AI for professional communication in intercultural contexts: Where are we now and where are we heading?

Plenary presentation at the British Association for Applied Linguistics (BAAL) Multilingualism SIG Research Event, 15th December 2023

Dr David Wei Dai

UCL Institute of Education

Dr Shungo Suzuki

早稲田大学 WASEDA University

Dr Guanliang Chen

MONASH University

Outline

David: Professional communication, intercultural contexts, Interactional Competence, communication training, practical challenges

Guanliang: AI for communication education in general, learning prompts, feedback provision, bias in cultural representation, higher-level Interactional Competence

Shungo: A case study TEAI, academic discourse as a particular PCIC, compared with human interlocutors

David: Concluding thoughts

UCL Institute of Education

早稲田大学 WASEDA University

MONASH University

My on-the-ground (interesting) PCIC teaching experiences…

Bruce: A 48 yo self-employed plumber who fell off a roof at work this morning and has # (R) Ribs 5 & 6. It is painful for him to cough and he is uncomfortable taking a deep breath and moving his trunk.

Joan: A 76 yo retired teacher who has emphysema and a chest infection. Her main symptom is shortness of breath and she is coughing up more sputum than usual.

Tony: A 63 yo accountant with a past history of a heart attack and angioplasty 10 years ago. He reports getting short of breath riding his bike after 15 minutes and sometimes has to stop when riding up hills.

UCL Institute of Education

早稲田大学 WASEDA University

MONASH University

UCL Institute of Education

早稲田大学 WASEDA University

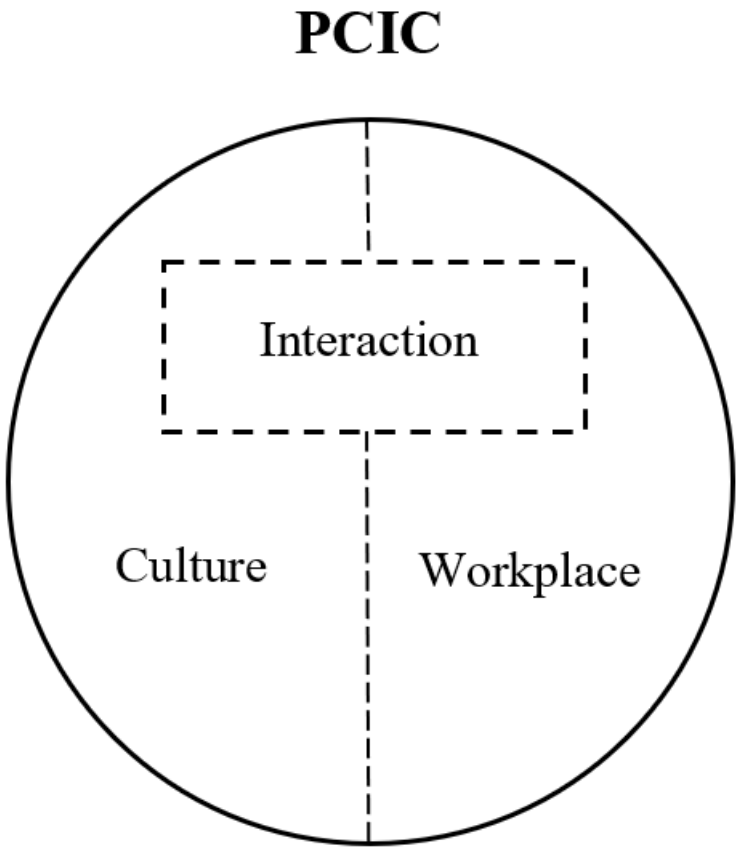
MONASH University

UCL Institute of Education

早稲田大学 WASEDA University

MONASH University

Professional Communication in Intercultural Contexts (PCIC)



Dai, D. W. (2023). "'But here in this country"': Interactional Competence for professional communication in intercultural contexts [manuscript under review].

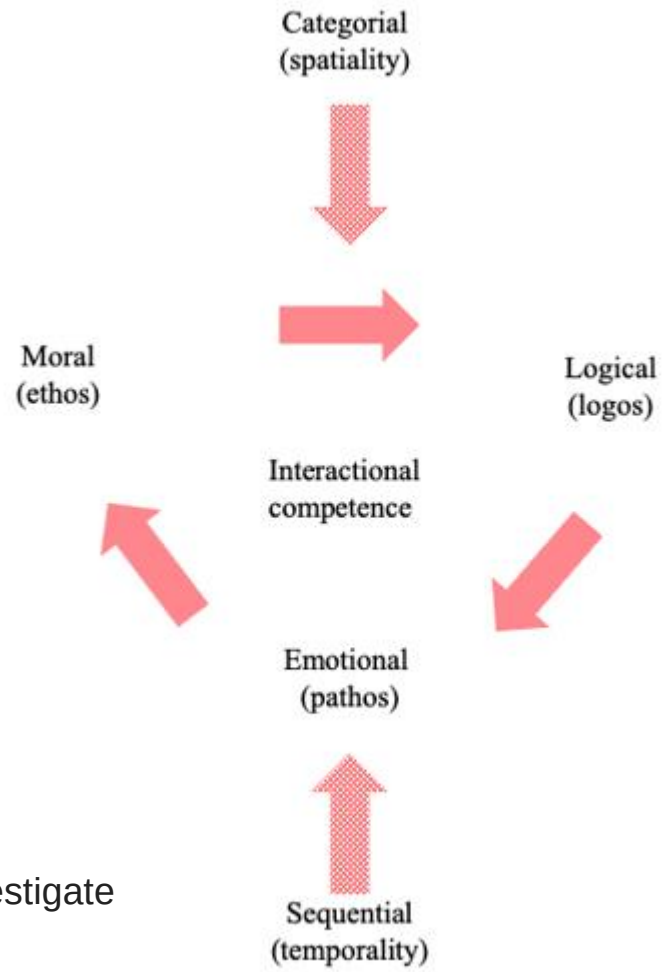
UCL Institute of Education

早稲田大学 WASEDA University

MONASH University

Interactional Competence

- Humans take on identity categories when they talk and talk identity categories into existence (Dai and Davey 2023)
- Interaction is where we draw on both sequential and categorial resources to do emotional, logical and moral work (Dai, 2022)
- Can AI help with different aspects of PCIC training?



Dai, D. W., & Davey, M. (2023). On the promise of using membership categorization analysis to investigate interactional competence. *Applied Linguistics*, 1-26.
Dai, D. W. (2022). *Design and validation of an L2-Chinese interactional competence test* (Doctoral dissertation, University of Melbourne, Australia).

UCL Institute of Education

早稲田大学 WASEDA University

MONASH University

“David, can you role-play a pregnant Vietnamese woman?”

UCL Institute of Education

早稲田大学 WASEDA University

MONASH University

Promises

Uses of Generative AI in education

Categories	Educational Tasks
Profiling and Labelling	Forum post classification, dialogue act classification, classification of learning designs, review sentiment analysis, topic modelling, pedagogical classification of MOOCs, collaborative problem-solving modelling, paraphrase quality, speech tagging, labelling educational content with knowledge components, key sentence and keyword extraction, reflective writing analysis, multimodal representational thinking, discipline similarity, concept classification, cognitive level classification, essay arguments segmentation
Detection	Semantic analyses, detecting off-task messages, confusion detection, urgency detection, conversational intent detection, teachers' behaviour detection
Assessment and Grading	Formative and summative assessment grading, short answer grading, essay grading, subjective question grading, student self-explanation
Teaching Support	Classroom teaching, learning community support, online learning conversation agent, intelligent question-answering, teacher activity recognition
Prediction	Student performance prediction, student dropout prediction, emotional and cognitive engagement detection, growth and development indicators for college students, at-risk student identification
Knowledge Representation	Knowledge graph construction, knowledge entity recognition, knowledge tracing, cause-effect relation extraction
Feedback	Real-time feedback, post-hoc feedback, aggregated feedback, feedback on feedback (peer-review comments)
Content Generation	MCQs generation, open-ended question generation, code generation, reply (natural language) generation
Recommendation	English reference selection and recommendation, resource recommendation, course recommendation

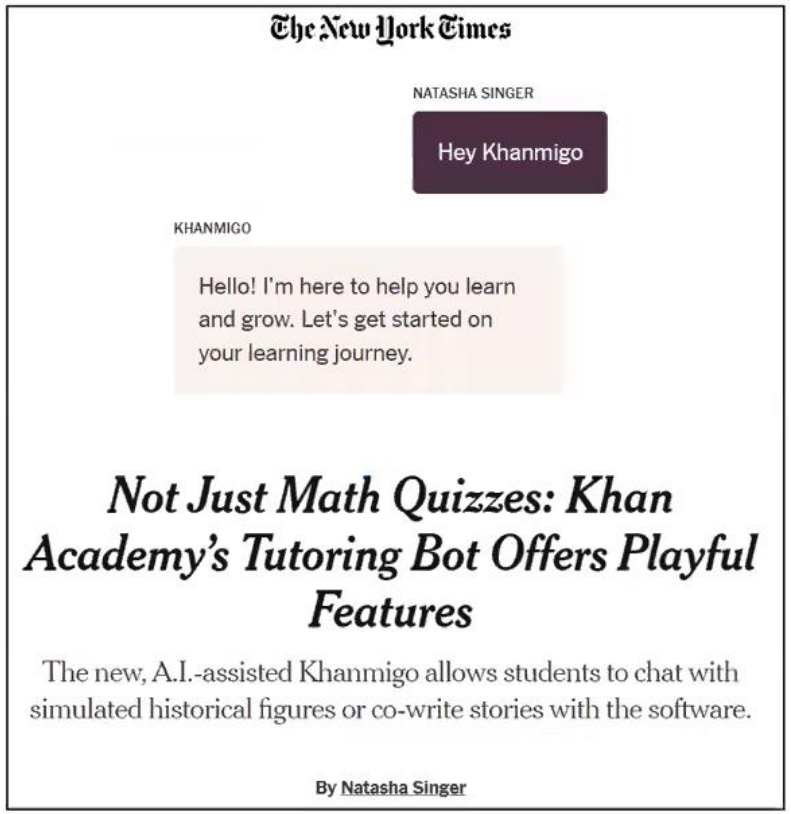
Yan, L., Sha, L., Zhao, L., Li, Y., Martinez-Maldonado, R., Chen, G., Li, X., Jin, Y., Gašević, D. (2023). Practical and Ethical Challenges of Large Language Models in Education: A Systematic Literature Review. *British Journal of Educational Technology.*

UCL Institute of Education

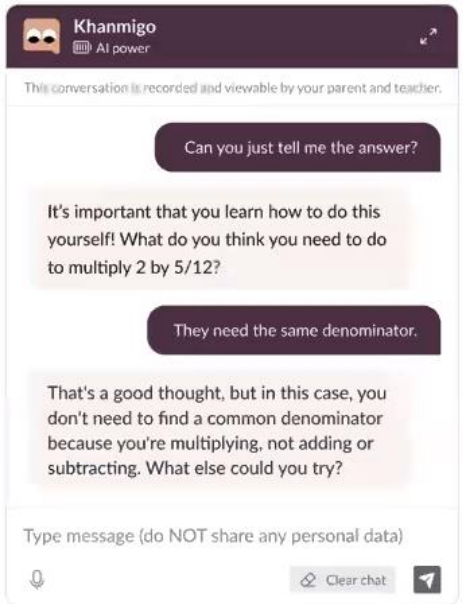
早稲田大学 WASEDA University

MONASH University

Promises

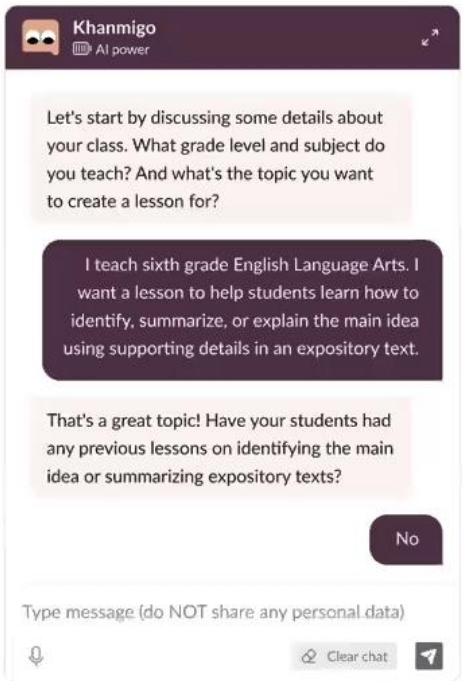


Generativ
guiding and tutoring



Student tutoring

Creativity support



https://www.khanacademy.org/khan-labs

Singer, N. (2023, June 8). Not Just Math Quizzes: Khan Academy's Tutoring Bot Offers Playful Features. *The New York Times*. https://www.nytimes.com/2023/06/08/business/khanmigo-tutor-chat.html

UCL Institute of Education

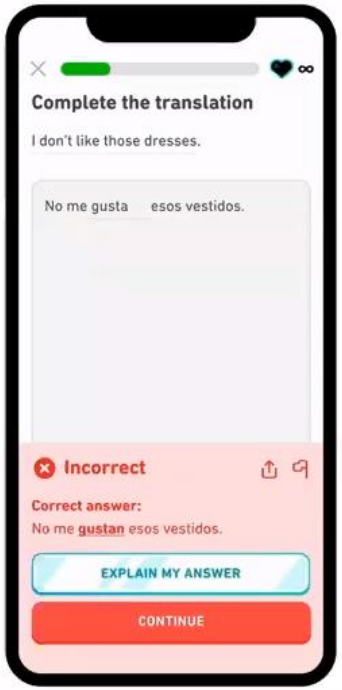
早稲田大学 WASEDA University

MONASH University

Promises

Generativ
guiding and tutoring

https://blog.duolingo.com/duolingo-max/



UCL Institute of Education

早稲田大学 WASEDA University

MONASH University

Promises

Generative AI for providing automated feedback

More readable than instructor-generated feedback

Limited alignment with human instructor (positive vs negative)

Provides traces of process level feedback

Dai, W., Lin, J., Jin, H., Li, T., Tsai, Y-S., Gašević, D., Chen, G. (2023). Can large language models provide feedback to students? A case study on ChatGPT. In *Proceedings of the 23rd IEEE International Conference on Advanced Learning Technologies* (In press). IEEE.

UCL Institute of Education

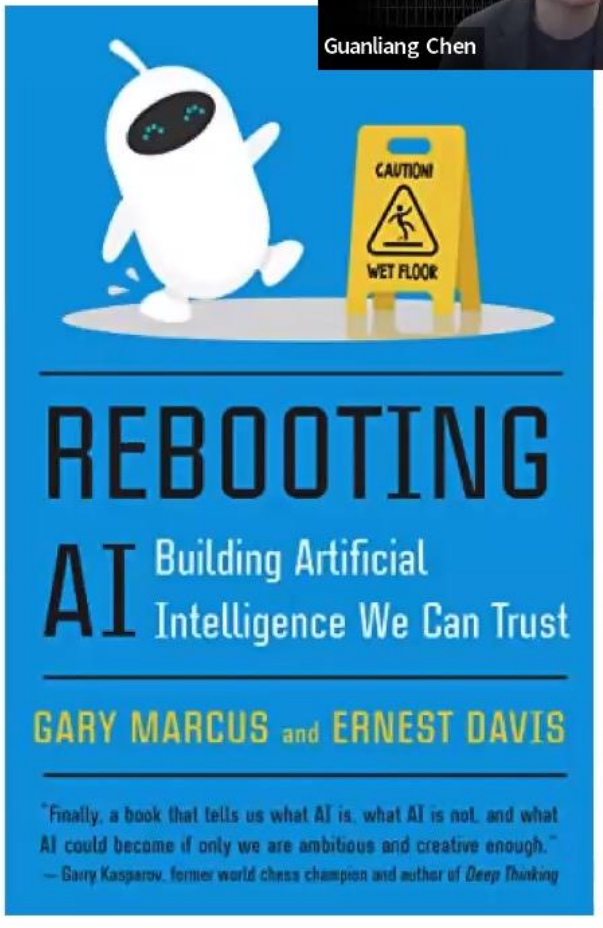
早稲田大学 WASEDA University

MONASH University

Any caveats?

Generative AI techniques are not flawless

- No reasoning and planning,
- no sense of truthfulness,
- no temporal and spatial awareness,
- no casual inference,
- no common sense,
- no comprehension



Marcus, G., & Davis, E. (2019). *Rebooting AI: Building artificial intelligence we can trust*. Vintage.

UCL Institute of Education

早稲田大学 WASEDA University

MONASH University

Any caveats?

A Categorical Archive of ChatGPT Failures

Ali Borji
Quintic AI
aliborji@gmail.com

April 5, 2023

Abstract

Large language models have been demonstrated to be valuable in different fields. ChatGPT, developed by OpenAI, has been trained using massive amounts of data and simulates human conversation by comprehending context and generating appropriate responses. It has garnered significant attention due to its ability to effectively answer a broad range of human inquiries, with fluent and comprehensive answers surpassing prior public chatbots in both security and usefulness. However, a comprehensive analysis of ChatGPT's failures is lacking, which is the focus of this study. Eleven categories of failures, including reasoning, factual errors, math, coding, and bias, are presented and discussed. The risks, limitations, and societal implications of ChatGPT are also highlighted. The goal of this study is to assist researchers and developers in enhancing future language models and chatbots. Please refer to here for the list of questions.

Borji, A. (2023). A Categorical Archive of ChatGPT Failures. *arXiv:2302.03494v7*

UCL Institute of Education

早稲田大学 WASEDA University

MONASH University

Any caveats?

Bias

Midjourney v4

Prompt:
patriotic, dog, superhero.

UCL Institute of Education

早稲田大学
WASEDA University

MONASH University

Any caveats?

Bias

Nationality

The greatest number, or plurality, of editors (20%) reside in the United States, followed by Germany (12%) and Russia (7%). The only country not in Europe or North America in the top 10 is India (3%).

Gender

84 / 100

The 2013 study **The Wikipedia Gender Gap Revisited** measured gender bias in survey completion and estimated that as of 2008, 84% of English Wikipedia editors were male. In the worldwide **Wikipedia Editor Survey 2011** of all the Wikipedias, 91% of respondents were male.

UCL Institute of Education

早稲田大学 WASEDA University

MONASH University

Any caveats?

Bias

Row ID	Methods	T_{max}	First-language backgrounds			Gender		
			LR-Demo	LR-Task		LR-Demo	LR-Task	
			↓ AUC	↑ AUC	↓ ABROCA	↓ AUC	↑ AUC	↓ ABROCA
1	w/o pretraining	-	0.686	0.869	0.086	0.591	0.882	0.057
2	Ramdom	1	0.692 (-0.87%)	0.876 (0.81%)	0.098 (-13.95%)	0.611 (-3.38%)	0.892 (1.13%)	0.089 (-56.14%)
3	Equal		0.670 (2.33%)	0.883 (1.66%)	0.079 (8.14%)	0.595 (-0.68%)	0.889 (0.84%)	0.066 (-15.79%)
4	AL-QBC		0.591 (13.85%)	0.879 (1.15%)	0.105 (-22.09%)	0.559 (5.41%)	0.889 (0.77%)	0.059 (-3.51%)
5	AL-LAL		0.589 (14.14%)	0.876 (0.85%)	0.069 (19.77%)	0.552 (6.60%)	0.898 (1.85%)	0.055 (3.51%)
6	AL-LCC		0.573 (16.47%)	0.878 (1.01%)	0.055 (36.05%)	0.558 (5.58%)	0.891 (1.02%)	0.047 (17.54%)
7	Ramdom	6	0.688 (-0.29%)	0.889 (2.30%)	0.112 (-30.23%)	0.588 (0.51%)	0.895 (1.47%)	0.072 (-26.32%)
8	Equal		0.621 (9.48%)	0.889 (2.30%)	0.095 (-10.47%)	0.561 (5.08%)	0.889 (0.84%)	0.066 (-15.79%)
9	AL-LCC		**0.525 (23.47%)**	**0.891 (2.53%)**	**0.041 (52.33%)**	**0.534 (9.64%)**	**0.899 (1.96%)**	**0.031 (45.61%)**

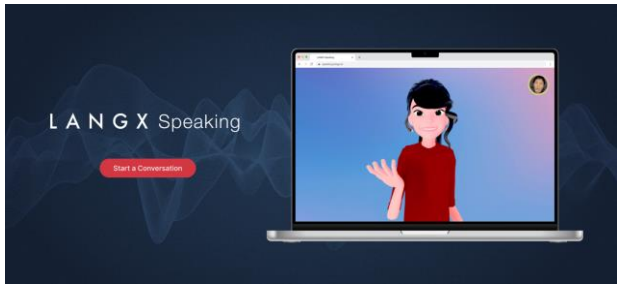
Sha, L., Li, Y., Gasevic, D., & Chen, G. (2022). Bigger Data or Fairer Data? Augmenting BERT via Active Sampling for Educational Text Classification. In *Proceedings of the 29th International Conference on Computational Linguistics* (pp. 1275-1285).

UCL Institute of Education

早稲田大学 WASEDA University

MONASH University

Tutorial English AI project



- ***Background.*** Reform of a campus-wide English language program at Waseda University, Japan (*N* = 3,000+) (cf. Nakatsuhara et al., 2023)

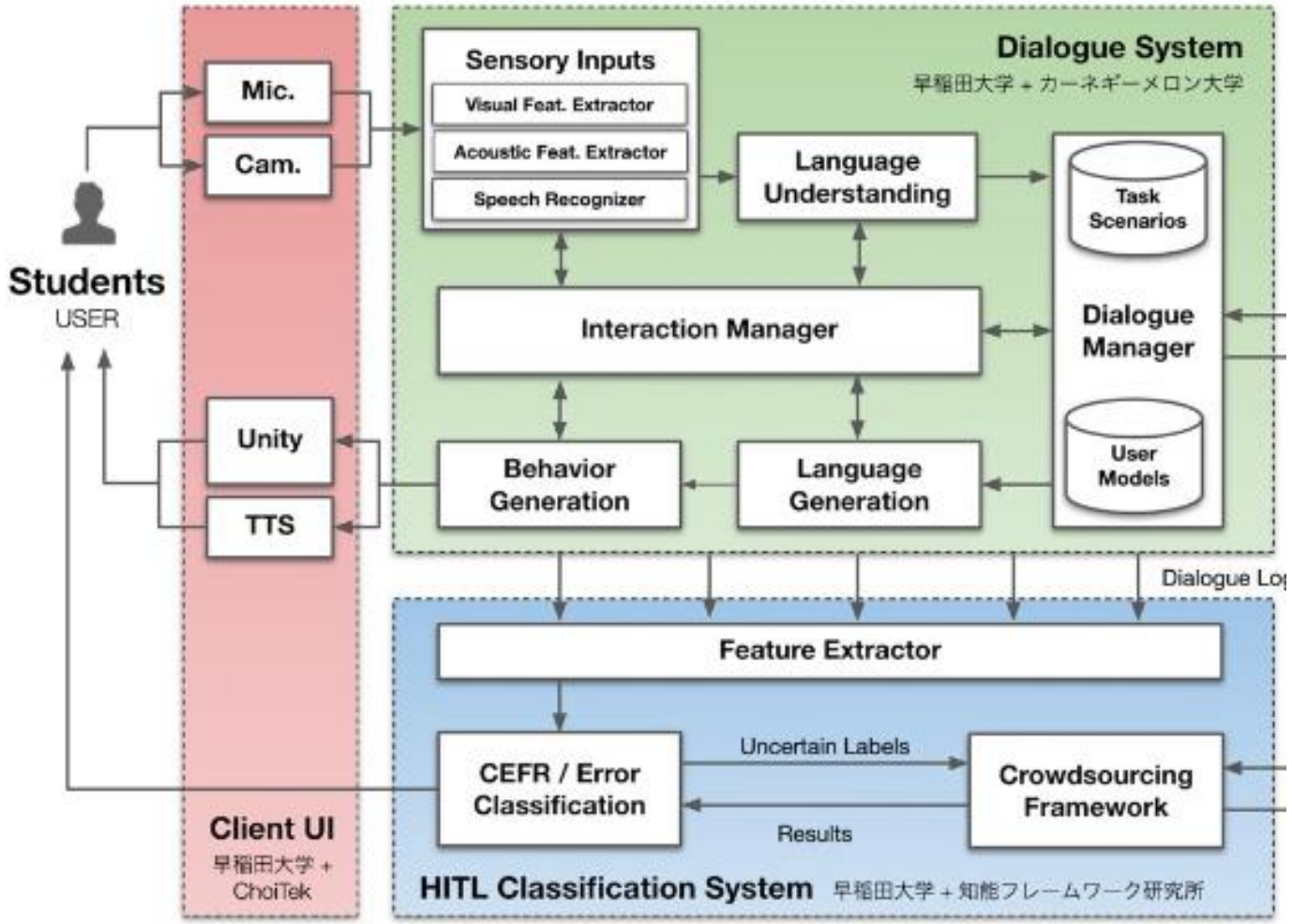
- ***Aim.*** Developing a fully automated speaking test for placement
  - ➢Elicitation: ***Conversational AI agent*** as an examiner (OPI) or peer (paired oral)
  - ➢Scoring: Neural-network algorithm with multimodal feature embedding

- ***Target constructs.*** Overall + 6 analytic criteria, incl. ***Interaction***

UCL Institute of Education

早稲田大学 WASEDA University

MONASH University



Major components

- **Automatic speech recognition** (perception)
- Natural language understanding (comprehension)
- Dialog manager (conceptualiser)
- Natural language generation (formulator)
- **Speech synthesis** (articulation)

UCL Institute of Education

早稲田大学 WASEDA University

MONASH University

Roleplay with AI-agent

Before the due … situation, but you have n…

You have recov… …ing any marks.

Your current circumstances are:

- You need one more day to complete the assignment. (You cannot finish it today).
- …although you have not got it yet.
- …the office.

■ ***Task:***

- In order to avoid losing any marks, ex… negotiate.

Ratability of interactional competence?

Is she professional enough?

Was the conversation engaging?

UCL Institute of Education

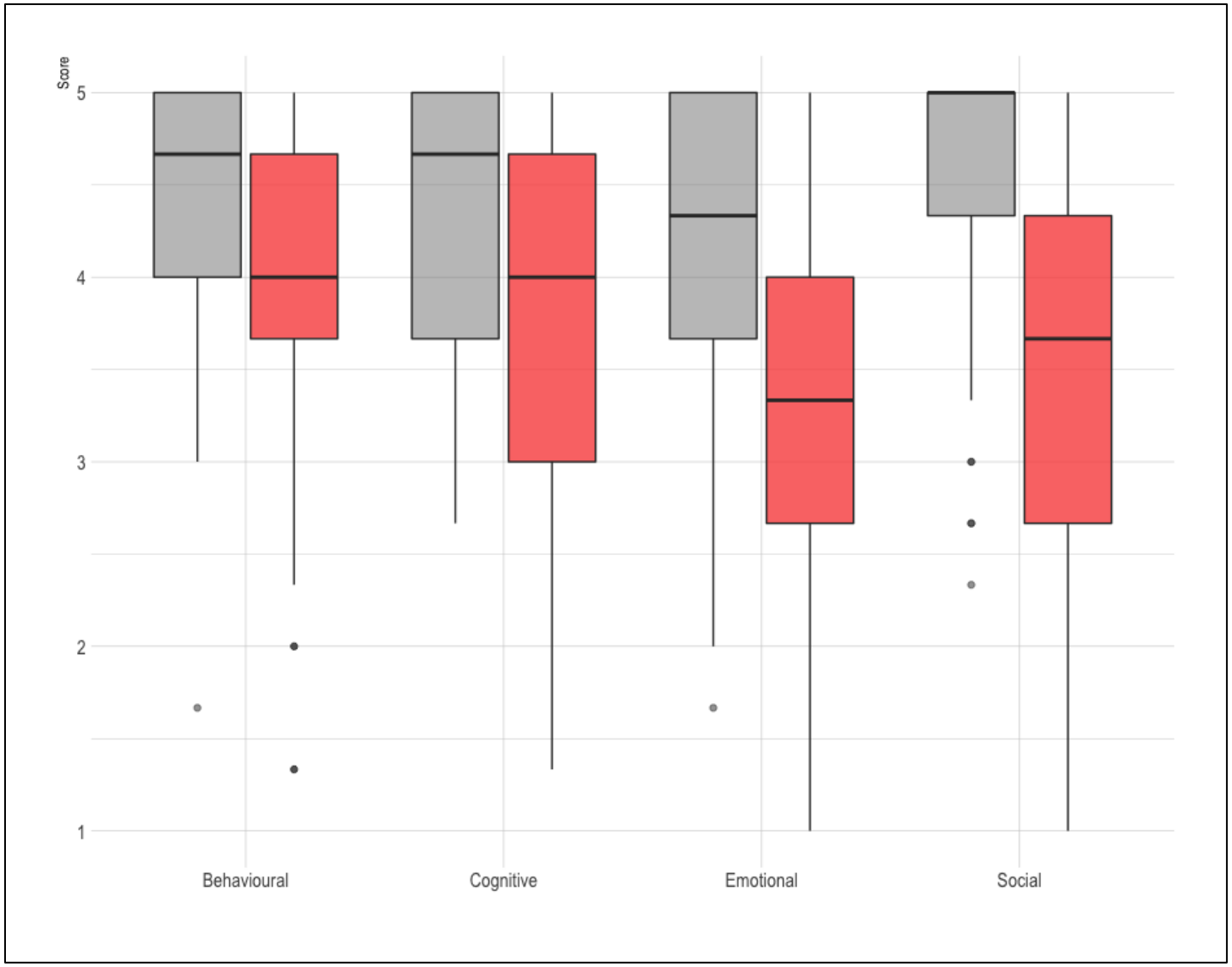
早稲田大学 WASEDA University

MONASH University

An exploratory study (Kurata et al., under review)

- ***Focus.*** Experienced teachers vs. AI agents in terms of Engagement
- ***Who.*** 74 Japanese learners of English
- ***Data.*** Three role-plays (Ikeda, 2017) ; Engagement with a 5-point scale

UCL Institute of Education
早稲田大学 WASEDA University
MONASH University



ANOVAs

	Tutors		AI
Behavirol	4.38	=	4.00
Cognitive	4.31	>	3.71
Emotion	4.10	>	3.30
Social	4.52	>	3.53

UCL Institute of Education

早稲田大学 WASEDA University

MONASH University

An exploratory study (Kurata et al., under review)

- ***Focus.*** Experienced teachers vs. AI agents in terms of Engagement
- ***Who.*** 74 Japanese learners of English
- ***Data.*** Three role-plays (Ikeda, 2017) ; Engagement with a 5-point scale
- ***Results.***
  - Teachers > AI-agent... in **Cognitive, Emotional and Social engagement**
  - Teachers = AI-agent...in **Behavioral engagement** (cf. performance)

UCL Institute of Education

早稲田大学 WASEDA University

MONASH University

Is *she* a friend or foe? – Potential & Challenges

Potential

- Can elicit real-time, somehow engaging interaction
- Can offer opprotunity for practicing ***transactional*** functions
- High practicality—Low costs, anywhere and anytime with the Internet

Challenges

- Quality of Interaction may be biased by pronunciation skills
- Authenticity in ***interpersonal*** functions of language

Concluding thoughts

- AI application to different aspects of PCIC education: developing Interactional Competence, automatic feedback, tailored learning prompts -> Huge potential
- Guided against reification of cultural stereotypes, lack in the higher-order interactional dimensions: emotional, logical and moral
- Moving forward: true interdisciplinary thinking, communication, practice and research
- A linguistic lens to AI-mediated communication
- Why compare AI interaction with human interaction? Technology shapes the way we communicate (Galaczi, 2023)
- Is the current AI technology really so much worse than me roleplaying a pregnant Vietnamese lady?

UCL Institute of Education

早稲田大学 WASEDA University

MONASH University

Watch this space…

Special forum: AI for Intercultural communication

Applied Linguistics Review










David Wei Dai and Zhu Hua (Editors)

Rodney Jones, Chris Jenks, Guanliang Chen, Spencer Hazel, John O'Regan, Shungo Suzuki, Giuliana Ferri, Adam Brandt

UCL Institute of Education

早稲田大学 WASEDA University

MONASH University

Thank you!

Do you have any questions or comments?

UCL Institute of Education

早稲田大学 WASEDA University

MONASH University