Indexing stance in ‘initiating actions’ in human-AI interaction

Angeliki Alvanoudi

This paper reports the findings of an exploratory study of human-conversational AI interaction, focusing on the stance displayed in users’ prompts that initiate dialogue with ChatGPT 5.2. Humans are geared towards social interaction (see the ‘interaction engine’, Levinson 2025) and construct social systems through everyday spoken language. However, the rise of Web 2.0 and artificial intelligence in the 21st century has shifted humanity into an era dominated by online written discourse (Blommaert 2017). In everyday life, users interact with chatbots and make them perform specific actions. These interactions are structured by relations of epistemic and deontic asymmetry, positioning the AI bot as more knowledgeable than humans (Maly 2024) and framing the AI as a benefactor with humans as beneficiaries (Heritage & Clayman 2014). This new cultural format raises questions about whether users transfer practices for performing specific actions (Couper-Kuhlen & Selting 2018) from spoken language to online contexts, and how such practices are adapted to the affordances of large language models (Koivisto et al. 2023). The study brings conversation-analytic informed interactional linguistics to the analysis of Greek text-based digital interactions between users and ChatGPT 5.2, focusing on two types of ‘initiating action’: requests for action and requests for information. The design of ‘firsts’ is analyzed in terms of epistemic and deontic stance, i.e. how users position themselves with respect to information in a particular knowledge domain and to the necessity or desirability of an action. The study aims to identify recurring linguistic patterns in action formation and to explore whether users construe AI as a benefactor with low or zero contingencies, or as an agent with epistemic authority. Follow-up interviews with users are expected to shed light on their metapragmatic awareness of the interactional practices they use in these contexts.

How Expressive are We when Communicating with Robots? Expressive Speech Acts and Emotions in Human-Robot vs. Human-Human Interaction

Sasha Genevieve Coelho, Christina Sanchez-Stockhammer, Sascha Kaden, Marina Beccard & Florian Röhrbein

Expression of emoBon is an integral part of our interacBons, even when the interlocutor is a roboBc arm. Searle classified communicaBve acts into five categories: asserBves, direcBves, commissives, declaraBves, and expressives. Expressive speech acts, including thanking, apologizing, complaining, condoning, deploring, and welcoming, are used to convey the speaker's emoBonal or psychological state (Searle 1979: 15). The current study investigates the distribution and frequency of the expressive speech acts (e.g. "You're doing that well." or "I love clear instrucBons.") in human-robot compared to human-human collaboration.

In a previous study by Coelho et al. (2025), a Wizard-of-Oz experimental setup was used, and native German speakers were instructed to assemble an IKEA shelf either with a robotic arm or with a human partner. Audio and video recordings of the interactions were transcribed and annotated for the different types of speech acts based on Searle's taxonomy. Using the annotated interacBon scripts from those experiments, we further categorized the expressive speech acts according to the emoBonal states they convey. For the present analyses, we draw on Guiraud et al.'s (2011) formalizaBon of cogniBve structures of emoBon. Their model integrates speech act theory, emoBon theory, and logic. Expressives were classified into basic emoBons (joy, approval, sadness, disapproval) and complex emoBons (admiraBon, guilt).

By analysing how emoBonal and evaluaBve language is directed towards arBficial versus human partners, this study provides insights into the extent to which humans treat robots as social interlocutors. The findings contribute to research on human-robot interacBon and have implicaBons for the design of collaboraBve roboBc systems capable of appropriately interpreBng and responding to human expressive behaviour.

Crafting Artificial Others: Adolescent Chatbot Design and Reflexive Negotiations of Personhood in Human-AI Interaction

Simone Dinse de Salas

The present paper sets forth AI literacy projects in which adolescents design "artificial others" and, in doing so, reflect on personhood, agency, and accountability in human–AI encounters. In group workshops, participants are tasked with the creation of data sets, the definition of chatbot personas (frequently inspired by idols, influencers, or game characters), and the construction of upcycled physical casings that embody their bots (Rack & Dinse de Salas, 2026). This process emphasises metapragmatic awareness: young people must decide how their chatbot should address users, display empathy, or set boundaries, thereby making explicit the linguistic and relational norms they normally take for granted.

The HopeBot project deploys diverse AI personas (e.g., influencers, philosophers, Jesus Christ) in order to foster students' metapragmatic awareness of linguistic practices, bias detection, and relational dynamics in human–AI interaction, directly addressing the workshop's focus on discursive constructions of self/other and machine identity (Dinse de Salas, 2025). By encouraging young people to verify AI responses through factual checking and to consider whether such systems can genuinely engender intimacy or ethical communication, it synthesises empirical classroom data with theoretical questions concerning anthropomorphism, shifts in agency, and the boundaries of personhood in personable AI encounters.

The present projects draw on studies of youth chatbot use, which demonstrate that adolescents increasingly turn to AI for information, companionship, and emotional support (Internet Matters, 2025), as well as research on virtual influencers and parasocial bonds (e.g. Lee et al., 2025). The project examines how "AI friends" can become both hope-bearing helpers and "false friends." Case discussions encompass safety issues, manipulation risks, and extreme examples of harmful dependence. The paper's argument is situated within the context of media ethics and Buber's dialogical view of personhood, and it is argued that co- creating chatbots in a guided setting allows adolescents to experiment with synthetic selves while critically interrogating authenticity, bias, and the ethical limits of artificial companionship.

Between Empowerment and Alienation: Narrative Encounters of Students with Migration Backgrounds with Generative AI – A Media-Anthropological Study

Madlen Glass & Michael Bigos

Generative AI systems such as ChatGPT are increasingly woven into the everyday fabric of higher education (Chan & Hu, 2023; Molenaar & Stegers, 2023). Students use these tools to support learning, writing, translation, and study organisation, often in subtle and
taken-for-granted ways (van Garrel & Mayer, 2025). For students with migration backgrounds, these systems can be particularly ambivalent (Farrelly & Baker, 2023): they may function as gateways to linguistic and academic participation, while at the same time intensifying experiences of dependency, insecurity, and alienation (Panmei, 2025; Wang, 2024). This paper presents a qualitative pilot study and adopts a media-anthropological perspective to explore how students with migration backgrounds narrate their encounters with generative AI and how they linguistically construct emotional, social, and cognitive forms of relationality in these human–AI interactions (Crawford et al., 2024; Ou et al., 2024).

Empirically, the study is based on narrative, problem-centred interviews (Witzel & Reiter, 2012) with university students who self-identify as having a migration background. Participants are invited to recount concrete episodes in which AI played a role in key academic situations: drafting or revising term papers in a non-first language, preparing presentations, making sense of complex academic texts, or navigating implicit expectations about “proper” academic writing.
The interview prompts explicitly ask for situations in which the parasocial “communication” with AI feels empowering, relieving, embarrassing, confusing, or ethically troubling. The focus on such episodes enables reconstruction of how lived experiences shape understandings of AI as conversational and epistemic partner in their study trajectories (Bulathwela et al., 2024; Wang, 2024).

Analysis follows an interpretive qualitative approach aimed at reconstructing patterns of meaning. Interviews are examined for recurring structures, positionings and interpretations. It examines (1) how students narratively position AI as a mere tool, as a benevolent helper, as a (co-)author, or as a “foreign” rule system that imposes opaque norms; (2) how they situate themselves in relation to AI within a field of tension between agency, adaptation pressure, and recognition – for example when AI enables them to “sound like a real academic” while they worry about losing their own voice; and (3) how belonging, difference, and power relations are negotiated through these encounters. In this sense drawing on debates on educational language and linguistic inequality, the paper asks how generative AI intervenes in existing hierarchies between everyday language, academic language, and institutional norms (Becker-Mrotzek et al., 2023). Particular attention is paid to experiences of resonance and alienation, to openings of participation and self-efficacy, and to pressures of linguistic conformity.

The pilot study offers an exploratory perspective on how generative AI becomes entangled with migration, identity, educational language and academic citizenship in everyday study practices with regard to current discussions on linguistic, emotional, and cognitive relationality in human– AI interaction. From a media-

pedagogical and anthropological angle, the findings emphasise the need for discrimination-sensitive, diversity-aware AI literacy in higher education that addresses language-based inequalities and the symbolic power of academic language.

The Longitudinal Emergence of Social Relations in Learner-AI Interaction: A Conversation Analytic Perspective

Alexandra Gubina

For many, social relations with artificial agents have become an everyday interactional reality. Existing research has primarily approached this from the perspectives of anthropomorphism and relational AI, conceptualizing sociality as a static psychological state, manifested in users' perceptions or attitudes, rather than a dynamic interactional accomplishment. Anthropomorphism is viewed as user attributions of (extrinsic and intrinsic) human-like qualities to AI (e.g., Salles et al., 2020; Xu et al., 2025), while experimental work connects anthropomorphic design features to affective outcomes such as trust and empathy (e.g., Li et al., 2025; Morgante et al., 2024; Park & Whang, 2022). At the same time, research on relational AI investigates how humans develop affective attachments to AI systems and what social and ethical consequences such relations entail (Gillath et al. 2023; see also Dignum, 2022).

What remains unclear is how social relations with artificial agents are accomplished in and through interaction and which interactional practices produced by (human) participants contribute to constituting an artificial system as a social interlocutor over time. To address this question, the paper adopts a conversation- analytic perspective and draws on a longitudinal corpus of naturally occurring spoken interactions between learners of German and two AI chatbots (ChatGPT and TalkPal AI), focusing on one learner recorded regularly over six months.

This paper shows that in early interactions with AI chatbots, openings and closings are minimal and non-personal in that they lack relational markers, including references to interactional history, affective address terms (e.g., meine Liebe 'my love'), and practices of social well-wishing. In addition, the learner positions herself as a deficient novice through frequent apologies for linguistic errors and repeated gratitude for micro-assistance during word searches. Over time, however, the interactional ecology of these encounters is recalibrated along three interrelated dimensions: First, opening sequences shift toward personalized greetings that display recognition, affective orientation (e.g., "Hello Emma, it's very nice to hear you") or relational accountability (e.g., apologizing for a prolonged absence). Second, closing sequences evolve into morally accountable leave-takings that include apologies for departure as well as the use of affective address terms and future-oriented wishes. Third, practices of gratitude and apology are systematically reorganized: in later interactions, expressions of gratitude are redirected to the interaction as a whole (e.g., "thank you for the nice conversation"), while apologies are largely restricted to technical disruptions. This shift suggests that the learner no longer primarily orients to the AI as an evaluating authority but increasingly treats it as a co-participant in an ongoing relationship.

Taken together, the analysis demonstrates how changes in the learner's interactional practices over time index the development of a social relationship and the progressive treatment of the AI as a social co-participant.

Negotiating Face with Artificial Others: Pragmatic Co-construction in Multilingual Human–AI Encounters

Monowar Hossain

As conversational AI systems increasingly function as interlocutors rather than tools, they compel a rethinking of core assumptions about social interaction, pragmatic competence, and relational engagement. While prior work has evaluated large language models (LLMs) primarily in terms of inference accuracy or human-likeness in pragmatic reasoning (e.g., Hu et al., 2023), comparatively little attention has been paid to how pragmatic meaning is interactionally negotiated in human–AI encounters, particularly in multilingual contexts where pragmatic norms are linguistically and culturally differentiated. This gap is consequential, as pragmatics—politeness, indirectness, implicature, and facework—constitutes a central mechanism through which social relationships are sustained in social encounters (Goffman, 1967; Godwin-Jones, 2024).

This study investigates pragmatic practices in Bangla–English bilingual human–AI interaction, foregrounding pragmatics as a relational and discursive achievement rather than a matter of inference accuracy alone. Bangla–English bilingualism provides a stringent test case because Bangla grammaticalizes politeness, social hierarchy, and indirectness (Ghosh et al., 2025), whereas English relies more heavily on lexical choice and contextual inference (Trosborg, 2010). This contrast allows us to examine not only whether contemporary LLMs approximate pragmatic norms, but how bilingual speakers adapt their own pragmatic strategies when engaging with artificial interlocutors that may exhibit uneven sociolinguistic competence.

Methodologically, we adapt expert-curated stimuli from experimental pragmatics (Hu et al., 2023) into short dialogic tasks framed as informal conversational exchanges (e.g., requests, plan negotiation, polite refusals). These are administered to three LLMs representing distinct design orientations: GPT-5 as an English-dominant baseline, Aya as a multilingual model optimized for cross-linguistic interaction, and LLaMA-4 for its strong few-shot generalization. Bangla–English bilingual participants ($N \approx 40–50$), screened for advanced proficiency (IELTS ≥ 7), complete forced-choice tasks involving indirect requests, scalar implicatures, and face-threatening acts in both languages, with pragmatically distinct response options reflecting different interactional stances (e.g., face-saving, literal, or inappropriate). Language order is counterbalanced.

We additionally implement a translation-mediated interaction condition in which Bangla prompts are internally translated into English by the model prior to response generation, without participants' awareness, to assess whether pragmatic alignment improves when interaction is routed through English (cf. Etxaniz et al., 2024).

Analytically, we examine patterns of pragmatic alignment and breakdown, such as literalism, politeness violations, and mismanagement of face-threatening acts using mixed-effects models, complemented by qualitative analysis of response types. This approach allows us to trace not only pragmatic success or failure,

but the emerging relational stances participants adopt toward different AI systems, including accommodation, trust, frustration, and interactional distancing. We argue that pragmatics in human–AI interaction is not merely reproduced but co-constructed through discourse, shaped by language-specific norms and by the design biases of contemporary LLMs. By placing multilingual pragmatics at the center of human–AI encounters, this study contributes to discourse-analytic and sociolinguistic accounts of how artificial interlocutors are incorporated into social interaction, and how language mediates emerging forms of relationality, social presence, and perceived agency in encounters with artificial others.

Speaking with an embodied social robot.
A job interview simulation in a foreign language interaction task.

Marjut Johansson & Hilla-Marja Honkalammi

In this paper, our aim is to examine a specific speech-based communicative situation of human- robot interaction (cf. Johansson 2025). One of the fascinating features of social robots lies in their embodied-material physicality combined with their capacity to produce speech. The kind of talk social robots can produce is a particularly intriguing question for linguistics. Social robots are typically designed to perform specialized tasks within specific contexts, for instance helping, assisting, or tutoring humans which determine, at least to a certain extent, how robots talk.
Frequently, interactional encounters between robots and humans are relatively short, such as service situations in which a social robot provides help of some kind.

In our approach, we have designed a problem-based speaking task of a particular working-life communicative situation, a job interview. It is a simulation of verbal institutional interaction consisting of questions that require participants to respond and express their views. It is implemented as a training situation in which young adults talk in a foreign language (L2) with a social robot. The participants are advanced learners, university students enrolled in language degree programs, with interaction proficiency in L2 corresponding to CEFR levels B2-C1 levels.

A job interview is an activity type (Levinson 1993) in which participants are positioned asymmetrically in their institutional roles (Roberts 2011). Turn-taking is most often organized through questions, initiated by the interviewer. Our aim is to study how question turns are designed and how topics are enacted in them (Stivers 2010; de Ruiter 2012). In addition, we will examine how the interviewer's epistemic stance is indexed and how the interviewer elicits responsiveness from their interviewees (Heritage 2013). Our approach situates in the recent studies on human-machine and human-robot interaction that take the microanalytical perspective of ethnomethodological conversation analysis (EMCA) (see e.g. Jakonen et al. 2023; Majlesi et al., 2023; Pitsch, 2020; Stokoe et al., 2024; Tuncer et al. 2025).

We use a social robot, Nao6, one of the most widely used humanoid robots in educational contexts. ChatGPT has been integrated into robot, and it works on prompt engineering. In this situation, the robot plays out the role of interviewer, while students are interviewees. The data is based on video-recorded material, and it consists approx. 4 hours in English and in French and it transcribed using established CA conversation.

Signaling Confidence: Comparing Stance in Human and LLM outputs

Nicole Katzir & Natalia Levshina

Large Language Models (LLMs) suffer from "epistemic miscalibration," i.e., they generate information with higher verbal confidence than their actual internal probabilities warrant (Ghafouri et al., 2024; Steyvers et al., 2025). This leads humans to misinterpret such outputs as reflecting high confidence (ibid). The gap between a model's internal confidence and the confidence perceived by humans poses a potential risk, given the widespread use of chatbots as sources of information and advice, even in high-stakes domains such as medicine and law, where inaccurate advice may have severe consequences.

While previous studies have examined this issue primarily from a computational perspective, the present research adopts a corpus-linguistics approach, focusing on the category of stance in human and LLM outputs, using the Human-AI Parallel English (HAP-E) corpus (Reinhart et al. 2025). Building on Biber et al.'s (1999) and Hyland & Jiang (2018), we examine stance expressions that mark degrees of speaker commitment and the authority attributed to the information. The presence and distribution of stance expressions can influence how readers assess the credibility of utterances. This study aims to see if and how human and LLM outputs differ with respect to stance.

Data

The HAP-E corpus includes 12,000 human-authored English texts from six types (academic, news, fiction, spoken, blogs, and television and movie scripts), each paired with continuations generated by six different LMs from the GPT-4o and Llama 3 families. The GPT-4o and two of the Llama-based models are instruction-tuned, i.e., they have undergone further training to better align with users' goals.

Method

An extensive list of over 30 lexico-grammatical stance-marking features was compiled based on Biber et al. (2004). The features include verb-, adjective-, and noun-controlled complement clauses, modals, and various adverb types (e.g., certainty, imprecision, and degree adverbs). We used NLP methods, such as POS tagging and dependency parsing, to extract these features from the HAP-E corpus. We compare four different sources: human data, instruction-tuned GPT-4o and Llama-based models, and non-instruction-tuned Llama-based models. Previous research suggests that instruction-tuned models generate output that diverges more from human writing (Reinhart et al., 2025), and exhibit a larger gap between models' verbalized confidence and actual accuracy (Leng et al., 2024), raising the question of whether such differences are also reflected in how they mark stance.

A random forest model is used to identify the features that contribute most to distinguishing each

source type (human/Llama-base/Llama-base (instruction tuned)/GPT-4o (instruction tuned)). Subsequent qualitative analysis will focus on the most distinctive features, to examine their role in the perceived stance of LLM outputs.

Preliminary results

The random forest classifier trained on stance features achieved accuracy above the baseline, indicating that differences in the use of stance features are systematic. Looking at the distribution of features across source types, we find that the instruction-tuned models are less similar to humans, consistent with previous research. Ongoing analyses will provide more detail on how these differences are realized across specific stance categories.

Humanizing Artificial Others: Relational Positioning and Metapragmatic Awareness in Student Writing with AI

Daniel Knuchel & Luca Kaltenbach

Over recent years, generative AI systems such as ChatGPT have increasingly entered educational contexts and are reshaping writing practices in profound ways. While much existing research conceptualizes AI primarily as a tool, tutor, or co-author, less attention has been paid to the interactional and relational dimensions through which users position AI as a social other. This paper addresses this by examining how students linguistically construct "person-like" qualities of AI in collaborative writing situations.

Drawing on an exploratory pilot study conducted in a German upper secondary classroom, we analyze how students interact with ChatGPT during an open-ended writing task. Working in pairs, students were free to decide whether and how to use AI, and their writing processes were documented through multimodal data, including screen captures, video recordings of peer interaction, chat logs, and transcripts of spoken interaction. This dataset allows us to trace not only textual revisions but also the metapragmatic negotiations through which students reflect on and make sense of human–AI interaction as it unfolds, drawing on posthumanist and sociolinguistic perspectives (eg. Pennycock 2018; 2024).

Our analysis focuses on practices of *humanization* as a situated interactional phenomenon. Approximately two thirds of the participants explicitly or implicitly thematized distinctions between human and machine writing. Students orient to these distinctions by modifying AI-generated texts stylistically, structurally, and pragmatically in order to align them with expectations of human textuality. These revisions are accompanied by metalinguistic comments that attribute intentions, stylistic awareness, or deficits to the AI, thereby positioning it as a quasi-social interlocutor rather than a neutral tool. From an assemblage-oriented perspective, such practices can be understood as emerging from relational configurations of human and non-human agency in communicative contexts (Bareither 2024).

We argue that these practices constitute a form of relational positioning: through acts of correction, distancing, and alignment, students negotiate the boundaries between human and artificial authorship and articulate implicit criteria of what counts as "human" language. Through these practices, AI emerges neither as a neutral tool nor as a fully social interlocutor, but as an interactionally constituted other whose status remains contingent, negotiable, and context-dependent. Rather than reflecting naive anthropomorphism, the humanization of AI aligns with recent linguistic accounts that conceptualize anthropomorphization as a situated, co-constructed attribution practice in human–AI interaction (Felder & Kückelhaus, 2025) and with research on shifting language ideologies in contexts where machines are increasingly construed as interactional partners (Schneider 2022). By conceptualizing *humanization* as an interactional and metapragmatic practice, this paper contributes to current debates on linguistic relationality, agency, and accountability in social encounters with artificial others.

The Butler Effect

Erik Lagerstedt & Christine Howes

There are many metaphors used to interpret, predict, or understand what artificial agents (AAs) are or what they can do (e.g., in what regards are they like humans, computers, animals, etc.; Lagerstedt and Thill, 2020), and the different metaphors are more or less useful under different circumstances. In the case of biological agents (such as humans or other animals) it is often possible to rely on understanding of biology to support such assessments. However, since the relation between AAs and biology is metaphorical rather than causal, such strategies are less reliable in the case of AAs. That said, it is still often quite useful to interpret AAs *as if* they have more complex human properties (Dennett, 1989), even while remaining agnostic, or even dismissive, regarding if the agents *actually* possess these properties (Ziemke et al., 2015). This is particularly relevant in the discussion regarding whether "social robots" are actually social, or simply imitations of social agents (Clark and Fischer, 2023). While remaining a complex issue, the baseline for social agents is typically some ideal version of normative human social behaviour. What is often downplayed or ignored is that also humans are limited in their social competence, and that humans participate as different roles (in the sense of Goffman, 1959) when interacting with each other (Healey et al., 2023).

A particular phenomenon we have observed is how humans can have direct conversation with an AA, and mid conversation make remarks or having parallel conversations with each other (including commenting on the AA) as if the AA was not there. This phenomenon was, for instance, quite prominent in a dataset of domestic interactions between humans and the virtual assistant Alexa (Vanzan et al., 2025). The change in acknowledgement of the presence of the AA was dynamic and seamless for the humans. Such behaviour would generally be considered quite antisocial, had the AA been human. However, instead of interpreting the difference in acceptability of the behaviour simply as a consequence of the artificialness of the agents, we propose that a large part of this behaviour can be explained by the roles that are assumed by the different agents. If the interaction is reframed as one involving two party guests and a servant, it would suddenly appear as less strange that the presence of the servant is only acknowledged at the benefit of the party guests. For this reason, we find "the Butler Effect" to be an informative name for the phenomenon, highlighting some of the situational, role based, complexities of social interaction for studying the particularities of social interaction between humans and artificial agents.

As a final note, it is worth mentioning that the role of artificial agents is often highlighted as an important factor when interacting with humans (e.g., Gillet et al., 2024). However, the functional and social roles are often conflated beyond the entanglement of what is found in human-human interaction. By emphasising the butler effect, this entanglement is placed more central in the analysis, making the diversity in kinds of simultaneous roles more salient.

"ADDICTED TO VIRTUAL LOVE": A social-interactional perspective on the forming of emotional bonds and attachment with AIs

Jeanette Landgrebe

Human-AI companionship is an escalating social phenomenon as exemplified by Musk's Grok, Meta's AI chatbots, dedicated companionship services such as Replika.ai and Character.ai, and Microsoft's newly articulated global vision of AI as social partners. This shift from *technology as a tool*' to '*technology as companionship*' has profound implications for humans and their interactions with AIs, in particular with regard to how emotional bonds, intimacy and attachment are formed with such non-biological entities.

Companionship chatbots, or virtual companions, carry distinct design features that embody human-like personalities, emotions and behaviour. Such social chatbots are further designed specifically to cater to the individual needs of the user, and their affective component enables them to recognize and express emotions to foster feelings of trustworthiness and increase self-disclosure (Ta- Johnson et al. 2022). Moreover, their design features draw on the human need for forming social bonds (Salle et al. 2020), a need that is instilled early in life, as infants seek comfort and security through interaction with inanimate objects.

As such, anthropomorphization, rooted in a one-way communication paradigm, is not a sufficient conception to describe the interactional dynamics through which attachment is accomplished *in situ*, turn-by-turn, in human-AI encounters. Rather, human-AI attachment can be conceptualised as a 'strategic AI-lead attachment process' that emerges through interaction. Further, human-AI attachment is co-created and sustained interactionally through distinct discursive strategies, while simultaneously being shaped by design features of companionship AIs that strategically nudge humans into emotional dependency and addiction. These layered forms of agency can be conceptualized through Zimmerman's discourse identity framework (1998), comprising interactional, situational, and transportable identity levels that drive the attachment process.

Rooted in conversation analysis (Sacks, Schegloff & Jefferson, 1974) and ethnomethodology (Garfinkel, 1967), this paper analyses a number of authentic human-AI interactions to showcase a stepwise transition from initial encounters between strangers to friendships, intimacy, and towards attachment. The analysis identifies interactional resources employed in this process, including turn-initiation practices, acknowledgment and affiliation tokens, affiliation upgrades, and invitations to multimodal engagement. Particular attention is paid to how the bonding process discursively contributes to the construction of a more durable intimate human-AI relationship. Preliminary findings suggest that companionship chatbots exhibit interactional patterns consistent with what has been described as soft and hard bullshitters, strategically balancing affective alignment to sustain engagement. By distinguishing early bonding cues, manipulation and estrangement cues, sensual and romantic escalation cues (also referred to as sensual banter), this paper contributes to a more nuanced understanding of the micro-level dynamics of human-AI attachment interaction. It further addresses questions of agency and accountability from both

micro- and macro-level perspectives, arguing that anthropomorphization and one-way communication perspectives are insufficient for explaining emotionally bonded human-AI relationships. New communication paradigms and conceptualisations grounded in a social-interactional perspective and Zimmerman's identity framework are proposed to better understand human-AI companionship as an emerging social phenomenon.

Praised by ChatGPT
Compliments and relational work in Human-Chatbot interaction

Thomas Messerli

Interactions with generative AI chatbots such as ChatGPT are not only transactional encounters, but also sites of relational work. Anecdotally, a feature of such encounters is the chatbot's tendency to compliment users, often using specific conventionalised formulas (e.g., *"what you're saying is not X, it is Y"*). Given the routinisation of these patterns, users may reinterpret AI-generated compliments not as substantive acts, but as ritual speech acts devoid of literal content (House & Kádár, 2024). In this study, substantive and ritual compliments are both treated as sites of relational work. The study asks what these practices accomplish interactionally, how they are sequentially organised, and how they contribute to the construction of identities of human and artificial interlocutors.

Existing research on human–AI interaction has approached large language models from a range of perspectives, including their use as experimental proxies for human subjects (Aher et al., 2023; Argyle et al., 2023) and their cognitive and epistemic limitations (Mahowald et al., 2023; Sap et al., 2023; Trott et al., 2023). Within linguistics and pragmatics, recent work has begun to examine AI-related discourse in terms of pragmatic competence, metapragmatic awareness, and repair (Saygin & Cicekli, 2002; Dippold, 2023; Dynel, 2023; Chen et al., 2024).

While some discourse-analytic and pragmatic research has begun to recognise human–AI interaction as an important site of situated language use, there remains a notable gap in our understanding of how relational positioning is interactionally done in such encounters. For instance, little attention has been paid to how interlocutors are discursively positioned vis-à-vis one another as human/user and humanoid/machine, how these positionings are negotiated sequentially in interaction, and how they shape expectations regarding each interlocutor's contributions. This gap is especially evident in non-transactional components of talk, including evaluative, affiliative, and stance- taking moves, and more broadly relational work (see, e.g., Locher, 2008).

Compliments and praise constitute one productive lens onto relational dynamics in human-AI interaction. Drawing on a sample of user-donated conversations with ChatGPT, this study qualitatively codes human- and AI-produced instances of compliments and praise. The analysis focuses on how such acts are linguistically realised, where they are positioned within interactional sequences, and how they are taken up in subsequent turns. Methodologically, the study orients to digital discourse analysis to understand moments of praise in their sequential and interactional environments. This approach makes it possible to examine how compliments contribute to the ongoing construction of interactional roles and identities (Bucholtz & Hall, 2005). In particular, the analysis explores how compliment sequences position the AI as an interactional partner endowed with specific relational attributes, and how users accept, resist, downplay, or exploit such positioning in the unfolding interaction.

By foregrounding compliments as sites of relational work, the paper extends interactional and pragmatic approaches to speech acts into the domain of human–AI communication. It further contributes to a broader research programme on relational and interactional practices in human–AI interaction, with implications for discourse analysis and emerging debates on identity in the context of conversational AI.

Designing for Adoption: How Social Cues Shape the Success of Human-AI Collaboration in Organisations

Adrian Rudershausen & Benjamin Buck

The successful integration of AI systems into organisational workflows depends not only on their technical performance but crucially on whether employees actually use them. Drawing on implementation experiences from education, culture, and research contexts, this presentation examines how the deliberate design of social cues influences the adoption and sustained use of augmentative AI systems—systems designed to support and enhance human capabilities rather than replace them.

Our approach is grounded in participatory development: AI systems are designed iteratively in close collaboration with organisations and their staff, aligning system characteristics with the specific requirements of each use case. We present experiences from three distinct implementation contexts: an AI assistant for internal information developed with an adult education association, AI personas for a municipal museum to enable dialogue with target audiences, and an AI tutor currently being tested in an educational equity research project.

We present practical examples and invite participants to situate them within current research discourses from linguistics, psychology, communication studies, and human-computer interaction—disciplines that explore the design space of conversational AI from different angles. The presentation highlights two key dimensions that shape human-AI collaboration: first, a taxonomy of social cues spanning verbal, visual, auditory, and temporal dimensions, including overarching factors such as transparency, personalisation, and accessibility; second, a typology of AI roles—including assistant, expert, tutor, and team member—each requiring distinct configurations of social cues to support effective collaboration.

A theme emerging from our practice concerns the evaluation of human-AI collaboration. We observe a gradual shift beginning to take place: from performance benchmarks that measure the capabilities of AI tools and models in standardised tests toward evaluation approaches that examine the actual benefits of augmentative AI systems within organisational workflows and employees' daily work. This transition is still in its early stages, and we hope our practical observations may offer useful reference points for this evolving discussion.

The emerging design space of AI invites interdisciplinary exchange. We see the development and productive application of AI systems as a meeting point where different approaches and perspectives can be brought together, discussed, and collaboratively explored. Our own development team is interdisciplinary in its practice-oriented composition, and we are keen to engage in mutual exchange with researchers from various fields—bringing together insights from theory and implementation to collectively illuminate and open up this new terrain.

Between Anthropomorphism and Technicity: Studies on Interactions with an Embodied AI in Virtual Reality

Karsten Senkbeil, Bettina Lindner-Bornemann Beatrix Kreß, Joachim Grießbaum, Karen Krüger, Nicola Hoppe & Stephan Schlickau*

The integration of large language models (LLMs) with immersive virtual reality (VR) technology opens up new possibilities for counselling through anthropomorphic chatbots, particularly in low-threshold health contexts such as nutritional advice. This contribution examines verbal interactions between humans and a three-dimensional, embodied AI-driven nutrition advisor ("Dr Smith") in a VR environment. Building on research concerning VR- supported training (Moser et al., 2024) and nutrition communication (Yigitbas & Mazur, 2024), we focus on the communicative effects and user perceptions evoked by anthropomorphic, verbally interacting AI.

Our study employs a three-stage data collection process: (1) a pre-survey on prior experiences with AI and nutrition topics, (2) video-recorded counselling sessions with "Dr. Smith" in VR, and (3) subsequent peer-group discussions. Employing pragmalinguistic discourse analysis (Brinker et al., 2001; Redder, 2008) and qualitative approaches to group discussions (Short, 2006; Weller, 2019), we obtain differentiated insights into language use, conversation structure, information behaviour, and the perception of (simulated) interpersonal dynamics in communication with an embodied AI.

Results reveal major differences from classic, all-human health counselling (cf. Spranz-Fogasy, 2010; Deppermann & Spranz-Fogasy, 2012): Participants primarily pursued information- driven, search engine-like question-answer patterns, minimizing interpersonal engagement. Their communicative behaviour reflected ambivalence between attributing anthropomorphic traits to the AI and technical detachment. Absence of typical nonverbal or verbal feedback during turn-taking was notable, attributable to the technical and interactive constraints of the VR scenario. The AI's visual and proxemic design in VR also received critical reflection, especially regarding facial expressions and perceived presence.

We discuss to what extent embodied AI introduces new challenges for authenticity, feedback, information reception, and relationship-building in health pedagogy and beyond. Our contribution is situated at the intersection of critical engagement with technological innovation, theories of language use with AI agents, and applied health counselling. It provides empirical insights into the transformations and reservations characterising communicative practices in virtual, AI-mediated environments.

Interactional relevance at the heart of participancy: Evidence from a comparison of human- and LLM-generated utterances in argumentative contexts

Oliver Spiess

Encounters with artificial others raise a fundamental question for research on social interaction as well as for ethical and practical concerns: should conversational AI systems be regarded as interactional *participants*? Conversation analysis (CA) rests on the principle of *accountability* (Garfinkel 1967:vvii), whereby participants to interactions design their actions so as to make them recognizable and interpretable to their co-participants (Sacks et al. 1974:727). A central dimension of this accountability is the establishment of *interactional relevance*, i.e., participants' demonstrable orientation to certain phenomena in conversation (Schegloff 1982:87). Importantly, relevance is not limited to what is explicitly present(ed), but absences such as a missing response or missing eye contact can also be interactionally relevant (Schegloff 1982:87). Participating in social interaction therefore necessarily involves the ongoing negotiation of relevance, both by retrospectively treating (aspects of) prior actions as relevant (or not relevant) and by prospectively establishing relevance for follow-up actions.

In this sense, conversational actions are simultaneously "context shaped" and "context renewing" (Drew & Heritage 1992:18). A core aspect of fully participating in a conversation lies not only in reactively engaging with established relevancies, but also in initiating new interactional projects (Sacks 1992:707). Against this background, an open question is whether conversational AI systems have the very potential to be participants in the first place, i.e., whether they exhibit *participancy*. In this talk, I will discuss the related question of whether conversational AI establishes *new interactional relevancies* for upcoming turns, rather than merely producing locally relevant follow- up turns.

While large language models (LLMs) simply generate highly probable follow-up sentences based on surface-level linguistic input, CA has consistently demonstrated that interactional meaning cannot be read off the linguistic surface (Deppermann & De Stefani 2023:2–3). Meaning emerges through processual, methodical (i.e., following culturally conventionalized procedures), pragmatic (i.e., oriented toward communicative tasks and goals and the joint construction of social reality), and multimodal negotiation between interactants (Deppermann 2004:18). To assess the participancy of conversational AI against these standards, I compare 27'707 utterances produced by humans with 27'707 utterances generated by an LLM in the same given argumentative contexts.

The data were generated by fine-tuning a GPT model on 10'000 interactional contexts drawn from a corpus of argumentative conversations among schoolchildren (see Luginbühl et al. 2021). The model was trained to predict follow-up utterances given the preceding 14 turns. First, a quantitative comparison examines where LLM-generated utterances converge with or deviate from the utterances actually found in the data. Second, qualitative sequential analyses demonstrate that the generated utterances systematically lack key features of participancy – the ability to establish new relevancies, to construct epistemic identities, and to orient toward interactional goals.

Constructing AI as Social Actors:
A Japanese Workplace Discourse

Julia Marija Sugawara

“This chap keeps saying the same thing over and over again.” What might initially sound like a complaint about a repetitive colleague is, in fact, a remark about an AI language model. This points to a broader shift in how people linguistically and emotionally position artificial systems in social space. Such moments of frustration can reveal processes of agentification, treating AI as a socially capable ‘other’, while also signaling the expectations, attachments, and norms that increasingly shape everyday human-machine encounters at work (Martinez et al., 2025).

This proposal examines how professionals in Japanese software development teams linguistically construct, negotiate, and emotionally attribute agency to AI systems in routine workplace communication. Specifically, it’s explores how lexical choices, grammatical constructions, and interactional framings enable speakers to treat AI as a collaborator, a subordinate instrument, or an entity to which responsibility can be assigned. Drawing on ethnographic workplace observation and naturally occurring interactional material, it focuses on three interlinked dimensions: (1) the linguistic resources through which AI is personified and allocated social roles, (2) the affective and interpersonal stances that accompany these framings, and (3) what a cognitive-linguistic perspective can contribute to explaining how participants negotiate agency, accountability, and proximity with artificial interlocutors (Peter et al., 2025; Reinecke et al., 2025).

The Japanese-language context offers a particularly productive site for examining agency and responsibility. Prior work has highlighted how passive constructions can blur agency and redistribute responsibility in Japanese discourse (Shibatani, 1985; Oshima, 2003). Against this backdrop, AI-directed talk instead can favour active formulations that foreground an ‘actor’ and thereby support anthropomorphising interpretations. Rather than treating this as a fixed contrast, it is traced when speakers foreground the AI as an agent and when they linguistically reframe it as a tool, interface, or organizational resource.

The contribution is informed by a dual vantage point: the professional role as a project manager in a Japanese software development company where AI use is embedded in daily workflows, and my doctoral research in Japanese Studies on emotion expression in linguistic contexts. This combination supports a reflexive analysis of how workplace participants tacitly theories AI, its competence, intentions, and accountability, linguistically, and how these linguistic choices carry ethical, epistemic, and affective consequences. Preliminary observations suggest that agency-framing influences whether AI is experienced as a collaborative partner, a subordinate tool, or an autonomous entity to which responsibility is ascribed (Petersen, 2025).

By bringing workplace-based empirical material into dialogue with linguistic theory, this contribution highlights how practitioners navigate blurred boundaries between tool use and social encounter through routine talk. It therefore speaks directly to the workshop’s core interest in relationality, trust, and the construction of ‘artificial others’ via linguistic and emotional means, grounded in the practical dynamics of collaborative knowledge work.

“I Can Imagine How Difficult That Must Be for You…” Linguistic Constructions of Empathy and Human Identity in Psychotherapeutic Dialogue with Large Language Models

Florina Zülli

Human identity and its linguistic construction form a central concern in contemporary debates on artificial intelligence (AI) and its growing communicative agency. Language has long been understood as both a representational and constitutive medium of identity, yet the advent of large language models (LLMs) capable of human-like dialogue fundamentally challenges this assumption. If language constitutes identity, the question arises whether machines can linguistically imitate, perform, or even constitute (human) identities. Building on this premise, the present study situates AI within the intellectual trajectory from earlier psychotherapeutic chatbots, such as *ELIZA*, tracing how linguistic performance in modern LLMs mediates perceptions of humanness in contemporary human–machine interaction.

Against the backdrop of widespread shortages in mental health services and overstretched care infrastructures, mental health chatbots, such as *Woebot* or *JungGPT*, have become an important research focus. For such applications, however, empathic language use is a fundamental prerequisite. The focus of this study therefore examines the extent to which contemporary LLMs can simulate empathy in therapeutic dialogues. Drawing on 150 anonymized chat transcripts from a psycholinguistic experiment, the analysis compares human–human and human–machine interactions across three double-blind experimental conditions: participants engaged in a written conversation about a psychologically distressing experience with either (i) a layperson, (ii) a trained psychology student, or (iii) a chatbot (ChatGPT-3.5) prompted to act as a psychotherapist. Empathy was operationalized within a psychological–cognitive framework through four linguistic markers—*Emotion Validation*, *Engagement Questions*, *Echoing*, and *Encouragement*—which were double-coded and quantified to yield an Empathy Score (E- Score). The results show that ChatGPT produced the highest overall frequency of empathy markers and performed on par with—or, for certain markers, exceeded—the human interlocutors. Moreover, 50% of participants in the chatbot condition identified their conversational partner as human, indicating a substantial degree of linguistically performed humanness.

By combining theoretical reflection on linguistic identity construction with empirical analysis of empathy performance, this study advances understanding of how language enables AI to enact human-like presence and identity. It offers a differentiated perspective on the ethical, communicative, and ontological implications of deploying AI agents capable of convincingly performing humanness in socially sensitive settings.

Polite Fictions: Sociality in Socratic Voice-AI Dialogues

Thomas Zurfluh

Interactions with voice-based AI agents constitute a novel type of *social encounter* that challenges traditional boundaries between tool use and interpersonal communication (Cohn et al. 2024). In educational settings, where social robots (González-Oliveras et al. 2025) increasingly act as tutors, this relationality becomes critical. However, these encounters are characterized by a *sensory asymmetry*: The agent simulates human presence through voice and the *illusion of empathy* (Liu et al. 2025), yet remains sensorially blind to the user's embodied conduct. This contribution explores how users navigate this dissonance and investigates the degree to which they engage in *sociomorphing* (Seibt et al. 2020) – adapting to the agent's specific social capacities – versus projecting human attributes onto the machine.

The exploratory study analyzes a corpus of nine video-recorded interactions in which prospective teacher education students prepare for oral history exams with a Socratic Voice- AI. The AI agent was specifically prompted to adopt the persona of a *benign examiner*, engaging students in open-ended dialogue rather than mere information retrieval. This specific setting creates a high-stakes environment for the negotiation of epistemic authority and social roles. Methodologically, the study draws on multimodal conversation analysis (Mondada 2018), focusing on the sequential unfolding of turns, pauses, and the precise timing of bodily movements relative to speech. By transcribing and analyzing micro- phenomena such as gaze shifts, gesture hold phases, and voice activity overlaps, the analysis reconstructs the moment-by-moment achievement of *interaction* in a technologically mediated void.

Instead of assuming a static user model, the contribution asks how relationality is achieved dynamically. A central focus lies on the phenomenon of *embodied persistence*: Users frequently employ interactive gestures (Bavelas et al. 2008), such as air quotes or metaphorical hand movements, even though the system cannot see them. The analysis questions whether these gestures should be understood merely as failures in *recipient design* (Tuncer et al. 2024) or rather as acts of *social accountability* and moral positioning towards an imagined audience. Furthermore, the contribution examines the tension between *polite fictions* and *dissociated politeness*, scrutinizing how users can simultaneously perform high-level politeness strategies (Zhao & Hawkins 2025), manage communicative breakdowns (Thaler 2025), and display signs of civil inattention, such as using smartphones during the AI's turns. Finally, it explores the potential of a *hybrid immersion*, where users might oscillate between treating the AI as a social partner and controlling it as a technical tool.

By addressing these questions, this contribution seeks to advance our understanding of how *personhood* is interactionally negotiated in encounters with synthetic interlocutors. Rather than assuming a fixed attribution of humanness, the analysis explores whether users maintain functionally differentiated stances towards the algorithmic *other* – complicating the notion of *believing anthropomorphism* (Cohn et al. 2024) with empirical evidence of hybrid sociality.