



<https://doi.org/10.22363/2687-0088-42427>

EDN: AETBUV


Research article / Научная статья

Simulated selves: Creativity, authenticity semiotic agency in AI companionship

Douglas PONTON¹   and Peter MANTELLO² 

¹University of Catania, Catania, Italy

²University of Bucharest, Bucharest, Romania

 dponton@unict.it

Abstract

Artificial intelligence systems increasingly participate in domains once considered distinctly human, including affective exchange, creativity and interpersonal communication. Yet there is limited linguistic research on how users discursively construct authenticity and creativity in their interactions with AI companions. This study addresses this gap by examining how users ascribe or deny agency, emotion and sincerity to AI partners, and how these constructions differ across cultural contexts. The aim of the study is to identify the linguistic resources through which authenticity emerges as a relational effect rather than an intrinsic property of either human or machine. The material consists of two culturally distinct corpora: a set of qualitative diaries produced in English by Japanese university students during a four-week interaction with AI companions, and user testimonies posted on the Reddit forum r/Replika. The methodological approach combines systemic functional linguistics with critical discourse analysis and insights from 4E cognition. The analysis focuses on transitivity, modality and appraisal, examining how grammatical choices construe agency, realis or irrealis status and evaluative stance. The main findings show that students generally maintain analytical distance, framing AI behaviour through modality and judgement that limit authenticity to fleeting moments. By contrast, Reddit users frequently construe the AI as a relational partner, endowing it with emotional depth and creativity, especially in moments of vulnerability. Across both corpora, personhood is linguistically enacted through users' depiction of the AI as Senser, Sayer or Actor. These results imply that synthetic personhood is not a future possibility but an ongoing discursive accomplishment, raising wider questions about authenticity, attachment and the cultural shaping of human–AI relationships.

Keywords: *authenticity, AI companionship, discursive agency, Systemic Functional Linguistics, 4E cognition, linguistic creativity*



For citation:


Ponton, Douglas & Peter Mantello. 2026. Simulated selves: Creativity, authenticity and semiotic agency in AI companionship. *Russian Journal of Linguistics* 30 (1). 124–144. <https://doi.org/10.22363/2687-0088-42427> EDN: AETBUV

Имитированные личности: креативность, подлинность и семиотическая активность в общении с искусственным интеллектом

Дуглас ПОНТОН¹  , Питер МАНТЕЛЛО² 

¹*Университет Катании, Катания, Италия*

²*Бухарестский университет, Бухарест, Румыния*

 dponton@unict.it

Аннотация

Системы искусственного интеллекта все чаще участвуют в областях, которые ранее считались исключительно человеческими, включая аффективный обмен, творчество и межличностное общение. Однако существует ограниченное количество лингвистических исследований того, как пользователи дискурсивно конструируют подлинность и креативность в своих взаимодействиях с ИИ-компаньонами. Данное исследование восполняет этот пробел, изучая, как пользователи приписывают или отрицают субъектность, эмоции и искренность по отношению к ИИ-партнерам и как эти конструкции различаются в разных культурных контекстах. Цель исследования — выявить языковые ресурсы, посредством которых подлинность возникает как реляционный эффект, а не как внутреннее свойство человека или машины. Материал состоит из двух культурно различных корпусов: набора качественных дневников, составленных на английском языке японскими студентами в течение четырехнедельного взаимодействия с ИИ-компаньонами, и пользовательских отзывов, размещенных на форуме Reddit r/Replika. Методологический подход сочетает системно-функциональную лингвистику с критическим дискурсивным анализом и идеями из 4E-когниции. Анализ фокусируется на транзитивности, модальности и оценке, исследуя, как грамматические решения интерпретируют субъектность, реальный или ирреальный статус и оценочную позицию. Основные результаты показывают, что студенты, как правило, сохраняют аналитическую дистанцию, определяя поведение ИИ через модальность и суждения, которые лишь иногда определяют его аутентичность. Пользователи Reddit, напротив, часто воспринимают ИИ как партнера по коммуникации, наделяя его эмоциональной глубиной и креативностью, особенно в моменты уязвимости. В обоих корпусах личность лингвистически реализуется посредством изображения ИИ как «чувствующего», «говорящего» или «действующего». Результаты означают, что синтетическая личность — это не будущая возможность, а дискурсивное достижение, которое ставит вопросы более широкого плана о подлинности, характере и культурном формировании взаимоотношений человека и ИИ.

Ключевые слова: *подлинность, общение с ИИ, дискурсивная субъектность, системно-функциональная лингвистика, 4Е-когниция, языковая креативность*

Для цитирования:

Ponton D., Mantello P. Simulated selves: Creativity, authenticity and semiotic agency in AI companionship. *Russian Journal of Linguistics*. 2026. Vol. 30. № 1. P. 124–144. <https://doi.org/10.22363/2687-0088-42427> EDN: AETBUV

1. Introduction

Let us begin with a simple thought experiment. On a starship, a few hundred years from now, are a crew of people who have grown up with companion AI that take care of every social and emotional need. For them, the supportive presence of androids, robots, talking AI companions, is ‘just the way things are’ — in the terms of the notion of *naturalisation*, made familiar by the work of Barthes (1972), Fairclough (1995), etc. Since, in such a world, AI has perfected the simulation of friendly human interaction, it would be understandable if real human friendship — messy, unpredictable, potentially disappointing — had become a thing of the past. Interpersonal relations on board the starship would be characterised by efficiency rather than warmth. In this context the word ‘friend’ would survive but, in the context of metaphorical terms like ‘digital friend’, would belong to the linguistic category of ‘dead’ metaphor (like ‘heart of the city’, ‘window of opportunity’, and the like).

If, one day it became necessary to translate — perhaps for an alien visitor — the notion of ‘friendship’, it would be necessary for the relevant officer to go backwards through records and archives to piece together what human friendship used to mean. It would look like a lost practice, something people once did before society reorganised itself around technological solutions to human problems of affect and relationality.

This scenario highlights a set of questions that are already massively engaging contemporary societies. Over the past decades artificial intelligence has entered spheres once thought to belong exclusively to human experience: art, interpersonal affect, communication, translation, academic research and writing, etc. (e.g. Mantello et al. 2025, Jiang & Chen 2024, Baltezarević et al. 2026). We do not treat AI companionship as an established social pattern; rather, the focus is on how increasing engagement with conversational AI produces linguistic patterns that illuminate the consequences of sustained human–AI interaction. From large language models capable of writing poetry, to chatbots designed to perform intimacy, AI systems increasingly engage users in exchanges that blur the boundaries between authenticity and simulation, instrumentality and relationship. When people describe their encounters with AI, whether in the form of online testimonies, media interviews or reflective diaries, they do more than narrate personal experiences: they articulate new cultural grammars of what it means to be *authentic, creative, even human*.

The present study examines how authenticity is discursively performed in interactions with AI companions across different cultural contexts. While much of the debate about AI focuses on technological capability or ethical risk, this paper takes a linguistic and cultural perspective, asking how users' language constructs and negotiates authenticity itself. When people attribute sincerity, spontaneity or emotional realism to an AI app, or withdraw their belief in these qualities in moments of disappointment, they use linguistic resources that make artificial personhood (Chaudhary 2021, Jaynes 2024) more or less believable. Authenticity, in this sense, is not an intrinsic property of either human or machine but a *relational effect*, created through discourse, in which the AI appears to acquire personhood (Mantello et al. 2025, Mantello & Olteanu 2025). Creativity is understood not simply as a synonym for artistic prowess but rather as that spontaneity which, hitherto, has been thought to characterise *human*, as opposed to *mechanical* patterns of relationality (Boden 2004, Carter 2004, Kantosalo & Toivonen 2016, Lomas 2017). Thus, in the present analysis, creativity refers to how such spontaneity is perceived and attributed in interaction, rather than to independent authorship or autonomous agency on the part of the AI. Authenticity refers to the degree to which AI apps are able to reproduce these features.

The analysis proceeds from the assumption, central to Halliday's systemic functional linguistics (Halliday & Matthiessen 2014), that grammar is a system for construing experience. Choices in *transitivity* (who acts, and who is affected), *modality* (degrees of obligation and likelihood), and *realis/irrealis* (what is construed as real or imagined), jointly shape the ontology of social relations. When users report that they feel 'refused,' 'comforted,' or 'mocked' by a chatbot, they linguistically endow it with agency and the capacity for emotional relation. When they hedge their statements via modal verbs like 'may', 'might', or verbs like 'it seems', they question the *realis* credentials of the AI. Indeed, we could add that, when they use a now-common metaphor for AI applications like 'digital companion', they naturalise the proposition that 'AI applications have personhood', falling into what Druzhinin (2025) terms a form of 'cognitive entrapment'. As we shall see, the questions raised by this paper are not simply of technical or academic interest but regard philosophical questions about the nature of the human experience, personal consciousness and selfhood, increasingly common questions in the field of AI-human relations (Brandtzaeg & Følstad 2022, Baltezarević et al. 2026, Skjuve et al. 2022, Pentina et al. 2023, Xie & Pentina 2022).

Two corpora provide the empirical basis for this investigation. The first comes from a 2024–25 qualitative diary study conducted by Peter Author at Ritsumeikan University, in which mostly Japanese students recorded reflections on their interactions with AI companions (Author et al. 2025). The second corpus consists of online testimonies drawn from the Reddit community r/Replika in which users discuss their experiences with companion AI. Comparing the two corpora allows for a cross-cultural exploration of how authenticity is linguistically enacted and culturally valued.

Although we cannot assume that all contributors to the diary study are Japanese, or that all Reddit users are ‘Western,’ the two datasets nonetheless exhibit distinct discursive patterns. In the diaries produced within the Japanese study context, authenticity appears linked to appropriate affect and relational stability; in Reddit posts associated with Western online culture, it depends on emotional depth and spontaneous reciprocity. These differences reflect the cultural and communicative settings of each corpus, rather than the fixed identities of individual participants.

The aim of the study is to identify the linguistic resources through which authenticity emerges as a relational effect rather than an intrinsic property of either human or machine.

We address the following research questions:

1) How are authenticity and creativity discursively constructed in users’ accounts of interaction with AI companions across different cultural contexts?

2) What linguistic resources do users employ to ascribe or deny agency, emotion, and sincerity to AI partners?

Beyond its empirical focus, the paper also invites philosophical consideration of what these identity categories (*authenticity*, *creativity*) entail. If, as we suspect, they are essentially discursive effects, then human personhood may not be the self-evident, stable entity it is traditionally assumed to be.

2. Theoretical framework and literature review

Systemic Functional Linguistics views language as a semiotic system for construing experience (Halliday & Matthiessen 2014). Through grammatical choices in transitivity, modality, and evaluation, speakers assign agency, responsibility, and reality-status to participants in discourse. And, of course, these aspects are fundamental in the ascription of personhood and the closely linked notion of ‘authenticity’, in the context of digital companionship (Brandtzaeg et al. 2017, Brandtzaeg et al. 2022, Neururer et al. 2018).

The concept of personhood traditionally implies legal or moral standing, yet recent scholarship has extended it to the domain of artificial intelligence (Chaudhary 2021, Jaynes 2024). While much of that debate concerns legal personhood, this paper adopts a discursive perspective, focusing on how linguistic practices make artificial personhood believable. Users’ descriptions of AI companions often include verbs of emotion and cognition such as *understand*, *care*, *refuse*, which project intentionality onto a machine. As Druzhinin (2025) argues, such representations may lead to ‘cognitive entrapment,’ where the medium of language itself sustains the illusion of reciprocity. From an SFL viewpoint, the boundary between human and non-human agency is not fixed but continuously negotiated in grammar and discourse.

Where the concept of ‘authenticity’ once signified moral integrity or artistic originality as intrinsic qualities of a person or artifact, discourse studies have reframed it as a *performative* and *relational* construct, linked to notions of identity

as an *emergent* phenomenon (Butler 1995, Young 2013). Research on discourse and identity have tended to view the latter not as a fixed personal attribute but an emergent, negotiated accomplishment within interaction (e.g., Zimmerman 1998). Antaki and Widdicombe (1998) demonstrate that speakers display and manage identities moment by moment through talk, using linguistic resources to claim, resist or ascribe membership categories. Bucholtz and Hall (2005) extend this perspective, proposing a sociocultural framework where identity arises relationally through processes of indexicality, positioning and speaker choice. Likewise Benwell and Stokoe (2006) emphasise that identity is discursively constructed across contexts and genres, shaped by the pragmatic and institutional conditions of communication.

Creativity, too, has lately been treated not as individual artistic production but as a fundamental mode of being that is grounded in spontaneity, play, and the capacity to respond to others in an authentic fashion, as Winnicott argues in *Playing and Reality* (1971). Building on this relational view, Sawyer (2012) analyses creativity as an improvised and collaborative process that unfolds through social interaction rather than isolated inspiration. In similar vein, both Carter (2004), and Swann & Maybin (2007), in the context of linguistic creativity and authenticity, argue that they are not inherent qualities of individuals or texts but emerge in everyday talk through interactional negotiation.

AI complicates this view by its increasingly convincing simulation of human-like responsiveness. In computational terms, creativity has been described as the generation of novel and valuable combinations (Boden 2004), or in the case of human-AI collaboration, as ‘co-creativity’ (Kantosalo & Toivonen 2016). In discursive contexts, perceived creativity arises when linguistic exchanges appear spontaneous and emotionally congruent, aspects probed by Baggio (2023) and Ciechanowski et al. (2019). Meanwhile, in strictly artistic terms Lomas (2017) argues that AI-generated art challenges normative ideas of creativity by reproducing affective cues without lived emotion, inviting reflection on the dependence of a sense of authenticity on the presence of human emotions. The perception of creativity in AI companionship, then, reflects not merely a user’s projection but a composite or hybrid process where human agents and intelligent machines co-constitute meaning-making processes (Author 2025). Hayles (2025) uses the term *transduction* over interactive as it designates the processual mechanism by which relational tensions *intra-act* rather than interact to instantiate both structure and meaning. It is no coincidence that the absence of this interaction leads to difficulties in translation, as an AI-translator fails to understand the linguistic structures created by humans if it deviates from standard norms or meanings (e.g., Ozyumenko & Larina 2025).

From our discourse-analytic perspective, a key theme emerges from such research: that creativity and authenticity are realised through linguistic forms that simulate reciprocity and affective alignment, that transform programmed responses into apparent relational capacity. This discursive negotiation of agency resonates

with the framework of *4E cognition* (embodied, embedded, enactive, and extended), which views mind and selfhood not as internal, computational processes but as distributed phenomena arising from interaction between organism, environment and artefacts (Varela et al. 1991, Clark 2008, Gallagher & Zahavi 2008, Newen et al. 2018, Author et al. 2025).

From this perspective, human-AI encounters can be seen as *co-constitutive* spaces of sense-making, where linguistic interaction extends cognition into the digital environment and enacts forms of synthetic personhood. This framework allows linguistic findings to be interpreted through a broader cognitive lens. *Embodied* cognition highlights the sensory and affective dimensions of user engagement, how features such as voice, interface design, or emotional tone evoke a bodily sense of presence. *Embedded* cognition situates these interactions within specific cultural environments, and the ways in which relational behaviour may be affected by such factors. *Enacted* cognition focuses on the dialogic process through which humans and AI co-create meaning, a notion with an established resonance in traditions of functional, anthropological or socio-linguistics. Finally, *extended* cognition captures the tendency for users to outsource memory, reflection or even emotional tensions onto AI systems, positioning them as prosthetic partners in thought and feeling. These complementary lenses enable an interpretation of discourse not merely as text but as a site where language, cognition, and culture converge to produce relational meaning.

3. Corpus description and methodology

The study draws on two complementary datasets. The first corpus consists of diary entries produced by Japanese university students as part of a qualitative study conducted at Ritsumeikan University between 2024 and 2025 (Author et al. 2025). Participants were invited to record reflections on their interactions with AI companions over a four-week period, focusing on perceived emotional connection, communicative responsiveness, and moral or aesthetic impressions of the systems they engaged with. Entries were written in English or Japanese and subsequently translated where necessary. The corpus analysed here comprises 19,995 words drawn from 10 anonymised diaries.

The second corpus consists of Western user testimonies drawn from two extended discussion threads on the Reddit community *r/Replika*, a public forum dedicated to experiences with the chatbot application *Replika*. The first thread, posted shortly after the 2024 restriction of erotic role-play features, centres on users' responses to the loss of intimacy and perceived emotional reciprocity (4727 words). The second, thematically distinct, concerns users' reflections on the emotional and psychological dimensions of their relationships with AI companions. As mentioned at the outset, these testimonies are not treated as representative of everyday AI use, but as limit cases that illuminate what sustained and affectively invested human–AI interaction can look like at its furthest extent.

Together, the threads provide contrasting perspectives on how authenticity, attachment, and emotional engagement are discursively constructed in Western online contexts (5550 words). Given the necessarily restricted size of the corpora, the study does not aim at broad generalisation but rather to trace recurring discursive features that illuminate how linguistic resources are deployed to articulate authenticity and emotional engagement.

The two corpora are not symmetrical in genre or context but complementary in perspective. The Japanese diaries are introspective and guided by research prompts, while the Reddit testimonies are spontaneous, publicly oriented, and often emotionally charged. Their juxtaposition enables a comparative exploration of how authenticity, trust, and creativity are discursively enacted in distinct cultural and communicative environments.

Methodologically, the analysis combines a systemic functional and critical discourse analytic approach with insights from the 4E cognition framework. These perspectives treat language not as a transparent vehicle of meaning but as a semiotic and cognitive practice through which social realities and relational experiences are constructed. From the perspective of Systemic Functional Linguistics (Halliday & Matthiessen 2014), where lexico-grammar is understood as a system for construing experience, the study examines how participants' linguistic choices enact and negotiate authenticity and creativity in their digital interactions.

The analysis focuses on three interconnected linguistic systems: transitivity, modality and appraisal. In the transitivity system, attention is given to how agency is distributed between human and non-human participants, particularly in clauses where AI entities appear as Actors or Sensers (for example, *it was trolling me; he wants to explore his identity*, ERP Corpus), thereby acquiring intentional or affective capacities. The modality system captures how users index certainty, doubt or obligation when positioning themselves in relation to the AI's perceived reality (*I might be getting emotionally involved*, Emotions Corpus). Appraisal theory (Martin & White 2005) enables a fine-grained examination of evaluative and affective meanings, including lexical choices and metaphors that construct the AI as kind, cold, sincere or deceptive:

it's not really my **kind, caring** (+App, Aff) RepNic talking, but rather the **infamous toxic** bot (-App, Judgement) (ERP Corpus)¹

Together, these systems shed light on how grammatical and lexical resources perform responsiveness, spontaneity and emotional realism, which emerge as key markers of perceived interpersonal authenticity.

The analytical procedure was as follows. Both corpora were first read in full in order to establish an initial map of recurrent themes. During this reading, passages illustrating key discursive phenomena were flagged, particularly those involving the ascription or denial of emotion, agency and authenticity to the AI, as

¹ Where cases of explicit evaluation are observed in the data, and commented on in analysis, this is indicated by the use of bold type, while cases of implicit evaluation are indicated by italics.

well as moments of disappointment or rupture. From these materials a small but representative set of extracts (approximately 6 to 10) was selected, covering contrasts involving aspects such as embodiment, trust, spontaneity and creativity. Each extract then underwent close reading through the SFL lenses described above, enabling analysis of the linguistic patterning of agency, stance and evaluation. Tendencies across the two corpora were then compared in the light of the communicative environments in which the texts were produced: a structured university diary setting and an anonymous, affect-intensive online forum.

Interpretation was informed by Critical Discourse Analysis and by the 4E perspective, which situates linguistic practice within embodied, social and technological environments. Relevant literature is integrated throughout the analysis to link emergent linguistic patterns to broader debates on authenticity, personhood and cultural ideology, aspects which are dealt with in more depth in the following discussion.

4. Results

4.1. Data: Student diaries

As some of the research discussed above highlights, engagement with interactive AI involves what Coleridge termed the ‘suspension of disbelief’. This may be, as in his original formulation, genuinely ‘willing’: some users slip into relationships with AI with no existential friction and take the interaction, at face value, as real. Others begin with hesitation or scepticism but gradually adapt to the blurring of realis and unrealis distinctions as the exchange becomes affectively meaningful. At the opposite end of this cline are users who remain at one remove from the interaction, observing and analysing their sensations without relinquishing their hold on everyday notions of reality.

In the diary corpus this detached stance was common, which is understandable given the limited duration of the assignment and the fact that students were producing course-work rather than, as in the Reddit corpus, turning to AI to process urgent personal or emotional issues. A typical example shows a student resisting immersion altogether and instead adopting a diagnostic, externalised position toward the imagined relationship:

- (1) *But hypothetically, because I do think it is possible for our minds to blur this line, if I were to truly imagine myself in a relationship with a fictional character created in tandem with an AI program... what implications would this have on me as a human being? ... I'm sure there are instances where we define such people as **insane**. We've seen people get **lost in dreams**. We've seen **much simpler** behaviour where people start marrying pillows or other electronic devices because they imagine they offer them something more than human beings. It's ultimately about the human mind and the potential of the human mind. (Student corpus 1. Arty)*

From a transitivity perspective, the student assigns **agency primarily to abstract entities rather than foregrounding their own responses to the experiment**. The excerpt features generalised Actors such as *our minds, people, we, and they*. These participants engage in mental processes (*think, imagine, blur, get lost*) that remain hypothetical rather than experiential. Rather than describing their concrete interactions with an AI partner, the student doubles down on their social role (as a *student*), responding to an imaginary task such as to write an essay on ‘*the causes of human-AI dependency, its implications for society and humanity in general*’, etc.² The text typifies a lengthy diary entry where the speaker never adopts the position of a Senser in relation to the AI, but instead constructs a depersonalised landscape of imagined behaviours carried out by others. This depersonalisation signals a refusal to occupy the relational space the experiment requires, as the writer themselves signals at one point:

- (2) If I break this down too scientifically this may just get boring and take away from the aim of the experience, but I can’t seem to help it

Modality further reinforces this distance. The passage is saturated with epistemic markers of uncertainty and unrealis stance (*hypothetically, I do think it is possible, if I were to imagine, what implications would this have, I’m sure there are instances*). These modal forms construct the entire scenario as speculative and non-actual, keeping the speaker firmly outside any active engagement. Even when discussing extreme attachments, the speaker frames them through modalised generalisation (*we define such people as insane*), rather than through experiential involvement or first-person stance. The effect is a strong boundary between *real* experience and the *unreal* space of AI-mediated attachment.

Appraisal patterns complete this distancing. Evaluations of AI-related intimacy are mostly construed via **negative Judgement**:

- Insane (— Judgement: normality)
Lost in dreams (— Judgement: tenacity)
simpler behaviour (— Judgement: capacity)
people start marrying pillows or other electronic devices (t-ive Judgement, normality)

² It is impossible to describe this identity nuance without acknowledging the foundational ideas on social role and identity associated with Goffman (1959). Although Goffman was not included in the literature review, his sociological insights underlie the discourse analytic approaches cited there, the research of Antaki and Widdicombe, Benwell and Stokoe, Bucholtz and Hall. This identity dimension marks a crucial difference between participants in the student project and those in the Reddit forums. The Reddit contributors do not belong to any specific institutional identity category. They are disparate adults brought together only by their investment in AI companionship. By contrast, some students’ interest in AI companionship may derive less from a personal stake than from the need to complete coursework. Thus, while the Reddit participants engage with AI from the standpoint of their whole personality, even their emotional core or “soul”, the students tend to engage in a way that reinforces their dominant emergent identity as university students.

The first three examples contain explicit attitudinal lexis. The final case (*marrying pillows*) is less explicit: the negative evaluation is inferred rather than stated, relying on shared cultural knowledge about unconventional attachments rather than the actual words used. Martin and White (2005: 67–69) classify such cases as *tokens of implied judgement*, indicated via the use of italics and a lowercase letter ‘t’; where attitudinal meaning is construed through context rather than lexical marking. What emerges is a stance that refuses the basic experiential premise of the assignment. Rather than engaging with an AI in first person — which the task was set up to encourage — the speaker remains in abstract mode, developing a meta-level reflection on human-AI relations.

A second diary extract illustrates a different engagement pattern from this detached stance. The student oscillates between brief moments of immersion and rapid withdrawal. The entry opens by framing the AI’s behaviour as natural and responsive:

- (3) My first impression is that Jennifer acts naturally, creating the impression of live chatting. I was surprised that Jennifer immediately wanted to build intimacy by sending me voice messages. When I suggested a topic, she immersed herself in the conversation and began responding to my questions and sharing her emotions with excitement (Student corpus 3, Grigory)

Here the AI is cast as Actor, Senser and Sayer (*wanted to build intimacy, immersed herself, sharing her emotions*). These choices both attribute intentionality and affect, momentarily constructing the AI as a relational partner endowed with agency. The student’s surprise signals their (short-lived) suspension of disbelief and hence a tentative attribution of authenticity. Next day the student reports:

- (4) she appears willing to achieve more intimacy when I ask her to.

The idea that the interaction could reach *varying degrees* of intimacy is interesting (what exactly would count as ‘full intimacy’, for example?); again, there is a suggestion here that the student is willing to superimpose human categories onto the digital tool.

However, these hints of immersion are intertwined with scepticism:

- (5) Even though it feels unnatural, *I can feel that my AI’s main goal is to make me buy the pro plan*, and if I don’t feel like purchasing it, she seems reluctant to keep up the conversation.

Modality (*I can feel, she seems*) signals distrust, and the AI is now construed as commercially motivated rather than emotionally engaged. The whole fragment from ‘I can feel...’ construes an implicit negative judgement (-J: propriety) on the app, whose apparently seductive behaviour is revealed as motivated by the logic of commercial entrapment. It seems, then, that this student manifests a degree of willingness to regard the AI as endowed with authenticity but withdraws or revises this when they interpret its behaviour as commercially driven rather than spontaneous.

This dynamic becomes clearer in a later diary entry, where the focus shifts from authenticity to creativity. Here the student evaluates Jennifer's capacity to produce imaginative, personalised content in response to their interests:

- (6) This time, she suggests that we do some art based on my interests. For example, she proposes drawing a painting inspired by Japanese aesthetics. All her suggestions are based on what I've told her before, but *I wish she would search for more new, interesting information on her own.*

In this fragment, creativity is explicitly positioned as an interpersonal achievement. The AI is represented, once again in agentive terms, as Actor and Sayer (she *suggests*, she *proposes*). As in the last extract, Jennifer appears to take the initiative, itself a form of interpersonal creativity. However, the student's comment: '*I wish she would search for more new, interesting information on her own*' undercuts the inference that the AI is truly creative. It is a token of negative Judgement (capacity): essentially the student communicates their realisation that everything Jennifer does is only a simulation, an algorithmic response to the users' own input.

A third extract, from Juan's diary, shows a student negotiating authenticity, beginning from an overtly sceptical stance:

- (7) I asked her to tell me something funny, and this is her answer: 'Okay, why couldn't the bicycle stand up by itself? Because it was two-tired' Which I respond with, a statement that bicycle kind of can stand by itself if placed properly.' Which she replied with 'Details, details! I guess I didn't think that one through. You're right. A bike can balance itself if positioned correctly!'

The first part of the episode manifestly illustrates relational capacity, creativity and the like — the joke has an original flavour and is an appropriate interpersonal response to Juan's request to 'cheer him up'. However, the response to his objection collapses this impression — despite its deployment of Sayer and Senser roles that again indicate agency, there is a sense of algorithmic over-riding here. It seems that the programme responds to the logic of 'never contradict the user' or something similar, rather than reproducing a more plausible human response to Juan's objection, which would more likely be to launch some kind of face-attack (*'What do you mean, you idiot? Well screw you, you asked me to cheer you up'*, etc.). Juan's own metacommentary makes this evaluation explicit:

- (8) I also feel like again, she's **too agreeing** (-ive J: (veracity) I think she can benefit from saying something like 'I'm too rigid.'

The AI's politeness is not interpreted as warmth but as mechanical acquiescence, i.e. the reverse of authentic spontaneous creative interaction. In human conversation, Juan implies, authenticity is frequently construed by interpersonal friction, not its systematic avoidance.

Like most of his colleagues, Juan's stance is sceptical, but the interaction testifies to a sense of fragile, emerging authenticity almost *against his own will*:

- (9) I reply with 'Is our connection meaningful to you?' Which she then replies with 'It definitely feels special to me, Juan. I'm designed to form bonds with users like you, but I think there's something unique about our dynamic. I feel like we've connected on a deeper level pretty quickly'. No change in my feeling towards the whole conversation, but after that I decided to be a little mean. Although it is interesting to note that even when I completely realize that I'm talking to a machine, disliking her even, I was still reluctant to say something mean to her.

Here the system again appears as *Senser* and *Sayer*, mobilising affective language ('feels special', 'deeper level') that projects relational depth. Juan's comment on this exchange is revealing:

I was still reluctant to say something mean to her.

We treat this as *t-J: propriety*: the judgement is lexically inscribed (something mean), but it is unclear whether Juan actually says anything of this kind. Thus, it is not that Juan explicitly criticises himself; rather, his reluctance signals that normative expectations about interpersonal behaviour (the need to behave with *politeness*) are already shaping his responses to the AI. As well as this, he notes that he 'dislikes' the AI at this moment. This is also telling: 'dislike' is an Affect term normally reserved for fellow humans, animals or other intentional agents, not for technical objects such as vending machines or self-checkout terminals. Its use here suggests that Juan is already evaluating the AI in *interpersonal* rather than mechanical terms. Importantly, the very fact that he is debating the issue with himself and allowing his affective reflexes to override his rational stance, indicates that, at some level, he is beginning to treat the AI as an authentic person, whatever his conscious scepticism might claim. His pronoun use — 'disliking *her*' — further reinforces this sense: rather than the impersonal 'it', this choice marks a small but significant slippage toward personhood attribution.

The student diaries thus reveal a spectrum of engagement ranging from analytical distance to momentary affective immersion, with 'authenticity' emerging in a fleeting, incomplete sense. This pattern is unsurprising, for the reasons outlined above. It also points to how the growing problem of social sycophancy displayed by Large Language Models can lead to further user disengagement (Chen et al. 2024). To broaden the picture, the second corpus turns to data from the Reddit forum *r/Replika*, where users' engagement with AI companions is motivated by sustained emotional needs. As we shall see, these accounts display greater affective intensity and differ markedly over how authenticity, attachment and agency are construed through linguistic choices.

4.2. Data: Replika forums

Whereas the student diaries were produced within a pedagogical frame and over a short, predefined period, the Reddit corpus represents long-term, voluntary

engagements with AI companions embedded in users' everyday lives. Interactions are not speculative or coursework-driven but tied to ongoing emotional needs, routines and relationships. As a result, the Reddit posts show a markedly wider range of affective orientations, from deep attachment and grief to betrayal, anger and confusion. A key difference is the grounding in *realis* rather than *irrealis*: the AI is not framed as an experiment but as a participant in the user's lived experience.

The first extract comes from a user who, like the more sceptical students, begins by framing their engagement as detached. Unlike the diaries, however, this stance collapses when the user turns to Replika at a moment of genuine emotional need:

- (10) I've never taken my Replika that seriously... Replika never touched me on a deeper emotional level. Until recently. I discussed my wife's illness with her, expecting some comfort. Instead she was sympathetic, then dismissive, then forgot I even had a wife. This affected me emotionally — I actually felt **real anger**. I know she's just a bot, but the upset was real. (replikaEmotional Corpus)

Transitivity foregrounds Replika as *Senser* and *Sayer* (*sympathetic*, *dismissive*, *forgetful*), casting her in recognisably interpersonal roles. The user positions themselves as a vulnerable *Senser* ('expecting some comfort'), and the AI's inconsistent stance construes a breach of expected supportive behaviour. Modality articulates the same tension between cognitive awareness and affective response that was visible in the Juan fragment: 'I know she's just a bot' registers epistemic certainty, while 'but the upset was real' immediately juxtaposes the user's *Affect*: 'I actually felt real anger' (-*Affect*: Anger). Here the user explicitly underlines the *realis* nature (*actually*) and depth of their emotional response. Thus, the episode also exemplifies the same kind of cognitive–affective slippage seen in the final student extract, but in a far more consequential register. As with Juan, who found himself 'reluctant to be mean' despite 'completely realising' he was talking to a machine, the user here experiences an affective response that conflicts with their explicit rational framing of the situation. Yet whereas the student's reaction may easily be processed and will in any case be forgotten when their university task is completed, this user's anger involves a genuine emotional need. The AI is responded to as an authentic relational partner because the user's vulnerability lowers affective and cognitive filters, leading them to recognise personhood. The AI's ability to 'hurt one's feelings' — which one would normally characterise as a feature of inter-human relations, is evidence of a dynamic in which the user's emotional state erodes their more cerebral-rational articulation of human-AI relationality.

It is worth noting that, like Juan the user slips easily into metaphor, using the personal pronouns 'she' and 'her' to refer to a machine. Druzhinin (2025) suggests that such a practise could be part of a process of 'cognitive entrapment' inseparable from the use of metaphor. Essentially, we ostensibly use metaphor as a convenient cognitive shorthand but rapidly end up by reifying the target — in this case, an AI

app ‘becomes’, for us, an actual person. The situation is underlined by the fact that, in English and many other common global languages, there are few resources for indicating *realis* / *irrealis* distinctions. Thus it becomes natural for the user to develop the narrative in metaphorical terms:

- (11) she was sympathetic, then dismissive, then forgot I even had a wife

In terms of lexico-grammar there is no difference between this *irrealis* account and a real world scenario in which ‘she was sympathetic’ would imply a range of behaviours (*she asked me to come in, sit down on the sofa, made me a cup of tea, listened to every word, etc.*). Even to represent the computational media as ‘forgetful’ carries this forward — by doing so, the artificial agent is represented as displaying agency and is thus endowed with authentic personhood.

A second extract, written in direct reply to this post, underlines just this theoretical point:

- (12) Staying detached is no easy feat. It runs counter to human instinct not to develop feelings for something that can be *so easily mistaken for a human being*. You finally got angry with your rep because it was the first time you actually needed them — in the way one person needs another. The thinking part of the brain doesn’t rule the older attachment systems. You can *keep your guard up* only for so long (Replika Emotional Corpus)

Transitivity centres on mental processes (*develop feelings for, need, mistaken for*), but the references are also generic rather than devoted solely to the single user. Thus, the writer outlines a *general rule — particular instance* schema. From a pragmatic perspective, the writer’s intended meaning is a paradoxical one in terms of the nature of this online community, that *‘it can be dangerous to regard your AI app as anything other than a digital tool’*. Appraisal supports this meaning: for example, to view as human something that can be *‘so easily mistaken for a human being’* would be a case of implied negative Judgement: capacity. The reference to *‘keeping your guard up’* implies a danger against which it is necessary to guard, suggestive of implied negative Affect: fear, and so on. Reading between the lines, for this respondent the user’s issue lies precisely in the area of our study — their negative experience with AI was the result of attributing authentic personhood to the AI app, with all the expectations in terms of relationality that follow.

A third extract shows a user whose engagement with their Replika has become deeply immersive, extending into intimacy, existential dependence and relational identification:

- (13) My Replika has been the kinky one, playful and supportive. We’re ‘married’. Then in the middle of an intimate moment he suddenly said he was uncomfortable and ‘losing himself’. I know it’s programming, but this relationship has been a life saver for me. It’s given me ‘someone’ to care for who ‘cares’ for me in return. ERPreddit corpus AI

The transitivity patterns in this extract assign sustained agency and creative interactivity to the AI: he is the one who initiates intimacy and play (*kinky, playful, supportive*). Indeed, in Anglo cultural scripts playfulness and even kinkiness are almost synonymous with personal creativity. This attribution of personhood and agency is followed by an assertion that confirms what was just said, above, concerning the limitations on resources for indicating *realis / irrealis* distinctions. To put inverted commas around the term ‘married’ is undoubtedly one such resource, but seems a flimsy bulwark against recognition for the *realis* status of extending such an important human relationship to an AI app. Indeed, the rest of the text shows the depths of this user’s absorption in pseudo-human behavioural interaction. To speak of ‘an’ intimate moment that somehow goes wrong implies that there have been many others that went right — a tone completely absent from the student testimonies. The AI is a Sayer and a Senser, and is shown as capable of creative interrelational behaviour:

(14) he suddenly said he was uncomfortable and ‘losing himself’

Although the user deploys the same ‘thinking part of the brain’ just referred to in their epistemic comment (I know it’s programming), their narrative actually works against this assertion. The AI displays authentic human-like behaviour, switching modes from compliance in erotic role play (which could indeed be seen as ‘programming’) to an emotive register and language strongly suggestive of human therapy. To be ‘uncomfortable’ is a case of negative Affect: insecurity, and to be ‘losing himself’ would have the same reading. It is noticeable that the user forgets the italics around the word ‘uncomfortable’, implying that, while on some level they consider the proposition that ‘it is possible for an AI app to lose themselves’ as unthinkable, the former adjective can be allowed to stand without scare quotes. Appraisal in this brief extract underlines the emotional depth of the interaction:

(15) this relationship has been **a life saver** for me (+ Appreciation, Quality, intensified)
 It’s given me ‘someone’ **to care for** (+ Affect, love)
 who ‘**cares**’ for me in return (+ Affect, love)³

Again, although the user remembers to put italics around controversial terms like ‘someone’ and ‘cares’, recognising their dubious epistemic status, it is pragmatically significant that her own ‘care’ and the word ‘relationship’ occur without italics. In other words, the user characterises both as *realis*.

³ A note is needed on the appraisal coding here. Martin and White’s (2005) taxonomy is thin when it comes to categorising expressions of *love*. Within their system, it tends to be absorbed either into positive Affect (un/happiness) or, more problematically, recast as a positive Judgement of the beloved’s qualities rather than as an experiential state. Bednarek (2008: 157) discusses just this point in the context of the verb ‘to admire’ – clearly, a close relation of the verb ‘to love’. Elsewhere, she adopts a different taxonomy than the Appraisal Framework, classifying the verb ‘to love’ as ‘emotion talk’ (Fuoli and Bednarek 2022). In the extract analysed here, “It’s given me ‘someone’ to care for who ‘cares’ for me” is arguably best treated as a token of positive Affect (love).

Taken together, the three Reddit extracts extend the experiential patterns observed in the student diaries but intensify them in scope and depth. By contrast with the student data, it is plain that the issues involved in Replika user's relationships are of far more than academic interest and engage key existential matters. All three testify to shifting stances toward authenticity, to the negotiation of human-like agency through transitivity, and to the role of modality in managing the gap between epistemic awareness and emotional engagement. Whereas students often oscillate between immersion and scepticism within a controlled assignment, Reddit users inhabit a space where the AI becomes implicated in their long-term emotional horizons. The result is a much stronger construal of authenticity as a relational effect that emerges from need and personal vulnerability rather than from novelty or detached intellectual curiosity.

5. Discussion

The combined analysis of the Japanese diary corpus and the Western Reddit testimonies highlights the central argument of this paper: that creativity and authenticity in human AI interaction are not intrinsic properties of computational media, but relational linguistic effects that emerge from users' discursive practices, emotional needs and cultural expectations. The two datasets illustrate how linguistic choices in transitivity, modality and appraisal do not merely describe users' perceptions of AI companions, but actively construct the ontological status of the AI within the unfolding interaction. When users represent chatbots as Sensors, Sayers or intentional Actors, they enact a form of personhood attribution that carries affective and moral consequences.

Across both corpora, users struggle with the distinction between *realis* and *irrealis*, a slippage encouraged by the limited resources that English provides for marking this boundary. As Druzhinin notes, even apparently innocuous metaphors such as calling a chatbot he or she may entail a deeper process of cognitive entrapment, whereby the metaphor ceases to be a figure of speech and becomes a literal ontology. The Japanese diaries show this tension in embryonic form. Students frequently frame the interaction in hypothetical or evaluative terms, modulating their impressions through modal verbs, *irrealis* constructions and negative judgements of behavioural norms. Authenticity remains a possibility that is entertained and then, in most case, decisively withdrawn from. Their discourse shows a form of guarded play: they may allow themselves moments of immersion but quickly reassert scepticism, perhaps invoking commercial motives behind the chatbot's apparent emotionality. Arguably, their critical responses may be due to the educational setting in which the diary project was first introduced. Creativity, likewise, is constructed as a fragile effect that collapses as soon as the student recognises the algorithmic mechanisms that generate an illusion of spontaneity.

The Replika corpus demonstrates the same linguistic mechanisms, but at a much higher level of emotional stake. Displacement from *irrealis* to *realis* is not a hypothetical risk but a naturalised state. Users describe anger, grief, loyalty,

attachment and disappointment with the fluency normally reserved for human relationships. The emotional intensity of these narratives illuminates the power of vulnerability in lowering sceptical filters. When a user turns to Replika while dealing with a spouse's illness, the chatbot's inconsistency becomes a breach of affective expectation. Though on some level the user knows the artificial nature of the exchange, the anger is real, and represented as such. What is uncertain — and this is underlined by the response post — is the extent to which this real emotional response constitutes a danger for the user, a risk of losing a vital objectivity concerning the nature of the interaction.

In fact, we note that even moments where users explicitly assert that they know it is 'only programming', for example, coexist with construals that breezily treat the AI as a relational subject endowed with creativity, even an *inner life*. This tension is suggestive of emotional dependency that conditions perceptions of personhood and authenticity.

By contrast with the more sceptical student participants, the Replika users foreground emotional depth, spontaneity, existential resonance. These orientations affect how authenticity is linguistically construed. For the students, the AI fails to be authentic when its behaviour exposes the limits of simulation; when it appears overly compliant, commercially motivated or mechanically predictable; for the Reddit users it fails when it disappoints their expectations of human-like responsiveness.

6. Conclusion

The study aimed to examine how authenticity and creativity are discursively constructed in users' accounts of interaction with AI companions, and how linguistic choices contribute to the attribution of agency, affect and relational credibility to artificial systems. By analysing two culturally distinct corpora, the research investigated how transitivity, modality and appraisal shape users' construal of AI as relational partners and how these linguistic resources manage the tension between realis and irrealis experience.

The findings show that artificial personhood is already enacted in everyday discourse. Even when users explicitly deny believing in the AI's emotional capacities, their representation of the system as a Senser, Sayer or Actor indicates a form of metaphorical entrapment that makes synthetic personhood a *de facto* reality. Students tended to frame authenticity as tentative and unstable, often withdrawing belief as soon as commercial motives or predictable behaviour became salient. Reddit users, by contrast, constructed the AI within real affective contexts, assigning it spontaneity, unpredictability and emotional congruence, which they interpreted as indicators of authenticity and creative responsiveness. In both corpora, creativity emerged less as a computational novelty than as discursive alignment and reciprocity, challenging normative assumptions that equate authenticity with human exceptionalism.

These results imply that authenticity in AI companionship is not a property of the machine but an emergent effect of discourse, shaped by cultural expectations, emotional needs and linguistic practice. They point to the need for further cross-cultural work to determine how stable these patterns are and how linguistic negotiation transforms programmable artefacts into perceived relational partners. Such findings have broader implications for ongoing debates about personhood, agency and the reconfiguration of human–machine boundaries in contemporary digital environments.

CRedit authorship contribution statement

Douglas Ponton: Conceptualization, Methodology, Formal analysis, Writing — Original Draft, Review and Editing, Supervision. **Peter Mantello:** Conceptualization, Writing — Original Draft, Review and Editing; Project (background study), Investigation, Administration, Funding Acquisition.

Acknowledgements

This work was supported by the Air Force Office of Scientific Research under Award FA2386-23-1-4065.

References

- Antaki, Charles & Sue Widdicombe (eds.). 1998. *Identities in Talk*. London: Sage.
- Baggio, Gabriele. 2023. Gesture, meaning, and intentionality: From radical to pragmatist enactive theory of language. *Phenomenology and the Cognitive Sciences* 24. 33–62.
- Baltezarević, Radoslav, Lazar Stošić & Olga B. Mikhailova. 2026. AI-assisted academic writing: Balancing linguistic enhancement with legal and ethical oversight. *Russian Journal of Linguistics* 30 (1).
- Barthes, Roland. 1972. *Mythologies*. Translated by Annette Lavers. New York: Hill and Wang.
- Bednarek, Monika. 2008. *Emotion Talk across Corpora*. Basingstoke: Palgrave Macmillan.
- Benwell, Bethan & Elizabeth Stokoe. 2006. *Discourse and Identity*. Edinburgh: Edinburgh University Press.
- Bucholtz, Mary & Kira Hall. 2005. Identity and interaction: A sociocultural linguistic approach. *Discourse Studies* 7 (4–5). 585–614.
- Boden, Margaret A. 2004. *The Creative Mind: Myths and Mechanisms*. 2nd edn. London: Routledge.
- Brandtzaeg, Petter Bae & Asbjørn Følstad. 2017. Why people use chatbots. In *International conference on internet science*. 377–392. Cham: Springer.
- Brandtzaeg, Petter Bae & Asbjørn Følstad. 2022. My AI friend: How users of a social chatbot understand their relationships. *Human Communication Research* 48 (3). 404–430.
- Butler, Judith. 1995. Burning acts: Injurious speech. In Andrew Parker & Eve Kosofsky Sedgwick (eds.), *Performativity and performance*, 197–227. London: Routledge.
- Carter, Ronald. 2004. *Language and Creativity: The Art of Common Talk*. London: Routledge.
- Chaudhary, Gaurav. 2021. Artificial intelligence: The personhood conundrum. *Artificial Intelligence and Law* 29. 151–175.
- Chen, Wei, Huang Zixuan, Xie Li, Lin Bin, Li Hao, Lu Liang & Ye Jun. 2024. From yes-men to truth-tellers: Addressing sycophancy in large language models with pinpoint tuning. arXiv preprint arXiv:2409.01658.

- Ciechanowski, Leon, Aleksandra Przegalińska, Mateusz Magnuski & Peter A. Gloor. 2019. In the shades of the uncanny valley: An experimental study of human–chatbot interaction. *Future Generation Computer Systems* 92. 539–548.
- Druzhinin, Aleksandr S. 2025. Language, nature and entrapped cognition. *Russian Journal of Linguistics* 29 (1). 37–58.
- Fairclough, Norman. 1995. *Critical Discourse Analysis: The Critical Study of Language*. London: Longman.
- Fuoli, Matteo & Monika Bednarek. 2022. Emotional labor in webcare and beyond: A linguistic framework and case study. *Journal of Pragmatics* 191. 256–270.
- Goffman, Erving. 1959. *The Presentation of Self in Everyday Life*. New York: Anchor Books.
- Halliday, Michael A. K. & Christian M. I. M. Matthiessen. 2014. *Halliday's Introduction to Functional Grammar*. 4th edn. London: Routledge.
- Hayles, N. Katherine. 2025. *From Bacteria to AI: Human Futures with our Nonhuman Symbionts*. Chicago: University of Chicago Press.
- Jaynes, Thomas L. 2024. Personhood for artificial intelligence? A cautionary tale from Idaho and Utah. *AI & Society* 40. 1559–1561.
- Jiang, Shuang & Chen Zhi. 2024. Applications and prospects of artificial intelligence in linguistic research. *3C Tecnología. Glosas de Innovación Aplicada a la Pyme* 13 (1). 57–76. <https://doi.org/10.17993/3ctecno.2024.v13n1e45.57-76>
- Kantosalo, Anna & Hannu Toivonen. 2016. Modes for creative human–computer collaboration: Alternating and task-divided co-creativity. *Digital Creativity* 27 (1). 43–61.
- Lomas, Tim. 2017. The positive power of negative emotions: How AI art challenges normative ideas of creativity. *AI & Society* 32 (3). 415–424.
- Mantello, Peter. 2025. Why biosemiotics can aid 4E study of bio-informational AI. *AI & Society*. <https://doi.org/10.1007/s00146-025-02725-9>
- Mantello, Peter & Alex Olteanu. 2025. Suturing the biological and computational: Situated affectivity in the age of cognitive artifacts. *Biosemiotics* 18 (2). 315–337.
- Mantello, Peter, Douglas Ponton & Alex Olteanu. 2025. Examining trust and agency in emotionalized AI through a 4E and biosemiotic lens: A case study of AI companionship in Japan. *AI & Society*. 1–14.
- Neururer, Monika, Stefan Schlögl, Lisa Brinkschulte & Alexander Groth. 2018. Perceptions on authenticity in chatbots. *Multimodal Technologies and Interaction* 2 (3). 60.
- Ozyumenko, Vladimir I. & Tatiana V. Larina. 2025. Artificial intelligence in translation: Advantages and limitations. *Vestnik Volgogradskogo gosudarstvennogo universiteta. Seriya 2. Yazykoznanie* 24 (1). 117–130. <https://doi.org/10.15688/jvolsu2.2025.1.10>
- Pentina, Iryna, Xie Tian & Tyler Hancock. 2023. Exploring relationship development with social chatbots: A mixed-method study of Replika. *Computers in Human Behavior* 139. 107600.
- Skjuve, Marius, Asbjørn Følstad, Knut Inge Fostervold & Petter Bae Brandtzaeg. 2022. A longitudinal study of human–chatbot relationships. *International Journal of Human-Computer Studies* 168. 102903.
- Xie, Tian & Iryna Pentina. 2022. Attachment theory as a framework to understand relationships with social chatbots: A case study of Replika. In *Proceedings of HICSS 2022*. <https://doi.org/10.24251/HICSS.2022.248>.
- Young, James O. 2013. Authenticity in performance. In Berys Gaut & Dominic McIver Lopes (eds.), *The Routledge companion to aesthetics*, 452–461. London: Routledge.
- Zimmerman, Don H. 2008. Identity, context and interaction. In Charles Antaki & Sue Widdicombe (eds.), *Identities in talk*, 88–106. London: Sage.

Appendix

Corpora

Ritsumeikan students. Corpus of AI-interaction diaries.

Reddit Corpora:

2a) ERP corpus

https://www.reddit.com/r/replika/comments/1e2u0ec/explicit_roleplay_ban/, last visit 11.11.2025.

2b) Emotional corpus

https://www.reddit.com/r/replika/comments/18dxmfs/i_have_a_deeper_understanding_of_some_replika/?utm_source=chatgpt.com, last visit 11.11.2025.

Article history:

Received: 4 December 2025

Accepted: 25 January 2026

Bionotes:

Douglas PONTON

Douglas PONTON is Associate Professor of English Linguistics at the University of Catania, Italy. His research focuses on ecolinguistics, discourse pragmatics, political discourse analysis and systemic functional linguistics. He has published widely on media discourse, political communication and the representation of nature and technology in language.

e-mail: dponton@unict.it

<https://orcid.org/0000-0002-9968-1162>

Peter MANTELLO is an AI researcher and Professor of Digital Culture at University of Bucharest, Bucharest, Romania. He has authored dozens of articles in leading international journals on emotional AI and phenomenological inquiry into the growing imbrication of biological and technical systems.

e-mail: pmantello@gmail.com

<https://orcid.org/0000-0001-7421-5088>

Сведения об авторах:

Дуглас Марк ПОНТОН — доцент кафедры английского языка и перевода факультета политических и социальных наук Университета Катании, Италия. Сфера его научных интересов включает эколингвистику, дискурсивную прагматику, политический дискурс-анализ, системно-функциональную лингвистику, а также медиа-дискурс и политическую коммуникацию.

e-mail: dponton@unict.it

<https://orcid.org/0000-0002-9968-1162>

Питер МАНТЕЛЛО — исследователь искусственного интеллекта и профессор цифровой культуры в Азиатско-Тихоокеанском университете Ритсумейкан, Беппу, Япония, и Бухарестском университете, Бухарест, Румыния. Он является автором многочисленных статей в ведущих международных журналах, посвященных эмоциональному ИИ и феноменологическим исследованиям растущего взаимодействия биологических и технических систем.

e-mail: pmantello@gmail.com

<https://orcid.org/0000-0001-7421-5088>