building hybrid natural-language processing (NLP) applications in the fields of computer-assisted language learning and assessment. In these hybrid applications, the LLMs are combined with rule-based components that drive the LLMs in generating and understanding natural language. This permits the developer to retain control over the pedagogical or assessment procedures while simultaneously tapping into the power of the LLMs for flexible language generation and understanding. Specifically, I will discuss the affordances of LLMs in the context of three past and ongoing projects that focus on (1) the assessment of L2 interactional competence, (2) the learning of L1 argumentation discourse, and (3) the learning of text-production strategies for integrated writing tasks.



Plenary Sessions

ADAPTIVE LANGUAGE LEARNING IN THE NEW AGE OF GENERATIVE AI

Xiaoming Xi,

Hong Kong Examinations and Assessment Authority

Nowadays, generative AI (genAI) technologies such as ChatGPT and GPT 4 have become the new catchwords in education. The potential of genAI has been exploited in many domains of education, including adaptive language learning. Although genAI can support the development of some functionalities of adaptive learning, such as creating extended language inputs and associated learning exercises and assessments and providing feedback on language outputs, I argue that designing a robust, learner-first adaptive learning experience requires a principled, systematic approach, where genAI is only a facilitating tool, not a solution.

In this talk, I will discuss the science and technologies behind an adaptive learning system, especially how the emergence of genAI could potentially empower the development of materials and assessments in an efficient way. I will decompose the architecture of an adaptive learning system, focusing on the chain of inferences supporting its overall efficacy, including user property representation, user property estimation, content representation, user interaction representation, user interaction impact, and system impact.



I will provide an overview of different types of assessment used in adaptive learning and an analysis of the assessment approach, priorities, and design considerations of each to optimize its use in adaptive learning. I will then propose a framework for evaluating different aspects of an adaptive learning system. I will conclude with thoughts on high priority research and development to provide truly learner-first systems to fully empower our learners and the expanding role of genAI in future development of adaptive learning solutions.

Plenary Sessions

CRITICAL PROJECT-BASED LEARNING AND SOCIAL JUSTICE: IMPLICATIONS FOR DIGITAL CITIZENSHIP

Michael Thomas, *Liverpool John Moores University*

Computer-Assisted Language Learning (CALL) tends to be portrayed as a value-neutral field of practitioner research that is concerned with access to or use of digital technologies, particularly to enhance language proficiency, motivation or flexible learning. Such a view of digital education is often highly deterministic and may ignore the material realities and role of people in shaping the technologies we use. This struggle is again being played out in the hype that has greeted the emergence of ChatGPT and related AI technologies. Although confronted with challenges in testing-based educational systems, Project-Based Language Learning (PBLL) has become a more prominent pedagogical approach in recent years and research suggests that when used alongside CALL it may “engage language learners with real-world issues and meaningful target language use through the construction of products that have an authentic purpose and that are shared with an audience that extends beyond the instructional setting” (NFLL, 2022).



The “real-world” dimension of PBLL and digital education can, however, be easily assimilated into neoliberal notions of education as skills training or mere preparation for the world of work. This presentation considers the value of a more critical approach to CALL and PBLL through the lens of the ‘social justice and sustainability turn’ in language education to consider how it may aid language learners and teachers to understand social and economic inequalities and to develop more critical notions of digital citizenship.

Plenary Sessions

GENERATIVE AI: EXPERIMENTS IN WRITING AND LEARNING

Abram Anders, *Iowa State University*

Emerging generative AI tools like ChatGPT present intriguing opportunities to enhance writing instruction by helping learners manage cognitive load, exercise fluency, and receive real-time formative feedback. Yet, these technologies will also challenge educators to reshape traditional learning design and assessment practices. Adapting will involve fundamental shifts in our mindsets and how we conceptualize student learning. First, students will need to develop capacities for self-regulated learning as they increasingly act as the “human in the loop” for their own AI-assisted writing and learning processes. Second, educators will need to focus on promoting discipline-specific forms of evaluative judgment to help students self-assess and improve as they engage in AI-assisted work. Finally, both students and educators will need to develop AI literacies that allow them to productively and ethically interact with these technologies to support social learning and professional development.



This presentation will report on an experimental fall 2023 course on “Artificial Intelligence and Writing” that seeks to put these principles into action. It will share learning designs for the course’s weekly creative challenges which provide students engaging opportunities to experiment and reflect on the use of specific AI tools and techniques. Preliminary findings from these creative challenges will be used to reflect on the future of AI-assisted writing and learning.

Abstracts

POSTER-LOOK WHO'S SPEAKING

Kate Knill, Mark J.F. Gales, *University of Cambridge and Enhanced Speech Technology Ltd.*

Diane Nicholls, Paul Ricketts, Scott Thomas, *English Language iTutoring Ltd.*

AI has challenged and transformed the traditional educational system and it has become a valuable tool in the field of education. The recent advancement of natural language processing, machine learning and AI has a profound influence on the translation sphere as well. Though AI translation is not perfect in some cases, especially when it refers to low-resource languages, its development, and impact are noticeable. It is indisputable that in certain types of translation, AI translation is rather adequate and equivalent. So, an issue arises as to why we need translators and what we need to teach them. This paper discusses teaching methods and tools for teaching English and translation to translation students. The research focuses on AI use and its impact on their education with the example of Brusov State University. The aim of the paper is to discuss: (1) The needs of the translation student in the 21st century, (2) The role of the AI tools as translation aids, (3) The analysis of potential limitations and challenges associated with the use of AI tools, including issues related to accuracy and linguistic nuances, (4) The analysis of the potential and functionalities of the Multilingual Student Translation (MUST) platform to enhance the quality of translation, (5) The effectiveness of the use of the Multilingual Student Translation Corpus platform for the students of Brusov State University.

Abstracts

POSTER-CHATGPT IN LANGUAGE LEARNING: A CONTENT ANALYSIS OF BILIBILI VIDEOS

Shan Chen and **Weiwei Han** *Weifang University*;
Yanhong Liu *Yangzhou University*

Bilibili, a popular Chinese online entertainment platform that hosts user-generated videos, offers an open space for language learners and educators to post educational content with a focus on ESL and EFL. The emergence of ChatGPT has resulted in a surge of videos featuring GPT as a language learning tool been uploaded to the platform. This study aims to analyze and categorize these videos (n=107) to examine 1) how content creators portray and represent ChatGPT as a language learning tool, and 2) how viewers engage with such content by tallying the numbers of likes, shares, gifts and comments each video received. The results suggest that ChatGPT is mostly portrayed as a powerful language learning tool which enables ESL and EFL learners to improve their efficiency in vocabulary acquisition, serves as a personal tutor or speaking partner to practice oral English, assist in English writing, and helps with the targeted preparation for high-stake English tests. This research contributes to the understanding of ChatGPT's integration in language learning, highlighting its positive reception among learners. Insights benefit educators, learners, and developers, informing them about ChatGPT's multifaceted roles. Furthermore, it opens avenues for exploring the efficacy and pedagogical implications of AI language models in language education.

Abstracts

POSTER-EXPLORING ACCEPTANCE OF ASR TOOLS FOR PRONUNCIATION LEARNING: CHATGPT FOR QUALITATIVE DATA ANALYSIS

Agata Guskaroska, *Iowa State University*

This study explores the acceptance of Automated Speech Recognition (ASR) tools for pronunciation learning in second language education. Combining the Technology Acceptance Model (TAM) and the Technological Pedagogical and Content Knowledge (TPACK) framework, the research focuses on the factors influencing students' intentions to use these tools. Data analysis is conducted with the assistance of ChatGPT. Participants in the study are 11 ESL graduate and undergraduate students who received a two-week online training on ASR for pronunciation improvement. After the training, interviews were conducted and transcribed using WebEx ASR technology. ChatGPT assisted in data cleaning, organization, and assigning preliminary codes. The results indicate that learners expressed positive attitudes towards adopting the tool for pronunciation practice and found it easy to use. They discussed the perceived usefulness of the tool, highlighting learning gains, benefits, and design drawbacks. This research has pedagogical, research, and technological implications. Educators and students can gain insights into the acceptance of ASR tools and their benefits for pronunciation learning. Additionally, the use of ChatGPT for qualitative data analysis in applied linguistics research is emphasized. The study also contributes to ASR tool design by understanding the factors that influence acceptance, leading to the development of more effective educational technology.

Abstracts

POSTER-THE DISCOURSE STYLES OF CHATGPT: A CORPUS-BASED STUDY

Shangyu Jiang, *Iowa State University*

With the advancement of generative AI, there have been discussions about implementing AI chatbots such as ChatGPT in second language classrooms, where AI-generated text can function as an additional source of language input for learners. However, while AI-generated text is mostly cohesive and free of grammatical errors, it is yet to be discovered how AI resembles or differs from human writers in terms of stylistic production. This study addresses the gap by investigating the discourse styles of ChatGPT. A corpus was compiled using ChatGPT responses and human responses to the same questions found on four Q&A websites. An additive Multi-Dimensional (MD) Analysis (Biber, 1988) was applied to the corpus by mapping the ChatGPT and human responses onto Biber's (1988) dimensions of General English. The results showed that in general, responses generated by ChatGPT were more affective and interactional, more context-independent, and more overt in persuasion, while responses written by humans were more informationally focused, more dependent on the context, and less overt in persuasion. The results of this study provide empirical evidence for assessing ChatGPT's stylistic production, contributing to the evaluation of using AI chatbots in language teaching and learning.

Abstracts

DATA-DRIVEN LEARNING FOR PRONUNCIATION: A STUDY OF PROMINENCE AND LEXICAL STRESS IN THE EAP CONTEXT

Kevin Hirschi, Lia Martin, and Okim Kang,

Northern Arizona University

Data-Driven Learning (DDL) provides learners with language usage patterns with rich contextual information, often sourced from corpora of written or transcribed language (Pérez-Paredes, 2019). However, with aligned audio corpora, learners can also hear target words or phrases as well as connect pronunciation to the prosodic context. To date, few studies have incorporated audio corpora into pronunciation learning but have generally found positive perceptions and effects (Fu & Yang, 2019; Kartal & Korucu-Kis, 2020; Sardegna & Jarosz, 2022). However, no known study has investigated the impact of DDL on learner prosody within the English academic spoken context. To this end, this study introduces the Second Language University Speech Intelligibility (L2-USI) corpus, and the web-based L2-USI DDL interface equipped with audio-enable concordance searches of highly intelligible L2 speech targets and interactive practice tools. It then reports on an investigation of DDL for pronunciation amongst 10 L2 English learners in North American universities. After a diagnostic recording and pre-test, participants spent 90 minutes engaging with the L2-USI DDL interface to practice their pronunciation of target words and phrases identified as problematic in their diagnostic test. Participants then completed an interview and stimulated recall of their experience using DDL for pronunciation. The results of nonparametric Friedman pre/post comparisons indicate a positive impact of DDL on production of lexical stress and prominence ($p=.003$), but no effect on perceptual skills was detected ($p=.134$). Furthermore, thematic analysis of follow-up interviews outline learners' views of DDL being a valuable resource for language practice, particularly focusing on the amount of prosodic context provided. The results are interpreted within usage-based and speech acquisition theories, providing implications and resources to corpus linguists, teachers, and learners who wish to engage in DDL for L2 speaking skills.

Abstracts

RESHAPING COLLEGE ENGLISH TEACHING AND LEARNING IN CHINA: AN AI PERSPECTIVE

Kun Sun, *University Malaya*

With the widespread application of artificial intelligence in education (Baidu AI Open Platform, NetEase Cloud Classroom, etc.), college English classrooms in China are increasingly influenced and changed by AI. For teachers, the challenge lies in the change of skills and concepts. At the same time, as AI plays an increasing role in the classroom, it may cause anxiety and resistance among teachers. And it should be highlighted that students' learning habits and learning motivation will change with the AI too. This requires teachers to provide more autonomy in instructional design, focusing on learning process and experience, to stimulate learning interest and develop students' subjective initiative (Chiu, et al., 2023). Therefore, how to integrate AI into college English classrooms in China and how to enable Chinese college English teachers and students to correctly view and use AI are urgent problems that need to be discussed and solved. This article reviews the application of computer-based products (including AI) in college English and other higher education subjects over the past two decades in China. By comparing recent, cutting-edge domestic and international theoretical experimental literatures, this paper does not aim to solve the above problems but seeks to analyze the above in a structured, multi-level and objective manner.

Reference

Chiu, T. K. F., Moorhouse, B.L., Chai, C.S., & Ismailov, M. (2023). Teacher support and student motivation to learn with Artificial Intelligence (AI) based chatbot. *Interactive Learning Environments*.

Abstracts

DETECTING WRITING QUALITY IN L2 ENGLISH: AT THE CROSSROADS OF LEARNER CORPUS AND NLP

Hakan Cangir, *Ankara University*

Recent technological developments enable (semi) automatic and more reliable annotation of learner corpora, and these corpora have the potential to guide material designers, language instructors and assessors. Current trends in learner corpus research show that its integration with natural language processing (NLP) techniques can yield more powerful and pedagogically more convenient results. Following this trend, we aim to (a) introduce the intersection of learner corpora and natural language processing (b) elaborate on the related tools used to detect L2 writing quality, (c) and attempt to build a reliable score prediction model in L2 writing in this study. To achieve that, we use a learner corpus of 691 L1 Turkish-L2 English student essays and 200 more artificially created texts using n-gram models and ChatGPT. Our score prediction model consisting of linguistic features, such as lexical sophistication, textual cohesion, syntactic complexity, and grammar accuracy shows an accuracy rating of 78%. The mixed effects modelling results reveal several significant positive (e.g., word count, Mutual information score, concreteness rating, word2vec) and negative associations (connectives, grammar errors per 100 words) between the investigated linguistic features and writing quality in L2. We discuss the findings in light of the related literature in the field of second language writing and usage-based language acquisition. Additionally, we offer some suggestions considering the future potential of combining NLP and corpus techniques to illuminate second language acquisition research and enhance language teaching and assessment practices (e.g., automated grading), particularly in an environment where online or distance education has become the norm.

Abstracts

PROMOTING LEARNER–COMPUTER INTERACTION WITH EXTRACTIVE QUESTION–ANSWER TECHNOLOGIES

Joseph Collentine and Karina Collentine,
Northern Arizona University

How can SLA practitioners leverage AI solutions that do not require large-language models to create CALL-based TBLT while providing learning opportunities within interactionist conditions? To that end, we describe the (pedagogical) design principles and technologies of a web-based app we designed and implemented. It promotes learning through negotiation of meaning with chatbots built on an extractive question-answer framework (EQAF) (i.e., Hugging Face Transformers). We also provide examples of participants' interactions with the app from a pilot study involving learners of Spanish.

Learner-computer interactions approximating conversation have been difficult to implement using either rule-based or generative AI (Kim et al., 2022; Kojouharov, 2016). However, recent advances in EQAFs require only a small training set of the target language for emulating conversational interactions. EQAFs are a form of generative AI where user questions can be answered based on a small set of context/task-specific documents, which can be developed by the researcher/teacher for a particular language-learning task. The app implements CALL-based TBLT, which foments acquisition through situated cognition (e.g., real-world task goals). The present task occurs in a virtual environment, where learners explore a 3D world and interact with avatars. Most activities of this sort have been limited to providing learners with input (Author 1, 2022). With EQAF., learners obtain input via language production through negotiation of meaning with avatars (e.g., producing reformulations, requesting clarifications). Based on the implementation of the task and the pilot-study results, we conclude with suggestions for future research, design considerations, and curricular implementation.

References

Godwin-Jones, R. (2021). Big data and language learning: Opportunities and challenges. *Language Learning & Technology*, 25(1): 4–19.

Kim, H., Yang, H., Shin, D., & Lee, J. H. (2022). Design principles and architecture of a second language learning chatbot. *Language Learning & Technology*, 26(1): 1–18.

Abstracts

EVALUATING PRAGMATIC COMPETENCE OF ARTIFICIAL INTELLIGENCE WITH THE LENS CONCEPT: CHATGPT-4 FOR CHINESE L2 TEACHING

Danjie Su and **Kevin Goslar**, *University of Arkansas*

Artificial intelligence (AI) is transforming language learning (Lee, et al. 2023). ChatGPT-4, an AI system that uses large language models to generate human-like text responses, is integrated into the language learning app Duolingo to coach over 50 million learners. But how reliable is AI as a language teacher? In the quest to answer this question this study will evaluate AI's pragmatic competence, namely, the ability to use the appropriate form for a given context (Taguchi, 2018) to generate text that expresses the user's intended affective meaning. This question has not been addressed since the focus of AI language research (Kohnke et al., 2023; Farrokhnia et al., 2023) is on grammatical and semantic accuracy.

This study shows that the Lens concept (Su, 2017, 2021, 2022, 2023) —i.e., speakers' subjective evaluation of an event—can be used to evaluate AI's pragmatic competence. Lens reveals the (mis)match between form and context by determining whether AI chooses the right form to construct the intended lens. Using the discourse alternation method, corpus data (32 billion characters), and human conversations, this study develops a test set and evaluation model for the pragmatic competence of AI. The test includes tasks to judge pragmatic appropriateness, complete sentences expressing conflicting lenses, write a story from different perspectives, argue for/against a causative event, etc. The model compares AI performance to L1 patterns. We find that ChatGPT-4's pragmatic competence matches the level of native speakers, especially in creative writing. Areas for growth include a deeper linguistic understanding of pragmatics around speaker subjectivity. The findings suggest considerable promise in using AI to teach second language pragmatics, with additional work needed for non-creative writing. The findings contribute to the use of AI in second language teaching and learning by characterizing the reliability of AI for pragmatic teaching, as well as computational linguistics, especially methods to evaluate the pragmatic competence of deep learning AI.

Abstracts

GEARING UP FOR THE CHATGPT ERA: TECHNOLOGICAL TRAJECTORY IN SECOND LANGUAGE WRITING EDUCATION

Solbee Kim and Youngjoo Yi,
The Ohio State University;
Jinsil Jang, Dongshin University

As the field of artificial intelligence (AI) continually evolves, notably demonstrated by AI-powered chatbots like OpenAI's ChatGPT, there are growing concerns about the impact of these technologies on emerging multilingual students' development of writing. The advent of new technologies inevitably reshapes the world, and in the field of education, our mission is to guide students, facilitating their adaptation to this rapidly changing landscape, while equipping them to fully utilize the benefits of these emerging technologies. As we approach the era of ChatGPT dominating the classroom, this presentation seeks to acquire insights and prepare educators for the integration of this novel technology in education by exploring the historical evolution of technologies applied in second language writing instruction including technologies ranging from Google Docs to ChatGPT, alongside understanding their respective advantages and disadvantages. In accordance with this objective, our presentation will report the findings from a systematic research synthesis (Norris & Ortega, 2007) on the use of technologies for developing emerging multilinguals' writing. Collected empirical studies are analyzed based on (1) the type of technology employed, (2) the impact of technology on students' writing advancement, (3) the context of technology integration (i.e. research design, participants' grade level(s), class/research context), and (4) the reported potentials and constraints of each technology. By synthesizing these findings, this presentation will conclude with a discussion on the main themes of key findings across the collected studies, pedagogical implications for effectively integrating ChatGPT to enhance students' writing development, and suggestions for future research directions.

Abstracts

THE USE OF AI IN TEACHING TRANSLATION

Marus Mkrtchyan and Gayane Hovhannisyan,

Brusov State University

AI has challenged and transformed the traditional educational system and it has become a valuable tool in the field of education. The recent advancement of natural language processing, machine learning and AI has a profound influence on the translation sphere as well. Though AI translation is not perfect in some cases, especially when it refers to low-resource languages, its development, and impact are noticeable. It is indisputable that in certain types of translation, AI translation is rather adequate and equivalent. So, an issue arises as to why we need translators and what we need to teach them. This paper discusses teaching methods and tools for teaching English and translation to translation students. The research focuses on AI use and its impact on their education with the example of Brusov State University. The aim of the paper is to discuss: (1) The needs of the translation student in the 21st century, (2) The role of the AI tools as translation aids, (3) The analysis of potential limitations and challenges associated with the use of AI tools, including issues related to accuracy and linguistic nuances, (4) The analysis of the potential and functionalities of the Multilingual Student Translation (MUST) platform to enhance the quality of translation, (5) The effectiveness of the use of the Multilingual Student Translation Corpus platform for the students of Brusov State University.

Abstracts

VIRTUAL TOURISM IN AMPLIFYING STUDENTS' SPEAKING SKILLS: A CASE STUDY IN INDONESIAN EFL CLASSROOM

Laela Hikmah Nurbatra and Benny Dele Bintang Ananta,
University of Muhammadiyah Malang

One of the developments in artificial intelligence is virtual tourism. The incorporation of virtual tourism into language education has presented speaking classes with new opportunities and challenges. The practice of language skills can be greatly enhanced by participating in virtual tourism, which provides authentic experiences. In addition, it can be utilized as a tool for teaching literacy and improving cultural sensitivity. However, virtual tourism also presents the challenge of bridging the gap between the virtual and actual worlds in order to promote student involvement.

The purpose of this study is to examine the students' perspectives in virtual tourism-based speaking classes as well as the challenges they face and their solutions to these challenges. Employing qualitative research, the data was collected through questionnaires, focus group discussions, and observations. Students enrolled in a speaking for informal interaction class in the English Language Education Department, University of Muhammadiyah Malang, Indonesia, were selected as the research subjects. The results revealed that students have differing opinions regarding the use of virtual tourism in speaking instruction. They viewed it as an effective instrument for enhancing speaking skills, expanding cultural knowledge, and bolstering learning motivation. However, several issues were also noted, including technical ones like limited devices and unstable internet connection, as well as psychological ones like anxiety related to public speaking. As potential solutions, actions such as improving technical infrastructure, increasing the amount of time spent on educating individuals in public speaking skills, and providing feedback that is both detailed and personalized are required. This study has important implications for classroom teachers, curriculum developers, and educational institutions interested in making the most of virtual tourism to improve students' speaking skills. By considering students' voices and removing barriers, the use of virtual tours to develop students' speaking skills can be made more effective and beneficial.

Abstracts

AN EMPIRICAL ANALYSIS OF THE USE OF AI-BASED TEXT TOOLS IN ACADEMIC WRITING: ATTITUDES, USEFULNESS AND INTERACTIONS

Joanna Baumgart, Thomas Mandl, and Ulrike Bohle-Jurok,
University of Hildesheim

Writing is a fundamental component of academic success in mother tongue (L1) educational contexts and is also an integral part of the second language acquisition (L2). Writing is central to our social identities, and we are often evaluated by our control of it. AI tools for text generation are shaking these fundamental assumptions as computers can generate text without students going through learning processes first. Especially, ChatGPT has raised concerns about the integrity of academic writing.

We assume that the proficient use of AI-based writing tools is part of professional writing competence, which, however, is not acquired by oneself, but requires teaching and reflection. The goal of our interdisciplinary project is to embed AI-based tools in meaningful writing and subject didactic arrangements and to test and evaluate them in the context of a large compulsory course.

In this presentation, we report some results of an empirical study for academic writing in English as L2. Students were encouraged to work on a term paper in three stages using either a paraphrasing or a language generation tool. Our exploratory analysis uses a mixed-methods approach to show how AI technologies are used for academic writing. Text production skills are assessed using a grid at the beginning and after completion of the course. Furthermore, stimulated recall during an interview is used to reflect on revision. Pre- and post-questionnaires were used to analyze the attitudes and perceived usefulness of the tools and their integration into the information behavior. In addition, screen recordings are used to analyze the behavior during the interaction with the tools.

Although the overall attitude to the tools were positive, there were differences between tools as well as among diverse student groups.

Abstracts

WHICH INTERACTIVITY MATTERS IN TSLL? AGENCY, ENGAGEMENT AND NEGOTIATION IN CONVERSATIONAL AI

Serge Bibauw,

Universidad Central del Ecuador, UCLouvain

Interactivity has been defined in various ways: in instructed SLA, as distinguishing dialogic tasks from monologic ones (Ellis et al., 1994); in game design, as involving agency and user control. We propose a bidimensional model of interactivity in conversational AI, structured around user control and bi-directional adaptivity. This study addresses the question of how these different dimensions of interactivity are operationalized in conversational AI for language learning, i.e., when learners interact with a chatbot to practice an L2 (Bibauw et al., 2019; 2022), and how much they influence the learning experience. We conducted an experimental study with 215 teenage learners of French, who interacted with two versions of a serious game designed around guided conversations with AI-based talking characters. The two versions differed in interactivity, one being entirely free while the other replicated a dialogue completion task. We measured learners' perceptions, engagement, and learning effects on vocabulary knowledge. Contrary to our expectations, the free dialogue system was not perceived as significantly different compared to the dialogue completion task. However, there were significant differences in the perception of a pilot version of the system devoid of scaffolding and feedback mechanisms. On the other hand, the interactivity of the dialogue system increased behavioral engagement and production through trials-and-errors incentivized by the system feedback. Vocabulary learning shows similar effects for both conditions in receptive ($d = 1.16$) and productive knowledge ($d = 0.59$). Beyond certain limitations proper to the study design, we hypothesize that interactivity, understood as a set of game-like motivational characteristics such as agency, might have less impact on perceptions and effectiveness than expected. Our results refocus our understanding of the benefits of interactivity towards differences in engagement, feedback and scaffolding, and concretely realized in conversational AI through negotiation of form and meaning.

Abstracts

EFFICIENT AUTOMATED ESSAY SCORING USING FEATURE EXTRACTION FROM TRANSFORMER- BASED LARGE LANGUAGE MODELS

Massimo Innamorati,

Cambridge University Press and Assessment

I present a feed-forward neural network model that produces state-of-the-art performance in essay scoring on the Cambridge Learner Corpus First Certificate in English (CLC FCE). The model employs a combination of features extracted from a pretrained transformer-based large language model, as well as task-relevant features extracted directly from the input documents. Feature extraction from transformer models poses the question of how to reduce the hidden layer dimension and the token dimension to obtain a single most effective document vector representation. I compare a wide range of unsupervised methods for feature extraction across multiple pretrained transformer models. This development procedure allows for fast and efficient hyper-parameter optimization by reducing the number of trainable model parameters to less than 3 million with little reduction in performance. I also verify whether the task is best approached as a regression, multi-class ordinal classification, or ranking problem. I show that regression outperforms both classification and ranking in terms of evaluation metrics and computational efficiency. Lastly, I present results on the Kaggle Automated Student Assessment Prize Automated Essay Scoring (ASAP-AES) corpus and the Common Lit Ease of Readability (CLEAR) corpus. These suggest that my approach is competitive with similar models which fine-tune several million parameters, while it could still benefit from aspects of current state-of-the-art solutions such as supervised layer pooling or mixed loss functions.

Abstracts

DESIGN BASED RESEARCH: CHATGPT-MEDIATED SPEAKING PRACTICE

Aya Owada, and Sohyeon Lee,
University of Hawaii at Manoa

The purpose of this study is to investigate the potential of using ChatGPT, an AI-powered chatbot, as a scaffold in promoting language proficiency, and to investigate the extent to which AI and online technology-based task design promotes meaningful input and autonomous language learning. The study adopts the design-based research approach, employing a ChatGPT-mediated task design that aims to facilitate students' speech delivery. The major components of this research consist of three key aspects related to the use of ChatGPT and other online tools in educational settings. Firstly, it explores how task and prompt design using ChatGPT can provide learners with opportunities for meaningful interaction. The study aims to identify the specific elements in task and prompt design that encourage active engagement, expression of thoughts, and effective development of speaking skills among participants. Secondly, the research focuses on creating comprehensive input for students to improve their speech delivery. It examines the use of technology-mediated activities, such as DeepL/Google Translate for pronunciation familiarization and Google Docs voice typing for pronunciation checking. These tools are intended to enhance students' speech preparation and foster autonomous speaking practice. Lastly, this study digs into the impact of task design on participants' attitudes toward the use of AI in education. By analyzing learner feedback and reflections, we aim to understand how the interaction between AI chatbots affect attitudes, perceptions, and acceptance of AI integration in education. By addressing these research questions, this study contributes to our understanding of the effectiveness of task design using ChatGPT for meaningful dialogue, the creation of comprehensive input for speech delivery, and the impact of AI-mediated activities on participants' attitudes. The findings can inform the design and implementation of AI-powered educational tools and contribute to the broader discourse on AI integration in education.

Abstracts

EMPOWERING EFL LEARNERS AND TEACHERS: HARNESSING AI FOR WRITING INSTRUCTION

Thomas Stringer, *Kwansei Gakuin University*;
Lydia Eberly, *Konan University*

Traditional peer assessment activities, such as analysis or feedback, promote self-assessment and learning (Reinholz, 2016). AI may play similar peer-supportive roles with English as a Foreign Language (EFL) writing. Furthermore, writing instruction is a labor-intensive aspect of language teaching. AI could reduce that burden. This ‘feedback on writing processes and products’ presentation explores ChatGPT (OpenAI, 2023) as a beneficial tool for learners and teachers in EFL writing classrooms.

A study was conducted with three classes of 15 to 25 first- and second-year EFL learners at two Japanese universities. Learners with proficiency scores equivalent to CEFR A1 to B1 met two to four times weekly. ChatGPT was used in five stages to enhance instruction: 1. To generate typical EFL writing topics and genres, 2. to produce sample topic texts, 3. to rewrite sample texts using stylistic genre conventions, 4. to generate genre-typical feature analyses of the stylized texts, and 5. to revise examples of students’ writing to make them more target-like. Additionally, the presenters will highlight how learners interacted with AI-generated texts. Learners first read a ChatGPT writing genre exemplar, before planning and writing short passages. Learners then engaged in traditional peer assessment, analyzed and compared other students’ writing to an AI-revised version, before finally rewriting their own work and receiving peer feedback. This exploratory study investigates qualitative and quantitative differences between learners’ drafts before and after analyzing the AI-generated revisions, and describes teachers’ impressions of affordances and limitations with using AI-generated texts to guide EFL writing instruction.

References

- Reinholz, D. (2016) The assessment cycle: a model for learning through peer assessment, *Assessment & Evaluation in Higher Education*, 41(2), 301-315, DOI:10.1080/02602938.2015.100898
- OpenAI. (2023). ChatGPT (Mar 14 version) [Large language model]. <https://chat.openai.com/chat>

Abstracts

SPEAKING TO AN AI CHATBOT AS A FOREIGN LANGUAGE LEARNER: A PEDAGOGICAL PERSPECTIVE

Ulf Schuetze, *University of Victoria*

A general question is how communication can be carried out between a learner and an AI chatbot. Speech places a high demand on cognition. The aspects of fluency, accuracy and complexity that are used to characterize speech have been set into relation to tasks-types arguing that accuracy and complexity can increase together, but at the cost of fluency (Robinson, 2001, 2005), or that when any two of the three increase it is at the cost of the third (Skehan, 1998, 2009). Regarding fluency, turn-taking is required, that is, the learner and chatbot each ask and answer questions (Kim et al., 2022). With respect to accuracy, the chatbot would need to recognize errors, decide which ones to correct and how to correct (Kartchava & Nassaji, 2019; Loewen, 2012). Concerning complexity, the chatbot needs a corpus that constantly adjusts the vocabulary and grammatical rules according to the learner's progress to account for the development of the speaker. Further to that, one must look at pronunciation. The speech of the AI chatbot should be human-like, that is, without delay caused by generating a sentence through an algorithm. At the same time, the AI chatbot needs to account for the individual pronunciation of the learner, in relation to prosody (O'Brien et al., 2018). It also needs to accept that the pronunciation of the learner will change with advanced practice.

Presently, there is no AI chatbot with the features outlined here. Everyday AI chatbots for native-speakers are programmed differently: they only respond to questions, accept as many grammatical, vocabulary or pronunciation errors by the native speaker as possible and have limited feedback. That type of programming is not beneficial to a language learner. This paper provides details on the issues outlined above and offers suggestions on how to address these challenges.

Abstracts

EXPLORING THE PEDAGOGICAL POTENTIAL OF VR-BASED LANGUAGE LEARNING APPLICATIONS: IMMERSE VS. IMMERSEME

Roman Lesnov, *Oakwood University*

Virtual reality (VR) has gained significant attention in language education due to its ability to create immersive L2 learning environments; however, their pedagogical grounding is yet to be explored (e.g., Parmaxi, 2023). Specifically, more research is needed to assess the alignment between these VR-based environments and contemporary L2 learning theories. To address this gap, we examine two popular high-immersion VR language learning applications, Immerse and ImmerseMe. To achieve this goal, we employ a descriptive, judgmental, phenomenological, non-empirical evaluation of the two applications. We developed a rubric to evaluate how closely each application adheres to five key principles of L2 pedagogy: (1) focus on meaning, (2) comprehensible input, (3) focus on form, (4) interaction, and (5) motivation (Li, 2017). The rubric encompasses specific criteria for each principle to ensure comprehensive evaluation. Using a Meta Quest 2 VR headset, we observed, participated in, and compared ESL/EFL lessons from Immerse vs. ImmerseMe as well as Spanish lessons from Immerse vs. ImmerseMe.

After observing the lessons and applying the rubric, we evaluated the five principles on a scale of 0 (not present) to 4 (present fully) for each lesson and compared the averages between the applications with regards to each pedagogical principle. As our findings suggest, Immerse demonstrates a stronger alignment with the L2 pedagogical principles, especially in terms of focus on form and negotiation of meaning. On the other hand, ImmerseMe has certain pedagogical advantages related to learner autonomy, motivation, and classroom use. Based on these findings, we conclude that although Immerse has more pedagogical potential than ImmerseMe, both applications can be useful for independent L2 learning, depending on the context. Pedagogical recommendations derived from this study will be provided to the audience, along with handouts describing the applications, rubric used for evaluation, and our findings.

Abstracts

TECHNOLOGY-BASED L2 INTELLIGIBILITY FEEDBACK: BALANCING GLOBAL AND LOCAL PRONUNCIATION FEEDBACK

Okim Kang, *Northern Arizona University*

Kevin Hirschi, *Northern Arizona University*

John Hansen, *University of Texas-Dallas*

Stephen Looney, *Pennsylvania State University*

Automated speech analysis technologies provide increasingly accurate but complex dimensions of data for pronunciation learning and teaching programs, including segmental and suprasegmental features that are important for second language (L2) intelligible speech (Cucchiari et al., 2002; Kang & Johnson, 2018). When these data inform salient technology-based feedback, which is in turn linked with actionable steps for L2 speech improvement, learners can engage themselves in productive and autonomous learning activities such as planning, monitoring, and self-correcting. However, for rich data on numerous speech dimension to effectively facilitate pronunciation development, informative and user-friendly feedback should be provided by balancing global (a discourse-level intelligibility score), and local (e.g., speech rate, lexical stress, pause errors) information.

This presentation reports on the refinement of objective speech intelligibility feedback for 72 International Teaching Assistants (ITAs) and Intensive English Program (IEP) learners at North American universities who received technology-based pronunciation assessment feedback. The data were analyzed using nonparametric Wilcoxon tests. During the first iteration of feedback, ITAs demonstrated minimal changes in their speech patterns before and after receiving feedback in terms of their speech rate ($p=.701$), lexical stress ($p=.593$) and overall speech intelligibility ($p=.657$). However, ITAs using the second version of the feedback, which increased the saliency of lexical stress errors and linked learners to online pronunciation resources resulted in a significant increase in lexical stress accuracy ($p=.012$, $r=0.24$), but not in intelligibility ($p=.072$) or speech rate ($p=.083$). Finally, IEP learners, who only experienced the second version of the feedback, demonstrated improvement in speech rate ($p=.031$, $r=0.20$) and overall intelligibility ($p=.025$, $r=0.21$), but not lexical stress ($p=.124$). Taken together, these results indicate that learners of different proficiency levels benefit differentially from pronunciation feedback designs. Implications for learners and researchers include careful consideration of the autonomous learning flow for various learner populations.

Abstracts

CHATGPT-3.5 AS AN AUTOMATIC SCORING SYSTEM AND PROOFREADER: STRATEGIES FOR GRADING AND REVISING ESSAYS

Ziqian Zhou and **Xinming Chen**,

*Beijing Normal University-Hong Kong Baptist University
United International College*

The widespread use of AI promises, or threatens, to transform language studies. There are heated debates on whether AI can benefit students' language learning. One way to approach this issue is to evaluate the effectiveness of AI tools in students' learning processes. One aspect directly relevant to the use of AI is the development of writing skills. While ChatGPT-4 is the most recent AI model available, ChatGPT-3.5 is currently the most widely used AI language model created by OpenAI. This study, therefore, explores the use of ChatGPT-3.5 as an Automatic Essay Scoring (AES) system and a proofreader to assess and revise students' IELTS essays. The assessment draws on the four descriptors in the IELTS writing rubric: task response, coherence and cohesion, lexical resource, and grammatical range and accuracy. Data in this research involved 20 essays collected from 20 Chinese EFL college undergraduate students whose IELTS writing scores ranged from 5.5 to 6.5. The experiment adopted quantitative and qualitative methods to compare the scores given to the essays by both former examiners of IELTS and ChatGPT-3.5. We also analyzed the strategies used by ChatGPT-3.5 in revising students' IELTS essays to achieve higher scores. The findings suggest that ChatGPT-3.5 could be a promising proofreader that raises the levels of formality and naturalness of the language of the students' IELTS essays. Learners' attention can therefore be drawn to a variety of writing strategies, such as using more advanced and academic words and making simple sentences more sophisticated by using adverbials, attributive clauses, or connectives. However, the band scores given by ChatGPT-3.5 are significantly different from those awarded by former IELTS examiners. For students who are at the upper-intermediate level of English, ChatGPT-3.5 cannot be considered a viable AES system for grading their IELTS argumentative essays.

Abstracts

DETECTING ABERRANT RESPONSES IN L2 SPOKEN ENGLISH AUTOMATED ASSESSMENT

Shilin Gao, *Cambridge University Press & Assessment*;
Mark Gales, *Enhanced Speech Technology Ltd*

Automatic marking systems have become a popular tool in spoken language assessments to meet the increasing demand from English learners. One of the challenges for automated systems is how to handle aberrant responses where, for example, the responses are not meaningful or the speech is not understandable. The guidelines that human examiners refer to often have clear definitions and special instructions regarding the marking of these responses. Although the percentage of these aberrant responses are typically small, as automated systems are applied to 'higher stakes' tests it is increasingly important that systems appropriately identify and handle these situations.

In this work two forms of classifier are developed to identify aberrant responses: a feature-based multi-layer perceptron classifier, whose input includes both linguistic features and the auto-marker output; and a BERT-based classifier, whose input uses the output of an automated speech recognition system. The performance of these classifiers is initially evaluated in terms of detecting aberrant responses, and precision-recall curves are plotted for this binary classification task. The impact of detecting these responses is then evaluated in terms of automated assessment performance, where all responses are required to be marked by the automated system. Here, once the aberrant responses have been detected, the appropriate mark, according to the guidelines, can be attributed. Performance metrics including Root Mean Square Error and Pearson Correlation Coefficient are measured on a test dataset consisting of over 70,000 candidate submissions. The datasets used in this work are extracted from the spoken section of Linguaskill, an internet-based English proficiency examination. The advantages of using these classifiers are clearly demonstrated, in terms of both detection and the impact on candidate score.

Abstracts

USING CHATGPT FOR EFL WRITING FEEDBACK: TEACHERS' PERSPECTIVES

Jining Han, *Southwest University*;

Mimi Li, *Texas A&M University-Commerce*

ChatGPT is a state-of-the-art Artificial Intelligence (AI) language model trained on a wide variety of text sources and aiming to generate human-like responses regarding a broad range of topics (Rahman & Watanobe, 2023). ChatGPT's ability to provide writing feedback offers multiple advantages, such as immediate feedback (Huang et al., 2022; Kuhail et al., 2023), grammar and syntax correction (Kohnke et al., 2023), style and tone suggestions (Kasneci, et al., 2023), and scalability (Lo, 2023). This presentation reports on an empirical study using a case study approach to examine four EFL teachers' reactions to ChatGPT-supported teacher feedback in their writing instruction. It is guided by two research questions: (1) How do the EFL instructors implement ChatGPT-supported teacher feedback in writing? and (2) What are the EFL instructors' perceptions of ChatGPT-supported teacher feedback?

In this study, the four teacher participants, upon training, provided feedback on their EFL college students' expository and argumentative essays, subsequently, based on ChatGPT feedback. The specific processes involved include (1) teachers using Ferris's 15 error categories (Ferris, 2006) as the first input command for ChatGPT to provide feedback; (2) teachers requesting ChatGPT to provide a second round of feedback focusing on rhetoric; (3) teachers reviewing the two rounds of ChatGPT feedback, making revisions, and then sharing feedback with the students. Following the completion of feedback for the two writing tasks, the researchers conducted a 30-minute semi-structured individual interview with each of the four teachers. The content analysis of the interview data, supplemented with the comparisons of ChatGPT feedback and the final feedback shared with the students clearly revealed how the instructors perceived the facilitating role of ChatGPT in teacher feedback. Providing writing feedback has been widely considered a time-consuming and laborious task for EFL teachers, and ChatGPT-supported teacher feedback, the novel approach, will have great potential for writing instruction in the digital age. This presentation ends with sound pedagogical suggestions on integrating ChatGPT for teacher feedback in L2 classes.

References

- Ferris, D. R. (2006). Does error feedback help student writers? New evidence on the short-and long-term effects of written error correction. In K. Hyland & F. Hyland (Eds.), *Feedback in second language writing: Contexts and issues* (pp. 81–104). Cambridge, England: Cambridge University Press.
- Huang, W., Hew, K. F., & Fryer, L. K. (2022). Chatbots for language learning – are they really useful? A systematic review of chatbot-supported language learning. *Journal of Computer Assisted Learning* 38(1), 237–257.
- Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., ... Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language model for education. *Learning and Individual Differences*, 103, 1022274.

Abstracts

BUILDING VALIDITY ARGUMENTS FOR THE USE OF MEASURES OF UTTERANCE FLUENCY IN AUTOMATIC SPEECH EVALUATION

Haiping Wang, *East China University of Political Science and Law*;
Zoe Handley, *University of York*

This paper contributes to the validity argument for the use of measures of utterance fluency as reasonable representations of oral proficiency in Automatic Speech Evaluation (ASE) through investigations of the extent to which measures of utterance fluency reflect the construct of oral proficiency and of the underlying cognitive validity of the measures and the implications for automated scoring. Seventy-three advanced Chinese learners of English participated in the study and sixty of them completed all the seven tasks in English. These tasks were: 1) an IELTS-style speaking task, 2) the productive levels test, 3) the word associates test, 4) a picture naming task, 5) a grammar knowledge test, 6) a sentence inflection and agreement task designed to measure morpho-syntactic encoding task, and 7) a sentence transformation task designed to measure syntactic encoding. A sample of eight native speakers rated the learners' oral production for functional adequacy.

Analyses of the data confirm the following: (1) articulation rate predicts functional adequacy, (2) breadth of lexical knowledge is the main predictor of articulation rate as well as functional adequacy, and (3) measures of utterance fluency only account for 34% of the variation in functional adequacy scores. These findings suggest that it is valid to use articulation rate in ASE models, but such measures of fluency should be supplemented with other measures that more directly represent the quality of the language produced. The current project can build validity arguments to either support or deny the algorithm used in those APPs that claim to assess learners' oral proficiency automatically and to help teachers properly use those tools to assess students' oral proficiency for learning and improving.

Abstracts

STUDENT ENGAGEMENT WITH TEACHER AND AUTOMATED WRITTEN CORRECTIVE FEEDBACK ON L2 WRITING: A MULTIPLE CASE STUDY

Sara Afifi, *Iowa State University*;
Mohammad Rahimi, *Shiraz University*;
Joshua Wilson, *University of Delaware*

The present multiple-case study, based on the multi-dimensional perspective on student engagement with CF proposed by Ellis (2010), set out to scrutinize students' behavioral, cognitive, and affective engagement with written corrective feedback (WCF) provided by two different sources: teacher WCF and automated written corrective feedback (AWCF) provided by Writing Mentor. To this end, four Iranian EFL learners—two limited proficient and two modestly proficient writers—were selected purposefully from two sections of an academic writing course, one providing teacher WCF and the other AWCF. Participants in both sections wrote five argumentative essays during an academic term, received feedback on grammar, usage, and mechanics, and made revisions. The results demonstrated that the participants had different engagement levels and were categorized as highly engaged, moderately engaged, and minimally engaged students. In section 1, both participants who received teacher WCF were behaviorally and cognitively engaged with the feedback; however, one participant spent more time, used more resources, and showed more revision acts. Regardless of their behavioral and cognitive engagement level, they both demonstrated deep affective engagement with teacher feedback. In section 2, while one participant who received AWCF demonstrated deep and active engagement in all three dimensions, the other participant was reluctant to respond to the feedback and demonstrated a minimal level of engagement. Findings indicate that students' engagement with WCF, whether provided by a teacher or automated writing evaluation system, is influenced by students' beliefs and attitudes toward feedback and the sources of that feedback. Students' writing proficiency was not clearly or consistently related to their degree of feedback engagement. The study has pedagogical implications for the effective delivery of WCF. An additional pedagogical implication is that it may be necessary to combine teacher WCF and AWCF.

Abstracts

EXPLORING THE AFFORDANCE OF CHATGPT AS A RATING TOOL IN ASSESSING INTERPRETING ACCURACY

Yichen Jia, *Nanyang Technology University*

The assessment of interpreting has been criticized for its subjectivity and variability. Although the use of multiple raters can enhance rating accuracy, it is impractical in everyday assessment situations. As an alternative, it has been suggested that analytical/rubric rating can improve rating reliability and accuracy, but this can be time-consuming and labor-intensive. Fortunately, the generative AI, ChatGPT, a large language model, has the potential to understand rubrics and facilitate the analytical rating process. This exploratory study aimed to investigate the affordance of ChatGPT in interpreting accuracy assessment. A total of 36 interpreting Chinese-to-English transcripts were rated on their accuracy with a rating rubric. The AI-generated scoring was then compared with the mean score assigned by three seasoned interpreting teachers, with inter-reliability over 0.7. The study also aimed to investigate the features that differentiate AI's and human raters' assessment decisions.

Preliminary findings show that ChatGPT can understand interpreting tasks and rubrics well, generate high-quality renditions, and provide detailed feedback. Specifically, 75% of the rating results fell into the same or proximate rating categories as human scoring.

Further analysis of discrepant rating samples revealed that unfinished sentences, false starts, and repetitions occur most frequently in the samples that received significantly different scores from human raters and ChatGPT. This observation suggests that ChatGPT may have an advantage over traditional automatic scoring tools in its understanding of context, identifying repetition and marking it as redundant information. However, repetition can sometimes be used as an interpreting strategy to avoid long pauses and is covered in the target language quality in interpreting assessment. Therefore, although ChatGPT can understand the rubric well and generate the most accurate scoring, it seems to incorporate linguistic features that are not covered in the rubric. To maximize its potential, designing more computer-specific rubrics may be necessary rather than replicating human-used rubrics.

Abstracts

EXPLORING MEANING-PACKING USE IN EFL LEARNERS' WRITING OVER GENRE-BASED PEDAGOGY IN AUGMENTED VIRTUAL REALITY

Esmaeel Ali Salimi, *Allameh Tabataba'i University*;
Seyedfarid Beheshtinezhad, *Allameh Tabataba'i University*;
Mohammad Mostafa Mohammadi, *University of Zanjan*

The application of Artificial Intelligence (AI) and Augmented Virtual Reality (AVR) technologies in academia can bring about numerous advantages, particularly in the field of language teaching and education. To capitalize on these benefits, a genre-based teaching and learning model of Hyland (2004) was implemented using virtual reality (VR) headsets programmed in C#. In this approach, culture-bound conventions of the Information Report genre were taught using mingled exemplar texts in the context of eukaryotic cell in the virtually generated world for the learners. Participants in the study, including an MD student, an MA holder, a BA holder, and three K-9 and K-8 school-level students, were able to explore inside and outside the modeled cell and its context through innovative, immersive learning experiences. The teaching and learning cycles consisted of modeling and deconstruction, joint construction, and independent construction of the text. The genre analysis of the produced texts before and after the implementation was evaluated based on macro features recognized by the Sydney School, as well as micro lexicogrammatical features such as lexical density, frequency of embedding clauses, and grammatical metaphors by AI. Finally, the participants' impressions and experts' feedback about the pedagogy and texts were considered to support the findings of the study. The initial findings demonstrate that learners' genre control expanded significantly after the implementation of the AVR-Genre-based pedagogy. The results have implications for the futurology of language teaching, suggesting that AI and AVR technologies can provide valuable resources for innovative and effective teaching and learning practices and research.

Abstracts

CHATGPT FOR WRITING EVALUATION: EXAMINING THE ACCURACY AND RELIABILITY OF AI-GENERATED SCORES COMPARED TO HUMAN RATERS.

Haeun Kim, Yasin Karatay, Shireen Baghestani, Jeanne Beck, Leyla Karatay, Sebnem Kurt, Shuhui Yin, Mutleb Alnafisah
Iowa State University

Since its introduction, ChatGPT (OpenAI, 2023) has garnered the attention of many language educators. One particular way that ChatGPT can be used to support students' language learning is through second and foreign language assessment. For example, ChatGPT has been found to be useful in evaluating student writing and providing formative feedback. However, questions regarding the usefulness of ChatGPT in generating "scores" in assessing second language writing based on integrated tasks are yet to be answered. In addition, while AI prompt engineering has emerged as a new important field of research (Teubner, et al., 2023), a limited amount of research shows the extent to which ChatGPT-generated scores can vary depending on the prompt the user provides.

Therefore, the current study's aims are twofold: (1) to investigate the accuracy and reliability of the scores that ChatGPT produces compared to human raters, and (2) to find optimal ways of prompting ChatGPT to generate scores based on a rubric for an integrated writing task.

First, a random sample of 90 argumentative essays was taken from Iowa State University's English Placement Test (EPT) Corpus of Learner Writing (2017). Each essay was independently rated by three raters, and the 70 essays which received the most consistent ratings based on a many-facet Rasch model were selected for subsequent use. Ten essays from the 70 were provided to ChatGPT as benchmark responses (i.e., two essays per each of the five levels on the scale) along with the EPT writing scale descriptors. ChatGPT was then asked to rate the remaining 60 essays, and the accuracy and reliability of its ratings were compared with the human ratings in various prompting conditions. In this presentation, we will present our comparison results and discuss the impact of prompts on the rating performance of ChatGPT.

References

Teubner, T., Flath, C. M., Weinhardt, C., van der Aalst, W., & Hinz, O. (2023). Welcome to the era of ChatGPT et al. The prospects of large language models. *Business & Information Systems Engineering*, 65(2): 95-101.

Abstracts

AI-POWERED TOOLS AS AGENTS OF CHANGE IN PRE-SERVICE ENGLISH LANGUAGE TEACHER TRAINING

Sibel Söğüt, *Sinop University*

Artificial Intelligence (AI) tools have emerged as pedagogical endeavors in English language teaching. The developing applications and facilities raise concerns about potential hindrances to critical thinking skills, academic integrity, and disruptions to teachers' instructional methods, procedures of assessment, and evaluation in L2. Considering that emerging developments have led to a paradigm shift in language education, we need to empower educators and students with AI literacy skills in language education. Motivated by this need and building upon current discussions and emerging needs in teacher training programs, this study aimed to document pre-service English language teachers' grounds and connections between AI tools and transformations of current pedagogical techniques and procedures in creating lesson plans to teach L2 writing skills. Upon the introduction of practical applications of AI tools in lesson planning and task design, and the employment of a qualitative research paradigm, a researcher-developed survey and elicitation of metaphors served to gather data to reveal 28 pre-service language teachers' perspectives in a teacher training context. Based on the initial analyses of elicited codes and themes, they highlighted their concerns regarding challenges in evaluating AI-generated content, detecting violations of academic integrity, and ensuring fair assessment in the absence of clear guidelines and regulations. They also foregrounded the role of AI in disrupting existing inequalities in accessing sources. Their conceptualizations of AI in content creation and improvement in their writing skills and lesson plans were documented to provide suggestions for adopting a critical stance over the employment of these tools. The potential of AI-powered tools, transforming current pedagogical techniques and procedures in creating lesson plans and materials, and ways to revolutionize teacher training for L2 writing instruction will be discussed. This study will further provide suggestions to inform policy and practice in teacher education, paving the way for more effective and technologically enhanced language instruction.

Abstracts

INVESTIGATING THE POTENTIAL OF A COMPUTERIZED DYNAMIC ASSESSMENT WITH A NEW TESTING FORMAT IN ASSESSING CRITICAL READING SKILLS AMONG INTERNATIONAL COLLEGE STUDENTS

Hamidreza Moeiniasl,

*The Ontario Institute for Studies in Education
of the University of Toronto*

This paper examines the potential of a new testing format, grounded in Vygotsky's Sociocultural Theory, in assessing critical reading skills among college students with the aim of understanding the extent this new testing format can diagnose their level of critical reading skills when compared to the DIALNAG reading test. The impetus for developing our innovative multiple-choice testing format was to investigate if dynamic assessment, as it is claimed by its advocates, has the potential to offer a diagnosis of test-takers' critical reading skills that is more effective, precise, and succinct by offering multiple-try feedback with hints after each failed attempt. We argue that the new testing format could adequately capture the test-takers' underlying critical reading as they have to mull over the options and decide which one best reflects the ideas and concepts discussed in the text.

To examine the efficacy of our testing format, we developed a computerized Critical Reading Dynamic Assessment (CRDA) where testees were required to assign a degree of truth to the five options with regard to their proximity to the correct answer, ranging from definitely true to definitely false. The CRDA and DIALNAG reading test were both administered to 120 first-year college students in a communication course. While the results demonstrated a strong correlation between students' CRDA scores and those of their DIALNAG reading test, the profile generated by the CRDA was better able to diagnose the students' mature and maturing critical reading skills based on their responses to feedback. The test also attempts to integrate instruction and assessment in one single activity.

Abstracts

ANALYZING THE POTENTIAL OF GENERATIVE AI FOR TEXT GENERATION IN READING ASSESSMENT

Zhang Wenxin and **Vahid Aryadoust**,
Nanyang Technological University

This ongoing study aims to delve into the potential of ChatGPT, an AI-based language model, by comparing the linguistic features and content quality of its generated texts with those from two distinct corpora: a general corpus of texts and a specialized sample corpus consisting of English proficiency tests in the field of language assessment. This research employs a Multi-Dimensional (MD) Analysis approach to compare various linguistic features, including lexical complexity, syntactic sophistication, and content quality between passages derived from the two corpora with those generated by ChatGPT. Moreover, this research investigates the extent of content coverage and topic distribution in order to gain insights into the subject matter addressed in both sets of texts. The primary objective is to evaluate the potential and affordances of ChatGPT as a tool for generating reading comprehension texts for assessments from linguistic and content aspects. Through an extensive comparative analysis, this investigation provides a comprehensive understanding of the capabilities and constraints associated with AI-generated texts in the realm of language assessment. We offer suggestions for using AI-generated texts for assessing reading comprehension.

Abstracts

HOW STUDENTS USE MACHINE TRANSLATION & WHY: INSIGHTS FROM A COMPUTER TRACKING STUDY

Kimberly Vinall and Emily Hellmich,
University of California, Berkeley

Neural-network-based machine translation (MT) tools have had significant implications for world languages teaching and learning. While previous research has documented instructors' and students' beliefs about MT and its use (Case, 2015; Clifford et al., 2013; Jolley & Maimone, 2015; Niño, 2009) little is known about actual student use of MT tools. How students use MT tools and what influences this usage are important components of the larger ecology surrounding MT and language teaching/learning: from an ecological theoretical perspective, multiple factors (e.g., experience, beliefs, platform design and functionality as well as policy) interact across scale levels (e.g., individual, classroom, institution and society) to impact how digital technologies are understood and used in language learning contexts. In this talk, we draw on our recent computer tracking study to examine how learners of French, Spanish and Mandarin use MT and what influences their usage. In our study, learners completed a short, written task in the target language while we observed and recorded their screen. Thereafter, we conducted a stimulated recall, asking students to narrate key moments in their writing task and we conducted post-interviews to dive deeper into student actions and motivations.

We documented a myriad of strategies for their use of machine translation for working with language input (intentional input, changing input) and output (changing course, rephrasing, seeking examples, rechecking/triangulating output). We present these strategies with video and audio data to showcase the complexity of the actions. We also identified additional factors that influenced these strategies: (1) knowledge of language, (2) specific beliefs about online tools, (3) learners' perceptions of their own role and (4) the roles ascribed to online tools and classroom policies. We conclude by considering how instructors and researchers might draw on this research to think differently about integrating MT into classroom practices.

Abstracts

COMPARING CHATGPT-GENERATED AND HUMAN-DESIGNED MULTIPLE-CHOICE TEST ITEMS

KJean Chun, Natalia Barley, and Umer Farooq,
DLIFLC

The recent development of ChatGPT shows great promise, particularly in its ability to create test items quickly. However, the degree to which AI-generated items match the quality of human-produced items remains largely unexplored, even though this is crucial for ensuring test validity. This study compares test items generated by human writers with those by ChatGPT to investigate the quality and characteristics of ChatGPT-generated items.

To conduct the comparison, two sets of 40 multiple-choice questions were developed using listening passages in Korean. One set was written by Korean teachers in an intensive Korean language program, and the other set was generated by ChatGPT. The teachers had received item-writing training, while ChatGPT was guided by the researchers to develop items following similar guidelines. Two raters conducted a blind review of these items using five criteria: (a) the importance of the information questioned in the passage, (b) the clarity of the questions, (c) the presence of a single correct answer, (d) the distinctiveness of the distractors, and (e) the plausibility of the distractors. These criteria were established using item writing guidelines (Haladyna, 2004) and stakeholder input into the program's achievement tests. The item quality was assessed using a five-point Likert scale, and inter-rater reliability was examined. The researchers performed the Wilcoxon signed rank test to analyze the differences in ratings between ChatGPT items and human writer items for each criterion.

Additionally, the items and passages were qualitatively reviewed to identify unique features of ChatGPT-generated items that distinguished them from items written by human writers, as well as to uncover any potential impact of the passages on item writing. The study's findings discuss the potential of utilizing ChatGPT in developing test items to enhance practicality without compromising test validity.

Reference

Haladyna, T. M. (2004). *Developing and validating multiple-choice test items*. Routledge.

Abstracts

USING DIGITAL STORY TELLING TO INTEGRATE TECHNOLOGY WHEN LEARNING ACADEMIC LANGUAGE AND LITERACY IN ENGLISH

Sibusiso Cliff Ndlangamandla, *University of South Africa*

The study of technology-mediated communication has the potential for understanding 21st century language learning and constructing models of digital language use. For example, Digital Story Telling (DST) can assist with course design, and language learning, especially in view of Artificial Intelligence (AI) that seems to invite plagiarism. This paper presents research in which students were asked to write short language narratives about technologies and lifestyles. DST may present a linguistic conundrum (Wu & Chen, 2020, p. 9). This research uses an ecological approach to explain the emerging outcomes from different components of tools and people interacting across different scales. According to Goodwin-Jones (2022, p. 17), "An ecological perspective on the use of AI tools and education invites a wider, societal consideration of issues like learner agency and equity." The aim was to find out more information about the digital language and literacy practices, and multilingualism among students in the online learning context.

The study uses computer-mediated discourse analysis as a method of analyzing data. The findings indicate that intercultural communication, copy and paste, sharing and other social media practices as well as languaging is shaping sociocultural practices, linguistic diversity, and language development amongst a group of students learning English through online teaching and learning in Higher Education. Students reveal that even when they use the technologies, language, and learning can index particular cultural identities associated with a mother tongue and English as a Multilingual Franca.

References

- Wu, J. & Chen, D. T. (2020). A systematic review of educational digital storytelling. *Computers & Education* 147, pp. 1 – 16, 103786, <https://doi.org/10.1016/j.compedu.2019.103786>.
- Godwin-Jones, R. (2022). Partnering with AI: Intelligent writing assistance and instructed language learning. *Language Learning & Technology*, 26(2): 5–24. <http://doi.org/10125/73474>

Abstracts

HOW DOES CHATGPT ENHANCE LANGUAGE TEACHING AND LEARNING? A RAPID REVIEW OF THE LITERATURE

Joel Meniado, *SEAMEO Regional Language Centre*

ChatGPT, an AI-based language model designed to generate human-like responses to text-based prompts has caught the attention of many people and sparked debates about its potential implications for language education. While many academics recognize its value in supporting English language teaching, learning, and assessment, some are worried and skeptical about its threats to academic integrity, knowledge building and skills enhancement, equity and fairness, data privacy, and security. In this presentation, the presenter will expand the ongoing discourse on the use of generative AI in English language education. First, he will synthesize what the literature has documented so far by presenting the results of his systematic rapid review of the literature. He will illustrate how the functions and affordances of ChatGPT are supportive of Nation's (2007) Four Strands of English Language Teaching and Learning. Then, he will discuss the common issues that emerged from the systematic review and offer corresponding solutions and actions. Participants in this session will gain valuable insights on how to leverage ChatGPT to enhance their pedagogy and assessment practices while adhering to well-established SLA theories and L2 learning principles.

Abstracts

EFL TEACHERS' AWARENESS AND OPINIONS ON THE USE OF TECHNOLOGY AS 21ST CENTURY LANGUAGE EDUCATION SKILL

Lazgin Barany and Rayan Omar Azeez,
Nawroz University

Awareness and opinions on 21st-century education skills by EFL teachers can lead students to be lifelong learners. EFL teachers need to be aware of and familiar with 21st-century foreign language education skills including technology. This study aimed at investigating the EFL teachers' awareness and opinions on the use of technology in EFL classes. It also attempted to find out whether there were any significant differences among the teachers' awareness and opinions in terms of gender and field of specialization (language and literature). In order to achieve the aims of the study and answer its questions, both quantitative and qualitative research methods have been used. Quantitatively, a 5- 5-point Likert scale questionnaire adopted from a survey made by Ravitz (2014) was used in this study. The participants were 60 teachers: 20 from each department of English language at the Colleges of Basic Education, Education, and Languages at Salahaddin University, Kurdistan of Iraq. Qualitatively, an interview was conducted with 30 teachers: 10 from each of the departments of English at the colleges under study. The results of the study revealed that the participants were aware of the use of technology in EFL classes, and they held positive opinions on its use and application in their classes. Furthermore, the study found that there were no statistically significant differences among the teachers' awareness of the use of technology in terms of gender, but there were on their awareness based on the field of specialization. Moreover, there were statistically significant differences among the teachers' opinions in terms of gender and field of specialization.

Abstracts

GENERATIVE AI TECHNOLOGY AND LANGUAGE LEARNING: LANGUAGE LEARNERS' RESPONSES TO CHATGPT VIDEOS IN SOCIAL MEDIA

Ziqi Chen, *Macao Polytechnic University;*

Wei Wei, *Macao Polytechnic University;*

Xinhua Zhu, *The Hong Kong Polytechnic University*

Generative AI technology has great potential for language learning and teaching (Godwin-Jones, 2022; Jeon, 2022). Since the launch of ChatGPT at the end of 2022, its actual impact on language learners remains unclear. This study aims to address this issue by analyzing commentary data (n = 1088) from the 153 YouTube videos on the use of ChatGPT for language learning between January 2023 and April 2023. With the assistance of the sentiment analysis tool TextBlob, this study is able to detect the emotions in language users' self-reflection of their learning experiences with ChatGPT. Furthermore, a coding system was developed based on the results of both the visualization tool WordCloud and human analysis to identify the patterns in these comments, primarily answering two questions: 1) what languages are being learned with the assistance of ChatGPT, and 2) what kinds of digital affordances have been reported by these language learners. The findings suggest that the most commonly learned languages using ChatGPT include English, Spanish, and Japanese. Moreover, the reported digital affordances of ChatGPT in language learning include: a) opportunities to practice language (e.g., corrections, practice, and partners), b) authentic input of language use (e.g., stories, vocabulary, and examples), c) individualized feedback (e.g., mistakes, grammar rules, and specificity), d) engagement and interaction (e.g., immersion, inspiration, interactivity, and interest), and e) real-time assistance on linguistic knowledge (e.g., always being available, translations and explanations).

Abstracts

AI-POWERED TECHNOLOGIES AND ENGLISH LANGUAGE TEACHER EDUCATION

Betul Kinik, *Inonu University*

There has been an expanding interest in using AI-powered technologies in language learning and teaching recently. While research studies have explored the use and potential of AI in language education, it is crucial to investigate language teacher education to ensure that future language teachers are prepared to incorporate AI into their teaching practices effectively. The current study seeks to reveal the extent to which language teacher education programs equip student teachers with the necessary skills to integrate AI-powered technologies in their future teaching. Following qualitative research design, the courses of English language teaching departments of ten universities from Turkey and the practices of twenty teacher educators working in those universities were examined. Document analysis and semi-structured interviews were used to collect the data. The interviews focused on gathering insights into the preparedness of student teachers to integrate AI technologies into their teaching practices. The findings of the current study will contribute to our understanding of the current state of language teacher education in terms of AI technologies.

Abstracts

HOW DOES ARTIFICIAL INTELLIGENCE AFFECT EFL LEARNERS' COGNITIVE LOAD AND LEARNING ANXIETY? A NEUROSCIENTIFIC PERSPECTIVE

Liwei Hsu

National Kaohsiung University of Hospitality and Tourism

The expeditious development of artificial intelligence (AI) has saliently changed our lifestyle, including the way we learn, particularly in second language (L2) learning. Nevertheless, more empirical insights are still needed to thoroughly understand the role AI can play in L2 learning. This present study designed an experiment to provide empirical evidence to explore this issue in greater detail. Thirty EFL learners (n=30) at the post-secondary level in Taiwan were invited to join this experiment. They were selected based on the requirements of neuroscientific study such as the fact that participants of this present study should not have undergone prior brain surgery, nor should they be on medication that may possibly influence their mental state. None of the participants had the aforementioned concerns and the data collected from them were valid for further analyses. As for their levels of proficiency in English, all these participants had taken TOEIC before with a score between 550–580 as the average TOEIC score of Taiwanese college students was 568 in the year 2021, according to the Education Testing Service (ETS). Moreover, all participants were briefed about the nature of this research before the onset of this study, and they were expected to sign the consent form at the end of this research. They could be withdrawn from the study at any time if they did not feel comfortable continuing, in accordance with research ethics. All participants went through a 15-minute-long conversation with an AI-based chatbot, and their brainwave oscillations were recorded with an electroencephalogram (EEG). Their cognitive load was retrieved and calculated with changes in theta and alpha oscillations while their learning anxiety was obtained with frontal alpha asymmetry (FAA). Results showed that EFL learners' cognitive load was higher and learning anxiety was lower when AI was used.

Abstracts

WHEN IVR MEETS AI: PRACTICES AND CHALLENGES FOR LANGUAGE EDUCATORS

Ilaria Compagnoni, *Ca' Foscari University of Venice*

Increasing developments in educational technology have called for redefinitions of language pedagogies introducing the use of digital tools enabling multi-user interactivity and content creation. Given the increasing use of AI to support collaborative work practices, it is expected that this technology will steadily shape students' language output in task-oriented learning activities. Therefore, there is a need to train language teachers to use AI in task-based language activities granting students' engagement and participation in contextualized linguistic practices. These competencies can be supported by preparing language teachers to use AI with immersive Virtual Reality (iVR) for conducting classroom-based language activities based on hands-on collaborative practices in virtual spaces. However, the current literature lacks inquiries on language activities grounded on AI and iVR-based collaborative group work. Attempting to bridge this gap, this study presents the results of interventions conducted at the University of Arizona on language educators attending teacher training on educational technologies and using the iVR platform Workrooms and ChatGPT. Data collected from observations of teachers' activities and a post-activity questionnaire will provide methodological suggestions for the integration of VR and AI in language learning contexts. This contribution will also give indications of the technological skill sets necessary to teach and learn languages with AI and iVR in order to socially and professionally interact in an increasingly digital world.

Abstracts

WHEN IVR MEETS AI: PRACTICES AND CHALLENGES FOR LANGUAGE EDUCATORS

Wiwik Mardiana, *Universitas Islam Majapahit*

This current research investigates prospective teachers' competencies in building Multimodal Chatbot-Based English Learning Media. This action research involved 15 prospective teachers joining digital training. Using a collaborative project-based learning model, the competencies examined were creativity, creation of contents or materials, and multimodal composition. The results revealed prospective teachers' competencies were enhanced. They became aware of considering multimodal aspects to create comprehensive English learning media. The details of the discussion, implications and recommendations are further discussed.

Abstracts

DIGITAL PROFESSIONAL DEVELOPMENT: CAN IT IMPROVE LANGUAGE TEACHER SENSE OF COMPETENCE, AUTONOMY AND RELATEDNESS?

Alison Porter, *University of Southampton*;
Suzanne Graham, *University of Reading*;
James Turner, *University of Southampton*;
Kate Borthwick, *University of Southampton*;
Travis Ralph-Donaldson, *Niter Ltd.*;
Pengchong Zhang, *University of Reading*

In the UK, as in many other contexts, primary school foreign language learning (PFL) suffers from teacher supply issues related to a lack of language-specific expertise or training for teachers who are generalist teachers, or a limited sense of relatedness to the wider school curriculum for those who are language specialist teachers (Seymour, 2018). These issues pose threats to PFL teachers' motivation, especially to their sense of autonomy, relatedness and competence (Deci & Ryan, 2000). Digital Professional Development (DPD) has been found to promote language teachers' sense of competence, autonomy and relatedness (Haukås et al., 2022). Yet there is little evidence on exactly how DPD has this effect or about the relationship between all these aspects of teacher motivation within a DPD context.

This presentation reports on a research project that aimed to address these research gaps by investigating the role of technology within a programme of DPD. Thirty-five PFL teachers participated in approximately 12 hours of DPD that focused on the development of learner PFL literacy, including through using digital storytelling and an AI-powered pronunciation training app. Analysing data collected from questionnaires completed before and after the DPD, alongside interactional teacher data from the DPD digital platform, we present findings that consider how far and in what ways the DPD improved teacher sense of competence, autonomy, and relatedness, and hence motivation to teach PFL.

Abstracts

ROUNDTABLE-EFFECTIVENESS OF A TECHNOLOGY ENHANCED PROJECT-BASED LANGUAGE LEARNING (TEPBLL) APPROACH IN L2 PRAGMATICS ENHANCEMENT

Ebtehal Asiri, Gulbahar Beckett, and Ali Garib

Iowa State University

Competence in target language and sociocultural community pragmatics is crucial for non-native speakers of any language. Second and foreign language acquisition research has explored various aspects of pragmatics, including identifying and developing effective strategies to enhance L2 learners' pragmatic competence for decades. Yet, research consistently shows that pragmatics remains problematic for ESL learners, even after exposure to a target language context (Nguyen & Pham, 2021; Taguchi, 2019). As a result, researchers (e.g., Ren, 2022; Attardo & Pickering, 2021; Qari, 2021) have emphasized the need for pedagogical intervention in teaching L2 pragmatics in L2 classrooms with alternative authentic approaches.

This roundtable focuses on a proposal to bridge the above issue with an authentic teaching approach called Technology Enhanced Project-Based Language Learning (TEPBLL). Participants will be invited to share ideas on and discuss the role of AI tools within the TEPBLL approach to teaching, learning, and researching pragmatics. We will ask and address questions such as: if and how the applications of the current generation of AI tools in TEPBLL pragmatics teaching and learning can make a ground-breaking contribution to advancing our knowledge and practice on pragmatics. The roundtable will also explore how AI tools in project-based pragmatics teaching and learning may be researched to help expand pragmatics research.

Abstracts

PRESERVICE TEACHERS' INSIGHTS ON AI IN LANGUAGE EDUCATION

Alba Paz-López and Boris Vazquez-Calvo
University of Malaga

Recent developments in AI generative software have taken many researchers and educators off guard. In a recent tweet thread (1), some educators expressed varied opposing views regarding AI in education. First, there are those who understand AI as a model that requires an overhaul of the education system. Then, there are those who advocate for a revival of rote memorization and exams with no access to any technology whatsoever. Such heated debates inspired us to explore how younger generations of preservice language teachers perceive the irruption of AI in and for language education. As part of a larger study exploring how technology shapes the identities and practices of future and early-career language teachers, we incorporated questions about AI during in-depth interviews. We present preliminary findings from three interviews with three master's students from Spain who specialize in teaching EFL. Using CAQDAS software, the thematic analysis of the interviews followed an inductive pattern to reconstruct an incipient narrative around AI. Findings underscore AI's transformative potential in education. One participant likened AI to a 'fire' in schools, signaling both renovation and upheaval. Participants saw AI as a tool to enhance teaching practices in lesson planning, vocabulary acquisition, brainstorming, and transcription. Their responses suggest the acceptance of AI to streamline design tasks. However, they expressed hesitation and a lack of training regarding the direct use of AI in language instruction. For instance, one participant was adamant that AI complexifies the nature of plagiarism and plagiarism detection, challenging conventional notions of authorship. While the participants recognized the potential of AI in language education, they emphasized the irreplaceable value of face-to-face interactions. This underscores the importance of balancing technology integration with other pedagogical methods. The participants also highlighted the need for training in the effective use of AI in classroom settings.

This roundtable focuses on a proposal to bridge the above issue with an authentic teaching approach called Technology Enhanced Project-Based Language Learning (TEPBLL). Participants will be invited to share ideas on and discuss the role of AI tools within the TEPBLL approach to teaching, learning, and researching pragmatics. We will ask and address questions such as: if and how the applications of the current generation of AI tools in TEPBLL pragmatics teaching and learning can make a ground-breaking contribution to advancing our knowledge and practice on pragmatics. The roundtable will also explore how AI tools in project-based pragmatics teaching and learning may be researched to help expand pragmatics research.

Abstracts

ROUNDTABLE-PBLL @ ISU: A ROUNDTABLE ON PBLL PROJECTS, WORKS IN PROGRESS, AND AI CONNECTIONS

Jeanne Beck, *Iowa State University*;

Ali Garib, *Iowa State University*;

Junghun Yang, *Iowa State University*;

Hwee Jean (Cindy) Lim, *Iowa State University*

Project-Based Language Learning (PBLL) is a widely researched area of applied linguistics worldwide, often integrating technology as a way to assist language learners across a variety of contexts. In this roundtable, five Iowa State University Ph.D. students in the Applied Linguistics and Technology (ALT) Program will share insights from their PBLL research projects and works in progress, centered around some of the challenging elements of effectively implementing PBLL. These include supporting teachers through training, enhancing learning by connecting form and function, incorporating AI and other tech tools, and enhancing assessment of PBLL project implementation. Dr. Gulbahar Beckett, who will be serving as the round table's discussant, will guide the panelists in connecting the wider issues related to the field as it relates to their projects and work. Time will be allotted for questions for the whole group or individual speakers.

Abstracts

PRIMARY LEVEL LEARNERS' PERCEPTIONS OF BILINGUAL VOICEBOTS IN INDIAN ESL CLASSROOMS: A PILOT STUDY

Harshitha H, *The English and Foreign Languages University*

Young ESL learners in India, especially in rural contexts, are not provided sufficient opportunities for speaking in English both within and outside the language classroom. Such learners often experience language anxiety when they attempt to speak in the target language. This study explores the learners' perceptions of a tailor-made voice robot that can recognize the learners' speech in English and respond appropriately to provide additional opportunities for target language use. Unlike earlier studies with similar approaches, this study uses an interactive story robot that can provide support to learners in their own language (Hindi). 5 ESL learners at a primary school in India were given the opportunity to interact with the prototype. The study employs a qualitative research design where semi-structured interviews and learner perception questionnaires are used to collect data on learner perception post-interaction with the robot. The data was analysed using Davis's Technology Acceptance Model (1989) to see if the perceived usefulness and ease of use of the robot would translate to an intention to utilise them for ESL learning in the future. The pilot study shows positive results in terms of learners' acceptance of voice robots for language learning as well as the incorporation of their own language use in such interactions.

Abstracts

ROUNDTABLE-INTEGRATING AI TOOLS INTO INSTRUCTED SLA

Robert Godwin-Jones, *Virginia Commonwealth University*;
Jim Ranalli, *Iowa State University*;
Errol M. O'Neill, *University of Memphis*

Artificial intelligence is not new, but recent approaches to AI development based on machine learning have enabled the creation of “generative AI” tools capable of creating texts, images, or videos in response to a brief user request/prompt. The public availability of ChatGPT and other AI tools is having a transformative effect on our lives in a variety of domains, including education. Many language learners have for some time been avid users of AI-enabled tools such as Google Translate, “smart” text editors like Grammarly, or voice assistants (Siri, Alexa). While some uses of AI products in SLA have generally been seen positively (such as automated writing evaluation), the use of others has been controversial, with some language teachers, for example, attempting to ban student use of machine translation. This roundtable will address issues surrounding the integration of AI tools into instructed SLA, focusing on machine translation, chatbots, and automated written corrective feedback. The panelists, drawing from both experimental studies and theory-based analysis, will address practical uses of AI tools in instructed SLA and implications for the future of second language learning in an AI world.

Abstracts

USING AN AI PROGRAM TO ASSESS ACCURACY IN SECOND LANGUAGE WRITING RESEARCH

Charlene Polio, *Michigan State University*;

Adam Pfau, *Michigan State University*;

Yiran Xu, *University of California Merced*

Researchers in second language (L2) learning and teaching have been concerned with measures of syntactic and lexical complexity, accuracy, and fluency for over two decades. While, over the past ten years, automated tools have become available for assessing both lexical complexity and syntactic complexity, accuracy has been the least studied and the least assessed in studies of L2 writing, presumably because of reliability and feasibility issues. We argue and empirically demonstrate here that AI programs such as ChatGPT are a viable option for researchers who want to assess accuracy in L2 writers' written texts.

To test ChatGPT, we created a corpus of 100 essays with five proficiency levels represented. The essays were manually coded for precision and recall with regard to ChatGPT's identification of errors. We found that for our global measure, errors per word, ChatGPT correlated highly with human coders (i.e., .94), but there was some variation among proficiency levels. While ChatGPT rarely misidentified a structure as an error, ChatGPT identified far fewer errors than human coders in terms of absolute number of errors (or errors per word). In addition, although its classification of errors was inconsistent in terms of labels, it was mostly correct. We highlight the issues and challenges of using ChatGPT to assess L2 writing accuracy for research purposes and offer suggestions such as instituting manual checks, dealing with less proficient writers, and creating appropriate prompts for ChatGPT. Lastly, we discuss a study currently underway that involves re-examining data from a previously published study that included manually coded data. The goal is to determine whether or not the results are replicated using ChatGPT.

Abstracts

ESL GRADUATE STUDENTS' USE OF CHATGPT FOR TEXT REVISION: INSIGHTS INTO BEHAVIOR, COGNITION, AND EMOTION

Svetlana Koltovskaia, *Northeastern State University*;
Hooman Saeli, *the University of Tennessee, Knoxville*;
Payam Rahmati, *Oklahoma State University*

Extensive research spanning over a decade has investigated the utilization of AI-powered tools, like Grammarly, for text revision in L2 writing classrooms (Ranalli & Yamashita, 2022; Guo et al., 2021). While the research suggests that such tools hold potential for text revision, their feedback is often limited to grammar and mechanics (Koltovskaia, 2020). ChatGPT, a state-of-the-art AI-powered tool, offers suggestions that go beyond the confines of grammar and can significantly benefit learners by aiding them in directing their attention toward global aspects of writing. This consequently can help students develop independent self-editing skills, which is the ultimate goal of writing instruction (Ferris, 2014). To gain a comprehensive understanding of ChatGPT's benefits, however, studies should focus on students' use of this tool during the revision process, as this type of investigation can reveal the degree to which learners acquire metacognitive skills necessary to notice, evaluate, and improve their writing (Stevenson & Phakiti, 2014).

The present study employs a multidimensional framework of learner engagement to explore how six Iranian graduate students studying at a US southcentral university use, process, and react to ChatGPT's suggestions when revising their research papers. Following previous literature (Han & Hyland, 2015), the concept of learner engagement is operationalized through three interrelated dimensions, including behavioral engagement (reflected in revision operations and strategies), cognitive engagement (represented by the use of metacognitive and cognitive operations), and affective engagement (manifested in immediate emotional reactions and attitudinal responses to ChatGPT's suggestions). Behavioral engagement is investigated through analysis of screen recordings of students' usage of ChatGPT when revising their texts. The cognitive and affective dimensions of engagement are assessed through an examination of students' comments during the recall of the aforementioned screen recordings and semi-structured interviews. The findings reveal that all six participants utilized ChatGPT behaviorally for grammar, clarity, and word choice, relying on prompts from the training session. Cognitively, participants generally understood the feedback but questioned its accuracy when necessary. Affective engagement indicates that all participants expressed high satisfaction with ChatGPT, particularly favoring its use for paraphrasing to enhance the professionalism of their writing. The study findings provide implications for the meaningful use of ChatGPT in L2 writing classrooms.

Abstracts

GENERATIVE AI, RESEARCH WRITING, AND ETHICS IN APPLIED LINGUISTICS: HOW DO WE IDENTIFY HUMAN-PRODUCED WRITING

J. Elliott Casal, *University of Memphis*;
Matt Kessler, *University of South Florida*

Applied linguists are celebrating methodological strides made in recent decades (Gass et al., 2021; Plonsky, 2014) and calling for further commitments to elevate methodological rigor and critical interrogation of issues of research ethics (e.g., De Costa et al., 2019; Kubonyiova, 2008). At the same time, and like in other fields, applied linguists are grappling with the emergence of freely available Large-Language Model (LLM)-powered AI chatbots (e.g., OpenAI's ChatGPT, Google's Bard), which have the power to quickly generate text that is designed to resemble human-produced language. There has been considerable discussion in the field regarding the extent to which these technologies can and should impact language teaching and scholarship, with research starting to emerge on the issue. However, most studies have explored such tools' general capabilities and applications for language teaching purposes, with less attention directed toward the methodological and research ethics issues associated with these emerging technologies.

The current study examines issues pertaining to human judgements, accuracy, and research ethics. Specifically, it investigates 1) the extent to which linguists/reviewers from top journals can distinguish AI- from human-generated writing, 2) what the basis of reviewers' decisions are, and 3) the extent to which editors of top Applied Linguistics journals believe AI tools are ethical for research purposes. In the study, reviewers (N= 72) completed a judgement task involving AI- and human-generated research abstracts, and several reviewers participated in follow-up interviews to explain their rationales. Similarly, editors (N= 27) completed a survey and semi-structured interviews to discuss their beliefs. Findings suggest that despite employing multiple rationales to judge texts, reviewers were largely unsuccessful in identifying AI versus human writing, with an overall positive identification rate of only 38.9%. Additionally, many editors believed there are ethical uses of AI tools for facilitating research processes, yet some disagreed.

Abstracts

EXPLORING BENEFITS AND CHALLENGES OF QUIZLET'S Q-CHAT IN INDONESIAN LANGUAGE INSTRUCTION FOR MILITARY PERSONNEL

Gatot Prasetyo, *University of Montana*;
Sisilia Kusumaningsih, *University of Montana*

This proposal presents a research study investigating the use of Quizlet's ChatGPT-powered Q-chat in Indonesian language instruction for Army Special Forces. The study aims to examine the perceptions of teachers regarding this innovative technology. Surveys and interviews will be employed to explore the potential advantages and obstacles associated with integrating Q-chat into the Indonesian language curriculum for this unique learner population.

The study aims to explore the benefits of utilizing Q-chat as a supplemental tool in Indonesian language instruction for Army Special Forces. Surveys will gather quantitative data on the perceived effectiveness of Q-chat in improving language skills, enhancing vocabulary retention, and fostering communicative competence. Qualitative interviews will provide valuable insights from teachers about their experiences and suggestions regarding the integration of Q-chat into the curriculum.

Additionally, the study seeks to identify challenges associated with the implementation of Q-chat in Indonesian language classes for Army Special Forces. Surveys will collect data on technical difficulties, time constraints, and potential distractions. Interviews will allow participants to express their perspectives on adapting to the Q-chat platform, managing its interactive nature, and addressing concerns about privacy and security. By examining both benefits and challenges, this research aims to inform language educators and curriculum designers about integrating Q-chat into Indonesian language instruction for Army Special Forces. The findings will contribute to evidence-based practices and recommendations for utilizing Q-chat in this specialized context, enhancing language learning experiences and outcomes. This study also paves the way for further research and implementation of AI-enhanced technologies in language education for military personnel.

Abstracts

EXPLORING THE IMPACT OF CHATGPT ON ENGLISH LANGUAGE LEARNERS' EMAIL WRITING PROCESS

**Widya Kusumaningrum, Kate Challis,
Hwee Jean (Cindy) Lim, Jeanne Beck, Carol Chapelle,**
Iowa State University

ChatGPT, a new artificial intelligence chatbot built on a large language model, has grabbed the attention of academics around the globe due to its ability to synthesize prompts and rapidly generate responses. Both teachers and students alike are finding ways for the tool to help learning and increase efficiency in carrying out common tasks in English. The purpose of this study is to examine how English language learners (ELLs) use ChatGPT to write professional and informal emails. Our research examines how ELLs interact with ChatGPT during the email drafting process in terms of Baaijen & Galbraith's (2018) framework which evaluates: 1) time spent on each task, 2) production cycles (i.e., the number of prompts used for each task), 3) conceptual similarity, 4) sentence similarity, and 5) text modification. Establishing how ELLs currently interact with ChatGPT lays the groundwork for future research on what teachers need to do to optimize and empower ELLs to use this technology. This research also provides valuable insight into the text register of ELL email writing, a skill that is crucial for academic success and is currently understudied.

Abstracts

A FIVE-ECOSYSTEMS APPROACH FOR APPLYING AI TO THE LANGUAGE CLASSROOM

Thor Sawin, *Middlebury Institute of International Studies*

When applying generative AI to the language classroom, teachers and curriculum designers need far more than technical proficiency with AI tools. In van Lier's ecological approach (2004, 2011), technological tools must be evaluated within the rich ecosystem of learner processes, social expectations, and other tools in the environment, keeping pedagogy first, curriculum second, and technology third. Seemingly artificial, AI nevertheless can fit within ecological language learning, offering learners play, agency over topics and content, and abductive learning (van Lier, 2011). Yet for these affordances to outweigh stakeholders' fears (Chen, 2023) – of isolated learners, inauthentic language, and unoriginal thought – designers must carefully attend to the complex ecosystem surrounding the uses of AI.

This practice-based presentation explores five ecosystems which teachers and curriculum designers must consider, supporting each with research and illustrating each with practical examples of AI applied to classroom learning. The five ecosystems are: 1) the pedagogical – classroom processes AI performs; 2) linguistic – those aspects of language learnable via AI; 3) acquisitional – elements of the acquisition process (e.g. input, feedback, vocabulary provision); 4) technological – the other tools AI need interface with (e.g. screenshotting with drawing tools, social annotation, search engines) to be maximally effective, and 5) sociocultural – stakeholders' (such as parents and administrators) ideologies about language learning (e.g. privacy, studenthood, standardness, creativity). Designing within these ecosystems helps ensure that AI facilitates, rather than substitutes, classroom language acquisition.

Abstracts

CHATGPT AS AN AI-POWERED WRITING ASSISTANT: HOW RELIABLE IS IT IN EDITING L2 WRITING?

Shuyuan Tu, *Georgia State University*

The emergence of Large Language Models (LLMs), such as ChatGPT, has demonstrated remarkable abilities in generating human-like language and providing immediate and customized feedback, which makes it promising to serve as a valuable and accessible writing assistant (Pavlik, 2023; Mizumoto & Eguchi, 2023; Kuhail et al., 2023). Nevertheless, relatively few studies further extend this topic by investigating ChatGPT's feedback on writing products and its editing moves. Therefore, this study investigates ChatGPT's performance by analyzing the extent to which ChatGPT can contribute to lexical and syntactic complexity in editing learner essays. This study also explores the reliability of ChatGPT's contribution to linguistic complexity by comparing its editing moves with human editors. The L2 learner essays in this study were extracted from the International Corpus Network of Asian Learners of English (ICNALE) (Ishikawa, 2018). A total of 80 learner essays and 80 of their human-edited versions were selected for this study. Similarly, ChatGPT was prompted using the same rubric and required to generate the fully edited versions of the 80 original learner essays using ChatGPT API. The lexical and syntactic complexity in the original learner essays, human-edited essays, and ChatGPT-edited essays were measured using the Lexical Complexity Analyzer (Lu, 2012; Ai & Lu, 2010) and L2 Syntactic Complexity Analyzer (Lu, 2017; Ai & Lu, 2013). Both quantitative and qualitative methods were employed in this study. Tentative findings indicate that ChatGPT suggests greater lexical variation and coordinating conjunctions when editing learner essays. Moreover, it also presents divergent editing preferences with human editors. This study could provide insight into the reliability of ChatGPT in editing essays in L2 writing by investigating its editing moves. It could also raise awareness of the potential of the AI-powered writing assistant in facilitating L2 writing from a linguistic complexity perspective.

Abstracts

EXPLORING OPEN AI IN LANGUAGE TEACHER PROFESSIONAL DEVELOPMENT

Carol Chapelle, Shireen Baghestani, Jeanne Beck, Sebnem Kurt, Widya Kusumaningrum, Hwee Jean (Cindy) Lim, Shuhui Yin
Iowa State University

The language capacities of generative open AI invite language teachers to consider its implications for their teaching and compel language teacher educators to introduce its capabilities and limits. This presentation reports our research exploring professional development materials on using ChatGPT to serve the basic needs of international English language teachers, such as creating texts meeting genre, content, and level specifications, assisting students in brainstorming during prewriting, offering corrective feedback, and engaging in real-time language interaction. We created demonstrations illustrating these functionalities for our global online course, "Using Educational Technology in the English Language Classroom." Our research represents the initial phase of an educational design research process with dual goals of increasing theoretical knowledge and improving practice (McKenney & Reeves, 2019). Our team of six English language teachers with experience in China, Japan, Indonesia, Malaysia, Turkey, and the United Arab Emirates independently watched three AI demonstrations, one each week, over a three-week period. They then responded to questions about their probable actions as teachers based on each video's demonstrations, instructions, and ideas. Before watching, the teachers responded to questions about their prior knowledge of and interest in generative AI tools. At the end of the session, they reflected on the clarity of the advice in the module, their experience in prompting ChatGPT to obtain useful results, their learning from the experience, and their advice for their colleagues about ChatGPT. The responses from the teachers yield data that we interpret for 1) building an understanding of the role of generative AI in technological pedagogical content knowledge (TPACK) that forms the basis of the course and 2) identifying areas for improvement in the demonstrations.

Abstracts

CHATGPT VS. HUMAN IELTS TUTOR: A COMPARATIVE STUDY OF IELTS TASK 2 WRITING SAMPLE

Andrias Tri Susanto, *Iowa State University*

The impact of artificial intelligence (AI)-based applications has garnered significant attention in the realm of English language instruction. One such example is ChatGPT, which can generate a sample for the argumentative writing Task 2 section of the International English Language Testing System (IELTS) within 19 seconds, comprising 361 words (at a rate of 0.053 seconds per word) while demonstrating a high level of lexical resource and grammatical accuracy and complexity. Meanwhile, a human IELTS tutor can take an average of 28 minutes to produce a 408-word essay (at a rate of 4.12 seconds per word) with some minor errors, resulting in an astonishing 77-fold difference in efficiency.

This study aimed to explore the perspectives of 14 participants regarding this phenomenon. The study utilized three different writing samples for IELTS Writing Task 2, collected from (1) generating the example using ChatGPT 3.5 as of the September 2021 knowledge cut-off, (2) the official IELTS website, and (3) an IELTS tutor with ten years of teaching experience. The participants were requested to rank six samples, consisting of two sets of three samples from each session of data collection for IELTS Writing Task 2. Subsequently, three participants were chosen randomly and asked to explain the rationale behind their preferred rankings.

The research findings indicated that most participants ranked the writing sample created by ChatGPT in the first place, followed by the sample generated by the IELTS tutor and the one from the official IELTS website in the second and third ranks, respectively. The outcome of the semi-structured interviews revealed that the participants preferred a more rigid structure for writing samples with a somewhat limited degree of creativity. Nevertheless, the role of the IELTS writing tutor was deemed essential in providing guidance and identifying the participants' writing errors.

Abstracts

EFL SCHOOL TEACHER'S AI ASSESSMENT LITERACY: AN EXPLORATORY STUDY

Hao Xu, *Beijing Foreign Studies University*

In this presentation, I shall report on an exploratory study that investigated six high school EFL teachers' AI assessment literacy as manifested through their use of ERNIE Bot, a Chinese AI-powered chatbot similar to ChatGPT, as they conducted assessments on and gave feedback to the students' written works. Data were collected via individual interviews and digital protocols, which showed how the teachers used the ERNIE Bot. As teachers' assessment knowledge underpins their assessment literacy, a framework for discerning teachers' knowledge that supported their assessment and feedback practices was employed to enable the categorization of the manifestations of teachers' AI assessment literacy into a three-tier knowledge construct, i.e., technicality, practicality, and conceptuality. Findings show that the layers of practicality and conceptuality, which encompass teacher's practical and conceptual knowledge of assessment and feedback, do not seem to be added with new components due to exposure to AI technologies. The layer of technicality, nonetheless, seems to have developed as a body of knowledge that could operate independently, i.e., it seems to have become less attached to practical and conceptual knowledge. Besides, technicality might not necessarily promote the development of practical and conceptual knowledge, although the lack of technical knowledge obviously posed a hindrance.

This re-conceptualization of teacher's AI assessment literacy may provide an alternative to the more popular view that AI assessment literacy is the enactment of assessment literacy in AI-powered digital environments. This re-conceptualization thus calls for more attention to the indispensability of technical knowledge beyond being epiphenomenal, which co-exists with AI-powered digitality as a permeating culture that characterizes teachers' assessment and feedback practices.