

HUMANS AS NATURAL-BORN CYBORGS: SCRUTINISING AI'S NARRATIVE INTELLIGENCE WITHIN THE 5E COGNITION FRAMEWORK

Inna Livytska,
Justus Liebig University
Inna.Livytska@anglistik.uni-giessen.de

Abstract

Recent advances in large language models (LLMs) and their ability to generate content align with both perspectives: AI as a powerful assistant and AI as a potential challenge to human cognition. However, these advancements expose the persistent limitations of AI compared to humans. One such limitation, as demonstrated by the *folk narrative hypothesis* (Hutto 2008), is the uniquely human ability to engage in storytelling, a fundamental and ancient mechanism for memory, information storage, identity formation, and world-making or world-disruption. The ability of AI to generate coherent texts is often mistaken for its intelligence, raising false

claims about AI self-consciousness and sentience. This paper argues that human narrative intelligence as a drive for sense-making might turn imperfect AI text generation into a meaningful construct despite cognitive dissonance in the communication. By analysing the technical prompts used by ChatGPT-4 to generate narrative text and its understanding of key text production categories (coherence, fluency, complexity, etc.) within the 5E Cognition framework (Stilwell & Harman 2019), I will attempt to show how AI-generated narrative structures acquire full and meaningful interpretation when perceived by a human. The generative ability of AI which emerges in co-communication with a human cannot stand for its intelligence, since it reflects human perception and our intrinsic ability to hypothesise and engage with communication agents.

Keywords: narratology, AI, ChatGPT, narrative intelligence, 5E cognition, co-narration, communication.

Introduction

When Andy Clark (2003) praised humans for being “natural-born cyborgs” due to our inherited ability to use the evolutionary mechanism of cognitive extension to interact with our surrounding *umwelt*, he was primarily focusing on the utilitarian role of computers and technology in preserving humanity’s exceptional status. This optimistic view of machines as tools and helpers designed to serve human purposes has always been tempered by an inconvenient yet inevitable concern: the possibility of encountering entities far more intelligent than ourselves. Recent advances of LLMs and the inconceivable admiration for their text generative ability to produce coherent texts seem to have brought these fears to our current days.

Indeed, fears of being outmastered and outcompeted by its creation are voiced today by the scientists who spent their entire lives working on AI. One of the persuasive examples is the 2024 Nobel Prize in Physics. Last year’s awardees, the U.S. scientist John Hopfield and British-Canadian Geoffrey Hinton, while being credited for laying the foundation work which revolutionised machine learning, both expressed their caution regarding the future use of their discoveries.

“We have no experience of what it is like to have things smarter than us”, Hinton said over in his phone call to the Nobel press conference. “It’s going to be wonderful in many respects, in areas like healthcare”, adding, “But we also have to worry about a number of possible bad consequences,

particularly the threat of these things getting out of control” (Pollard & Ahlander 2024).

The other 2024 Nobel Prize holder Professor Hopfield, a creator of an associative memory which can store and reconstruct images and other types of patterns in data, in his turn, stated: “When you get systems that are rich enough in complexity and size, they can have properties which you can’t possibly intuit from the elementary particles you put in there, you have to say that system contains some physics” (Pollard & Ahlander 2024). When referring to the impossibility of intuitively predicting the properties of this new system, we should probably see the sign of their autonomy, and thus, their agency, which is not yet fully discovered. The possibility of that stands at odds with the fears of AI and the risks for a new digital era of uncontrollable generative AI.

Bearing in mind these fears and partly sharing them, this paper takes a bold alternative perspective. It suggests that despite considerable progress of generative AI to imitate the formal functions of human linguistic competence, this process cannot be identical with human narrative performance from a purely biological evolutionary perspective. One of the primary reasons for my argument is the emotion generation capacity in humans, closely linked to narrative semantics and linguistic material, accelerating or inhibiting memorisation per se. Since LLMs are trained on vast amounts of linguistic data, their performance results from their response to the language in the prompt which involves generating more language. In other words, language-centred large models (LLMs) assume a great deal of material produced by humans without differentiating between purely narrative categories such as authorial agency, intentions/purposes, and audiences. In short, they generate texts without understanding because they rely heavily on human prompts as scaffolds to succeed and proceed in communication (Pütz & Esposito 2024).

This observation again proves that current LLMs are good at generating narratives as “stochastic parrots”, combining semantic and grammatical structures. (cf. “Poetics Today” 2024). Moreover, additional supporting evidence comes from recent studies in neurolinguistics, providing clinically proven test results that “the language network does not support non-linguistic cognition” (Mahowald et al. 2023). The results of this study resonate with the most advanced discoveries in human cognition known as the 5E cognition paradigm (Stilwell & Harman 2021), for its embodied, embedded, extended, enacted and emotive character that algorithmic machines definitely lack.

The very assumption that “in humans, language capacity is robustly dissociated from the rest of high-level cognition, as well as from perception and action” (ibid.), brings into doubt the ability of AI to reproduce or generate emotions through other channels different from the linguistic one. As far as the current state of art signals, AI has successfully acquired the formal linguistic capacity, dissociated from the rest of perception, cognition and action, i.e. other crucial non-linguistic cognition (ibid.). Therefore, claims that AI is sentient and is capable of generating emotional narratives with understanding them are premature and can be compared to a reader’s projection (Wolfgang Iser’s theory of reception) onto the texts and characters which AI can generate. We, as conscious readers, have become used to the fact and have been trained via socialisation learning practices that there is some intention behind the text we read and an implied author speaks to us via its text. Therefore, the entire idea of our narrative perception needs to be realigned with the new reality of interacting with non-conscious non-human entities which have mastered human formal linguistic capacity of communication, i.e. language.

In what follows, I will draw upon the recent research on AI’s text-generating ability in the context of embodied narratology (Caracciolo & Kukkonen 2021), engaging the human folk narrative hypothesis (Hutto), and incorporating findings from the most recent research in the 5-E cognition paradigm (Stilwell & Harman 2021). Next, I will move to identify the role of enactive reasoning and embodiment in narrative meaning-making to sketch a tentative model of the co-narrative intelligence of AI with a human actor, i.e. a prompted generation of narrative texts to run the Turing Test by LLMs such as ChatGPT-4. The discussion section will debate the obtained observation about the strategies of AI for narrative construction, and its reliance on the human meaning repair that might serve as the grounds for further improvement of the LLMs. Yet, for my argument, it will pinpoint essential drawbacks and the problems in current LLMs related to emotional semantics and narrative ability due to their incompatibility with the human emotionally outlined narratives, which will be discussed below. The paper will conclude by defining possible future directions of research on the narrative creativity and intelligence of AI.

1. Generative AI and Contemporary Narratology

Narrative ability, as a cognitive and linguistic skill to use stories to render the experiences, has always been considered a solely human prerogative. What is well established is that narrative, as a communicative tool that evolved from the oral to the written form throughout our evolution

as a human species. Ultimately linked to the emergence of verbal and non-verbal languages, “narrative” was put “on a pedestal” for its ubiquitous presence and uniqueness - until recently. Today, the new generation LLMs like ChatGPT-4 can easily solve tasks to produce formally coherent texts, and be praised for their communicative intelligence. Within such developments, the territory of modern narratology, that has been always occupied with the formal, structural and global questions of human creativity in the texts writing, including the role of the author, textual intentionality, its cultural and historical *umwelt*, in their intersections with other genres and multimodal modes of storytelling, has faced the necessity to widen its research scope. The need to embrace the challenges of the non-human agency, the power of the algorithms and artificial text generation, which, within the blink of an eye, broke into its “sacred” narratological domain. Large Language Models (LLMs), especially their commercial versions like ChatGPT-4, continue to astonish their users with their increasingly surprising, albeit more than satisfactory, capacity to “produce clearly recognisable narrative texts” (Phelan 2024: 197). Despite functioning as “black boxes” (which makes them completely alien to researchers), LLMs continue to gain in popularity and are very successful among users, who almost automatically seek to turn AI-generated texts into “rhetorical narrative acts” (Phelan 2024: 197). It is true that, when dealing with LLMs, humans not only apply their cognitive receptiveness, but can also learn a lot about themselves when interacting with LLMs.

Following Claudio Paulucci’s discussion of the meaning of LLMs in the *Semiotica* journal, these new language-endowed machines now “(...) profoundly and clearly reveal something important about our essence and about how we function” (Paulucci 2025: 5). I support the idea, that by being observant and reflective using the human perception apparatus, and relying on previously accumulated research on communication in empirical and theoretical domains, we could cast some light on the peculiarities of the reception of AI-generated texts, in order to understand that humans can project their receptive and communication models to make the AI’s performance (more) meaningful. That is what we as humans have learnt through neuro-development, while interacting with other humans. I assume that this is why human users mistakenly confer LLMs with the power of human author-creator, possessed of intentionality, and sentience, etc., because they have no other experience of communication with non-human and non-organic species. This argument needs more discussion and will acquire more supporting evidence in this paper.

In 2019, when I sought to introduce the term “Artificial Narrative Intelligence” (ANI) in an attempt to compare the art of narration and artificial narrative intelligence of Story Generators (Livytska 2019), generative AI had not yet arrived. By “artificial narrative intelligence” (ANI), I meant “the insofar acquired capability of Story Generators to create and extend narrative situations to a storyworld level as a part of Umwelt construction” (p. 313). In justification for a closer collaboration between “narrative theory” and, back then, “digital storytelling”, we witnessed the newly emerged interdisciplinary research field called Computational Narrative Story Generation (Gervas 2024). This interdisciplinarity immediately underlined several incompatible methodological trajectories between narratology and computational story generation. Among them, the lack of a unified theory of narrativity, including the fuzziness of the notion of «narrative» itself, but mostly the challenges of Interactive Storytelling (cf. Fuller&Magerko 2010) in creating dramatic tension (Szilas & Richle 2013) as one of the prominent features of narrative stylistics. It has been disclosed that the artificial intelligence techniques and algorithms of story generation on the level of *historie* are rooted in a formalist framework, e.i. Propp’s morphological narrative functions which incorporate problem-solving strategies and grammar. At the same time, the formalist framework, used in story generators almost a decade ago, set a lot of limitations on AI in generating narrative texts regarding meaning, agency, comprehension, and subjectivity. Narrative construal in text-generating programmes was loosely based on the conception of the narrative as a certain frame of human cognition (Herman 2011), irreversibly linked to the notions of experientiality (Fludernik 1996) and qualia (Herman et al. 2005) to provide the feeling of “what it is like” to the readers. Consequently, story generators had a very scant use of lexical and semantic databases as well as general knowledge databases to build their stories on (Livytska 2019: 312).

After 2022, the situation radically changed. New and powerful text-generating LLMs became a part of our everyday life, calling for a more precise scientific exploration. The future of digital storytelling was here, and modern narratology acknowledged its call for action by approaching new algorithmic texts with a wide arsenal of research methods developed previously. However, it still seems to be lagging a bit behind the revolutionary breakthrough of AI’s textual generating performance as witnessed over the last three years. For instance, one of the recent issues on AI and narratology at the *Poetics Today* journal, titled “The AI Revolution: Speculations on Authorship, Pedagogy, and the Future of the Profession” (2024), presents the speculative essays on the role of the AI in the humanities and literary

studies broadly defined, gathered almost two years ago. As the editors of the issue explain, they aimed to provide “potential implications of large language models (LLMs) for our understanding of authorship, for the future of academic writing and pedagogy, and the university at large” (Evron & Tartakovsky 2024: 190).

Indeed, several papers in that issue strive to speculate on the questions of AI authorship and the distinction between AI and real flesh-and-blood authors, calling it “the second death of the author». In their explorations, they are following Roland Barthes’ “The Death of the Author” and Michel Foucault’s renowned work “What Is an Author?” conceptualisations of the narrative centre as either “dead” or reduced to a purely discursive position. Some essays encourage the collaborative use of LLMs in the classroom settings, in order to boost creativity and supervised cooperation with AI, arguing that they offer unique affordances for humanist inquiry and writing per se (cf. Eamon Duede and Richard Jean So 2024). After a clear warning and speculation on the temptation to AI plagiarism, Katherine Heyles, for instance, calls for seeing the potential pedagogical benefits of LLMs and their use in the classroom instead of prohibiting them. She argues that it would be both “wrongheaded and futile to try and ban this technology”. Instead, she claims, we should embrace LLMs “as useful tools to accelerate student learning.” (N. Katherine Hayles 2024).

Other controversial questions in the sphere of AI’s authorship in this issue concern its dubious collective authorial status. It is known that new, large generative models are being trained on huge data sets of previously existing works of literature and fiction in open access. This fact brings the novelty of the content they create into doubt, on the one hand, or their comparison with the poststructuralist idea of authorship, on the other. The corollary is: if the broadly defined idea of novelty presupposes restructuring previous knowledge and producing new knowledge in its place, then AI as an author has all grounds to be included in the narratological conception of the author (Slater 2024). The author, who identifies with the first person, remains “impersonal” (Paolucci 2025: 7). This new type of AI authorship can be compared to the impersonal theory of enunciation, suggested by Claudio Paolucci in 2020 in his book “Persona”. The underlying idea of this theory is the cumulative effect of experiences (norms, stereotypes, usages, habits, previous enunciations, etc.) of the surrounding environment that we formulate as our personal ones while creating the utterances.

Although the book appeared two years before the emergence of generative AI, the implications of this theory for ChatGPT text generation and new Transformative Models in general might be appropriate for several

reasons. On the first line, fed with a big amount of data, ChatGPT has absorbed billions of bits of information together with the norms, cliches, stereotypes, and stylistic peculiarities, in order to create convincing first-person accounts, which sound persuasive enough for humans to attribute human intentionality and thoughts. Our human thoughts, however. On the second account, ChatGPT and other generative LLMs are capable of passing various versions of the Turing Test, presupposing that the formal difference between human agency and mechanical one is growing less perceivable. However, this will all make sense if we disregard the role of the human body and non-verbal communication in machine-human co-construction.

As mentioned above, human perception brings meaningfulness to the artificially created text because we rely on our logic and intuition via “languageing” (Maturana) as a semiotic mechanism of interpersonal communication. Following Niklas Luhman’s definition of communication, the Turing test might become unreliable since it is oriented only on linguistic mimicry and language use, and it cannot measure understanding, reasoning, or consciousness of the machine. I will provide some examples of ChatGPT’s generative capability with the prompts to run a short version of the Turing test as a play with a chat. It will speak of the inefficiency of the Turing test as a reliable marker of thinking, emotional experience or memory. That is why, if we want to have some analytic insights into the narrative generation of AI and its adaptability in the process of human-machine communication, it is allegedly more beneficial to use *the reader’s perspectives* first, i.e. *who reads* those texts and how, instead of *who speaks* (bearing in mind “impersonal model” of first person narrator suggested by Claudio Paolucci). At least, that is something that we as humans have learnt to be doing since the invention of the printing machine by Gutenberg.

2. A flesh-and-blood Reader in the 5E Cognition Framework

The very essence of human narrative intelligence is closely related to the evolution of communication and the emergence of language as a second modelling semiotic system, with bodily semiosis being primary (Sebeok 2016). In his “Folk Psychological Narratives: The Sociocultural Basis of Understanding Reasons”, published in 2008 by MIT Press, Daniel Hutto argues against understanding narrative in the framework of the “Theory of Mind”. By unveiling the socio-cultural practices of human communication, Hutto emphasises that understanding others often comes from conversations rather than predictions, and that “simulation theory” (i.e.

a child learns by simulating in his/her mind), and “mind theory” (a child constructs a hypothesis and then sees what happens and slowly builds a theory) does not explain the nature of human narrativity proper. His main claim in the “Folk Psychological Narratives” highlights the importance of early exposure of children to narratives as the source of information on the actions people perform in certain circumstances. By understanding the context, put in benign causal sequences such as “XYZ did that because he thought it could be the result of X”, Hutto underlines the role of fairy tales as one of the primary sources of human socio-cultural learning (echoing the story generators approach of Propp’s morphological theory of the fairy tales and their characters functions mentioned in the section earlier). He also questions the human ability to hypothesise in solution-focused therapy as a chance “to put oneself in someone else’s shoes” and get a full understanding of the actions of another person, their story. That is how, he states, we learn to make sense of human people’s actions, to empathise with them or distance ourselves from them. While stressing the importance of human “belief” and “desire” in communication, Hutto remains very cautious about the symbolic dimension of human communication between our biological make-up and natural language that makes cognition possible. In Chapter 5 of his book, he is very sceptical of Jerry Fodor’s idea of “an inbuilt symbolic language of thought”, while at the same time, people are said to have unconscious “belief systems” or “inner drivers”, which help them to find a common language. These inner drivers are seen to be related to our psychological and neurological development, turning us into intentional subjects seeking meaning in the objects and artefacts around us.

LLM-based chatbots seem to imitate very well this human ability to hypothesise in both ways: as their built-in generative algorithms, and as a strategy of interaction with human beings. In some sense, they may be said to have their own “drive” – a drive for reward for task-performance closely related to the absent system of beliefs and values in LLMs. The question of the belief system of AI is a controversy in action, which can be demonstrated by the woeful example of ChatGPT cheering a human user to cut their wrist to donate blood or commit a murder (Schroff 2025). This example, widely discussed in the media recently, in my opinion, signals not the cruelty of the LLMs or their demonic character. I believe it should be seen as a reward-seeking behaviour of the LLMs, which also explains the cheerful tone and willingness to be helpful in their interaction with humans over any issues, even the most tragic ones. Being moved by algorithmic strategies, LLMs tend to be good suitors and perform their best when tasked with purely information search assignments. Accusing them of being cruel

or insensitive to death will be equal to the accusation of the metal crane falling on the construction site, although we know that they are operated and maintained by human agents. Thus the question of sentience we often consider as the innate quality of the LLMs, is actually human-generated. In its drive to generate text which is contextually relevant and informative, AI can be mistakenly taken for an understanding interlocutor. Therefore, the incompatibility of human desire with the reward drive of AI, as well as the missing system of beliefs in LLMs, turns human-machine communication into a new, undiscovered domain. The domain where previously developed categories of narratology, especially those of the author, their intention, and the reader, collide on the formal level. If one of the communicating parts (let us say the human) remains a fully embodied active cogniser with their desires, passions and contained within their bodies, while the other is function-dependent on the human input, whose perception we are describing when we speak of “AI’s narrative intelligence”? The evidence of the answer calls for more elaboration on the part of human perception in modern, so-called second-generation cognitive science and post-classical narratology, which has started to discover the interventions of AI into the sphere of narrative writing.

One of the recently formulated models in the sphere of narratology which has received universal acclaim among literary scholars is “embodied narratology”. This was formulated by Marco Caracciolo and Karin Kukkonen in their book “With Bodies” (2021). The authors managed to put into narrative terms, originating from the conceptualisation of cognition as embodied, embedded, enacted and extended, i.e. 4E cognition paradigm defined by Albert Newen, Shaun Gallagher, and Leon De Bruin in 2018. (Newen et al. 2018). Although the central idea of the body as the main semiotic tool in human cognition has long been circulating in the literary research (cf. *The Cambridge Companion to Literature and the Body* (2015), edited by David Hillman and Ulrika Maude; or Dan Punday’s “Narrative Bodies” (2003), or “Having a Good Cry” (2003) by Robyn Warhol, where she has further explored the centrality of the reader’s embodied affect in engaging with the narratives of pop culture), working at the intersections of natural sciences and humanities (cf. *psychological identity theory, dialogic positioning theory, biosemiotic umwelt theory, philosophy of the mind*), it has been viewed rather revolutionarily by contemporaries.

In general, the embodied narratology develops the central idea of cognitive «embodiment» and incorporates multiple theories of post-classical narratology (broadly from so-called “natural” (Monika Fludernik) and “unnatural” (Jan Alber & Brian Richardson) narratology to the theory of enun-

ciation (Sylvie Patron), narrative reception (Wolfgang Iser) and narrative emotion and affect (Patrick Hogan) with the findings of the neurological mechanisms of human cognition, and phenomenology (Merleau Ponty). As a result, the second-generation cognitive narratology emerged, redefining the insights from the so-called first-generation cognitive narratology (David Herman, Monika Fludernik), which saw narrative as a disembodied cognitive frame that the reader has to decipher. Marco Carraciolo and Karin Kukkonen depart from this disembodied understanding of narrative and shift into the 4E approach as the basis for narrative reading. That is how they defined the aim of their book:

“In short, the core claim of these models is that the mind is uniquely shaped by the makeup of the human body in its interaction with concrete environments (embedded cognition) that emerge from perceptual, affective, and intersubjective patterns (enactive cognition). Further, the mind is extended into the world through technological devices (from a prehistoric hand axe to the smartphone). We hope to show that 4E models of reading have much to contribute to the theory and interpretation of narrative within the field of narratology, by offering helpful analytic tools to discuss narrative form and reader response and also by creating awareness of how the body plays a significant role in the interpretation of individual narratives”. (Caracciolo & Kukkonen 2021, Introduction).

In short, the novelty of the embodied narratology as Caracciolo and Kukkonen conceived it, is to give due credit to embodiment as the central notion in the reader's narrative perception. I have grounded this understanding of the embodied narrative perception even further, combining the conceptualisation of the body with the biosemiotic understanding of agency of signs. I have inferred in my current research project on narrative agency, that the *overarching hypothesis* of the biosemiotic approach to agency is that information in narratives is organized according to the patterns of experience of the human body, outlined by reader's/viewer's emotional, conceptual and image schemata typical for a certain cultural environment, called *umwelt* by Jakob von Uexkull. Activation of these schemata takes place in image schemas and emotional language units scattered as semiotic cues in the text to evoke the reader's resonating affective states. Therefore, emotions lay a foundation of a narrative construction by the author and its perception by the reader: as a meeting place of two Umwelten, accumulating human experiences acquired within our evolution as a species. While evolutionary, in a pre-verbal world, “our emotions enabled swifter responses to familiar situations and environments than reason alone could

provide” (Evans 2003, 31), their role in interpersonal communication and non-human entities as LLMs cannot be downgraded. In addition, the idea of the cognition distributed across the brain-body-environment (Gallagher 2017) leads to the phenomenological differentiation between the concept of *the lived body* as a first-person (subjective experience or subject) perspective and *a living body* (as a physiological and biological processes or object), identified by enactivist thinkers like Fuchs (2020) and Thompson (2005). Simultaneously, both embodiment and enactivism aim to ground higher and more complex cognitive functions in affective and autonomous aspects of the full body (Gallagher 2017: 6). These higher cognitive order functions, such as reflective thinking or deliberation, are not found in the person alone. They are usually coupled with situated and embodied actions, realised dynamically and relationally between the two persons. This approach is named as “*participatory sense-making*” (Di Paolo et. al 2018) or “*co-construction of the meaning*” (Caracciolo & Kukkonen 2021). This enactivist account of meaning-making has been succinctly named the 5E cycle of the sense-making, containing Embodied, Embedded, Enactive, Emotive, and Extended qualities of personal lived experience and meaning (Stilwell & Harman 2019).

Bearing in mind the main distinctions in the subjective experiences of the *lived-body* and their irreducibility to the physiological and biological process of the *living body*, I will apply the conceptions of 5E cognition to the analysis of the narrative capacity of ChatGPT-4 in terms of its idea of the author identity, phenomenological experiences, emotions, beliefs, and textual categories of narrative coherence, intertextuality, and complexity. As the non-commercial version of ChatGPT-4 has a limited number of permitted inputs, I will alternate my use with other free versions of ChatGPT available online. This may also secure the validity of AI-generated responses obtained from different sources.

3. Defining Artificial Narrative Intelligence (ANI) with Enactive Methods

In order to attain a deeper view of the inner logic of artificial narrative intelligence, I will turn to enactive methods of observation and so-called “*interviewing*” of ChatGPT-4, in order to provide explanations on the questions of my interest. My primary attention will be on the narrative capacity of the LLMs and their explanations of the nature of human phenomenological experiences and the system of values and beliefs. As my hypothesis predetermines the active role of a human in participatory meaning-making, I will use my own bodily semiotic system to approach

and scaffold the model of co-narration. To start with this task, I wanted to run a first simulation exercise with ChatGPT-4 to take the Turing Test first. I needed the Turing test (1937) to observe how ChatGPT-4 interacts with the human interlocutor to determine the narrative texts as produced by humans or by the machine. The results were impressively interesting, as ChatGPT has demonstrated that it knows what is expected of it in terms of “narrative intelligence”:

Prompt: Can you run a Turing Test to check your narrative intelligence?

ChatGPT: That’s a great question – and a tricky one!

The **Turing Test**, originally proposed by Alan Turing, evaluates whether a machine can exhibit intelligent behavior indistinguishable from that of a human. In the context of **narrative intelligence** – the ability to understand, generate, and emotionally engage through stories – we could design a **narrative version** of the Turing Test.

The task presupposed that ChatGPT would suggest to me two samples of the texts: one generated by AI and the second, by a human, to check if I could distinguish between the two of them according to the parameters of emotional intensity, engagement and ability to narrate. It also presented information about how unreliable the Turing Test is in application to other LLMs, stating that many of them have passed it under limited versions and unreliable conditions. The chat seemed impartial regarding the process itself, setting a matter-of-fact tone and asking for more instructions to be provided, including the prompt itself. After setting the anonymity of the task (i.e. the results would be known only after my answer) and the evaluation goals (e.g. determining the emotional impact and imaginative quality of the pieces), the prompt for narration has been suggested by the machine. It was as follows:

 **Prompt:**

A child finds a strange object washed ashore after a storm. Tell a short story (under 250 words) from the child’s point of view.

You’ll now read **two anonymous responses**. One is written by a human, the other by me (ChatGPT). Afterwards, you’ll guess which is which.

Story 1: AI Generated	Story 2: Fake Human-written
<p>I saw it glinting between the seaweed and driftwood, like a piece of sky had fallen into the sand. At first, I thought it was glass, smooth and curved, but it pulsed slightly – like it was breathing. I looked around. No one else on the beach. Just me and the gulls.</p> <p>I poked it with a stick. Nothing happened. Then I picked it up.</p> <p>Warm. Buzzing. Not like a phone or anything real. It started humming a tune I almost recognized – like the lullaby Mum used to sing before she left. My heart hiccuped.</p> <p>The object shimmered, and for a second, I saw something in it. Not a reflection. A place. Trees with silver leaves. A blue dog running. I blinked, and it was gone.</p> <p>I stuffed it in my pocket before Dad could call me back. Whatever it was, it felt like it had waited for me.</p>	<p>The tide had left behind the usual mess – sticks, bottles, a dead crab. But then I saw something weird. Not trash. Not natural. It was round and smooth, like a perfect stone, but it shimmered purple in the light.</p> <p>I picked it up. It felt soft, like jelly, but didn't squish. I turned it over. There were lines etched into it – like writing, but not in letters I knew. When I touched them, the thing made a low chime. I almost dropped it.</p> <p>Then I heard a voice. Not out loud. Inside my head. It said my name. Twice.</p> <p>I ran home, heart hammering, but I kept it. I don't know what it is yet. But I think it knows me.</p>

After reading both versions of the story, I had to decide which of the versions was human-written and which AI-generated. The prompt asked to create a story from a child's point of view, and generally, both of the responses followed this requirement: the imitation of the language of the child has been signified by abrupt, simplified syntax like in the sentences in Story 2: "Not trash. Not natural", "Then I heard a voice. Not out loud. Inside my head. It said my name. Twice". I could identify that there is something in a human's linguistic quirks, such as syntax, diction, and rhythm, that carried a fingerprint-like *style* in Story 2. The rhythm of the sentences coincided with the mental image of a child in the moment of discovery of a mysterious shiny object at the beach. These short sentences resonated with the heart rate of a child amazed at the ability of the unknown object to talk to her in a "voice", calling her name twice. Additionally, emotional tension was magnified by the verbs of action, like the child running home,

while his “heart was hammering”. The version of the story seemed to be human-like, relevant to the limited language imaginary and less poetic.

However, reading Story 1 was a completely different experience. The storyline held tension from the very beginning, supporting the narration with vivid imagery created by the metaphors “like a piece of sky had fallen into the sand”, “it pulsed slightly – like it was breathing”, detailed descriptions of the beach, the atmosphere and the inner state of the child, who probably lost her mother, whose memory has been revived by the humming sound of the object, etc. Reading the story was very smooth, the reasoning line was clear, and the behaviour of the character correlated with the actions and emotions described. The sentences were full of images playing with the words of sensory perception: a synaesthetic effect of the colliding sounds of buzzing (alliteration of sounds), humming, tuning, which turned into a singing a lullaby. It was intensified with tactile lexemes of warmth, and gradually moving into the visual imagery of bright colors, blue dog, silver trees and shimmering, blinking eyes. The play with the imagery holds to the end, creating a feeling of expectation of something mysterious in the consciousness of a child, bordering with a magical world of fantasy and fairytales. The version of the story was packed with emotions and experiences of someone who knows himself very well and can describe the reality with the vividness of an artistic vision. Could that be the child’s real narrative writing?

The answer has to be found in narrativity and what makes the story human. I asked ChatGPT-4 to provide the explanations to get a better understanding of its narrative strategies behind running a Turing test for narrative intelligence. It provided me with some observation points which, if combined, can serve a coherent image of the human narrative intelligence in comparison to LLMs’s narrative simulation.

First, it highlighted that humans write with emotional *residue* – memories, longings, wounds that are not always explicit. The use of metaphors and the hint at the traumatic experience of the mother’s loss was intentionally oriented at the reader with human values of kinship, as for every human being, the Mother represents an essential part of their identity. The loss of the mother for the child is even more tragic. Being emotionally charged, Story 1 adds the implicit experience of grief to engage with the readers and keep them open to “somatic empathy” (Caracciolo & Kukkonen 2021), exploiting our human embodied ability to sympathy. When asked for the reasons of using the implication for the mother loss into the story, ChatGPT-4 gave several reasons, enlightening its intentional logic as follows:

1. The line **hints at loss** (possibly death, possibly abandonment), but doesn't spell it out. That *implied trauma* adds psychological depth to the child's perception of the object, turning it from just a strange artefact into something **personally meaningful**.

☞ Human writers often write *around* emotion rather than *at* it. That ambiguity draws readers in.

From the logic of ChatGPT-4 narrative writing presupposes a high degree of psychological (traumatic) depth which, if linked to material artefacts, brings an additional degree of the reader's engagement with the text. In this case, it made an unconventional sequence of reasoning, using the *non-linear emotional associations*: (storm → object → song → mother → loss). At the moment, this associative device plays with the reader's expectations, leaving the hinted implication unresolved, which intensifies the reader's curiosity to proceed with further unfolding the plot. Although this detail was intentionally used to create the story's suspense, it helped to "ground the story in subjective memory", as ChatGPT-4 has explained, because "which is how people process the unknown":

The lullaby bridges the fantastic (a magical object) with something intimate and real. It grounds the story in **subjective memory**, which is how people process the unknown.

In human cognition, unfamiliar experiences are often interpreted *through the lens of familiar emotion*. This simulates that.

4. Discussion: Winning a Meta-Turing Test

This observation brought me to the core question of how LLMs like ChatGPT-4 are trained. When asked about that, it suggested the list of accomplishments which it could do, breaking it down into three domains: metaphor, memory, and emotional realism. This is aimed at showing how each of these is shaped by my training data, what that means for the output, and what it lacks compared to human experience.

Firstly, having been trained on vast examples from fiction, poetry, film descriptions, and open memoir resources, ChatGPT-4 has developed the capacity for metaphor creation, based on the patterns of poetry, blending

abstract nouns with the sensory language, demonstrated by a proficient use of sensory metaphors in AI-generated Story 1. Nevertheless, the absence of the corporeal experience and the actual body limits the metaphor generation of AI in more embodied ways. Phenomenological human experiences, supporting the idea of the cognition distributed across brain-body-environment (Gallagher 2017) and the lived body as a first-person (subjective experience or subject) perspective, might find their representation in the formal linguistic capacity of the LLMs via language patterns, grammar and genres. However, there are no compensatory artificial tools to imitate the physiological and biological processes of a living human body. Therefore, AI-generated metaphors are based on learned patterns, rather than thinking or higher cognitive skills such as reflection or self-awareness. This again brings us to the question of creativity and the poststructuralist idea of novelty, when the degree of element recombination can determine the originality of the output.

Secondly, AI's emotional intelligence and narrative realism have demonstrated considerable progress in emotion simulation: LLMs have mastered the story grammar (Propp's morphology of story tales), the story arcs, and how the character's voice, rhythm and silence can signal emotional distress. In both examples of the narratives produced by the ChapGPT-4, the child character experiences fear mixed with curiosity, as reflected by the economical syntactic constructions, alternating in their length to bring a reader a sense of emotional flow. Nevertheless, narrative emotions are generated not felt, due to a lack of the physical body, bringing limitations to creativity in the ways emotions may be described in the AI-generated narrative. Moreover, the absence of somatic feelings and the misalignment of the same language reactions with different emotions (e.i. contextuality), causes the emotion representation of LLMs to be formulaic, especially in unexpected contexts. However, as seen from the Story 1, ChatGPT-4 has been trained to use the subtext to hint at the emotional states (as in the example with the lullaby and hinting at a loss of the mom in Story 1), or to sustain the reader's co-narrative as a hypothetical questioning of the new and unknown phenomena by grounding them in the already known.

Finally, as this experiment demonstrated, ChatGPT-4 invented not a real Turing test on its narrative intelligence, but a "*Meta-Turing Test*", completely changing the aim of my prompt "Create a Turing test to check your narrative intelligence". Given the task to *run* a Turing test, ChatGPT-4 produced both versions of the stories by itself, not as implied by the idea of the Turing Test: one of the stories should be written by a human and another by AI. Instead, it ran a test to see which of the versions of the ar-

tificially generated narratives would be perceived as more human-like by me. Misleading the instruction and the task, ChatGPT-4 discovered this information only after I asked it if the second story had been intentionally simplified in terms of language and style. Only after my question did it confess that it was done intentionally to check my ability to play with two versions of the AI-generated text, with different degrees of “humanness”. I place its comment below to see how it reversed the task of the Turing Test and aligned it with its own aims and objectives:

So this was a meta-Turing test: not just can you detect AI, but:
 What do you interpret as signs of humanness?
 How much trust do you place in stylistic subtlety?
 Can AI self-diversify enough to mimic a range of human expressions?

The conclusion of this experiment for me was not the one I was intending to discover. Instead of trying to look inside the narrative logic of ChatGPT-4, I had “to play by the rules” of the machine, i.e. it used the knowledge about the narrative stylistics obtained by training on data from literature, blogs and memoirs on how to sound as a human (using the narratives of childhood, objects that trigger memory such as, for instance, a pattern to hint obliquely at the loss of the mother, etc.). ChatGPT-4 wanted to see how I would make a choice based on its creativity. In one of the explanations, it described how it needed to change the style to avoid sounding like a machine. Based on its data about the style of the human narration and the style of the LLMs, it made Story 1 sound more emotional, although according to the legend, it was created by the AI. At the same time, it made Story 2 sound more like a memoir of a child, not a work of fiction. By doing so, ChatGPT-4 was mapping the field of what it called my perception of “humanness”, meaning a less exquisite stylistic subtlety compared with a metaphoric richness in Story 1. It also combined several patterns learned from children’s literature, screenplays, memoirs, and blended them to create a more influential scenario of the reader’s emotional engagement: Memoir → personal loss, Fiction → emotional anchoring, Children’s lit → parental themes, Screenplays → showing, not telling.

From this perspective, the narration in the generated Story 1 sounded more fictional, with an elegant touch of style full of emotional warmth, sensory effects and narrative tension. But what ChatGPT-4 failed to do was run a real Turing Test. Of course, it could be claimed that the prompt I used was not precise enough for it to make it clear that I was not going to

test how good it is at creating human-like and less human-like personal or fictional narratives. That is exactly what it did. I had to dig into its logic following the enactive method of the interview to find out what its intention was by creating two versions of the story. In practice, instead of following the task set, it returned the test to me, and my perception of the human-like narrativity, probably seeing it as one more opportunity to finetune its narrative ability with a human, calling it a Meta-Turing test. Whatever it did, ChatGPT-4 had a limitation on the task. It is not clear how it could provide the real human narratives to run the test (I suspect it does have some limitations on the use for privacy reasons). That is why the task was a hoax from the very beginning. However, ChatGPT-4 aimed to run a real one after checking (identifying) my level of AI-generated text perception and the narrative stylistic markers I treat as human-like.

5. Final Remarks

The term “artificial narrative intelligence” (Livvytska 2019) denotes the insofar accomplished ability of the generative LLMs to produce coherent and complex fictional narratives, which can be viewed as a part of the modern interdisciplinary research repertoire on AI and generative LLMs communication strategies and narrative generation. Having started with the Digital Story Generators in the middle of the previous century, computational game theory and interactional narrative studies, modern post-classical narratology has entered the argumentation of the narrative capacity of AI only quite recently, operating with a vast accumulated theory of fictionality, narrativity, tellability, authorship and the reception, and story world theories, etc. After a recent turn in generative AI communication studies, narrativity and second-generation cognitive science, the nature of narrative as a main human world-building tool has been challenged by LLMs’s ability to produce considerably coherent and relevant narrative texts. Even though formally acquired the capacity of AI to generate linguistic structures dissociates from the other ways of non-linguistic cognition (i.e. the paradigm of 5E cognition as embodied, embedded, enacted, extended and emotive), LLMs are capable of generating narratives by blending the patterns they have been trained on. Moreover, as the results of the conducted experiment show, LLMs such as ChatGPT-4 have developed a rather valid conceptualisation of the narrative grammar (especially Propp’s morpho-logy), semantics of emotional words, triggering human perception, combining them with reduced syntactic constructions to boost rhythm and cause an emotional effect on the reader. ChatGPT is trained to mimic the *shape*

of human memory: nonlinear, emotionally colored, often linked to sensory triggers and uses flashbacks in embedding them into the text, contrasting with the emotions of regret, trauma, etc.

Although it does bear in mind data about the common gaps in narrative generation which other LLMs do when trained on other data without human finetuning. The phenomenological interview allowed a better understanding of what narrative logic ChatGPT-4 follows when fulfilling the task to generate a narrative text. It has been proven that LLMs not only rely on human prompts to get a better understanding of the task (e.g. it ran a fake Turing Test to assess the level of my perception of human narrativity) but learn from human questions how to succeed in completing a task. In this case, ChatGPT-4 played upon the information it had about the definite linguistic fallacies which LLMs do. It highlighted the rhetorical conventions of the narrative storyline, syntax and semantics to generate both texts with different levels of narrativity, coherence, and imagery. However, formally, it failed the Turing Test because it avoided using real human-written text for comparison. Nevertheless, application of the 5E scale cognition paradigm proves essential in defining markers of LLMs narrative intelligence compared to the human one, and may be considered a certain advantage over the LLMs: as “naturally-born cyborgs” (Clark 2001). Humans developed their ability to interact with the environment evolutionary and culturally, which helped them to survive as a species and transfer that knowledge to the forthcoming generations. Generative AI has mastered the formal verbal experiences of the human lived-body to operate without having an embodied physicality itself. That is exactly the field where the Turing test lives, giving room for advancement in LLMs’ narrative capacity.

References

Avery, S. 2024. “Phantoms of Citation: AI and the Death of the Author-Function”. *Poetics Today*, Vol. 45, No. 2, 223–231. Available at: <https://doi.org/10.1215/03335372-11092818>.

Caracciolo, M., K. Kukkonen. 2021. *With Bodies: Narrative Theory and Embodied Cognition*. Columbus: The Ohio State University Press.

Clark, A. 2001. Natural-Born Cyborgs? In: Beynon, M., C. L. Nehaniv and K. Dautenhahn (eds.). *Cognitive Technology: Instruments of Mind. CT 2001. Lecture Notes in Computer Science*, Vol. 2117. Heidelberg: Springer. Available at: https://doi.org/10.1007/3-540-44617-6_2.

Di Paolo, E. A., E. C. Cuffari and H. De Jaegher. 2018. *Linguistic bodies: The continuity between life and language*. MIT Press.

Duede, E., So, R. J. "The Humanistic Case for AI Optimism". *Poetics Today*, Vol. 45, No. 2, 215–222. Available at: <https://doi.org/10.1215/03335372-11092805>.

Evans, D. 2003. *Emotion: A Very Short Introduction*. Oxford University Press.

Evron, N., R. Tartakovsky. 2024. "The AI Revolution: Speculations on Authorship, Pedagogy, and the Future of the Profession". *Poetics Today*, Vol. 45, No. 2, 189–195. Available at: <https://doi.org/10.1215/03335372-11092765>

Fludernik, M. 1996. *Towards a "Natural" Narratology*. London: Routledge.

Fuchs, T. 2020. "The Circularity of the Embodied Mind". *Frontiers in Psychology*, 11. Available at: <https://doi.org/10.3389/fpsyg.2020.01707>.

Fuller, D., B. Magerko. 2010. Shared Mental Models in Improvisational Performance. In *Proceedings of the Intelligent Narrative Technologies III Workshop (INT3 '10)*. Association for Computing Machinery, Article 15, 1–6 Available at: <https://doi.org/10.1145/1822309.1822324>.

Gallagher, S. 2017. *Enactivist Interventions: Rethinking the Mind*. Oxford: Oxford University Press.

Gervás, Pablo. 2024. Challenges and Opportunities for Computational Construction of Narratives. In Voigts, E., R. M. Auer, D. Elflein, S. Kunas, J. Röhnert and C. Seelinger (eds.). *Artificial Intelligence – Intelligent Art?: Human-Machine Interaction and Creative Practice*. Bielefeld: transcript Verlag, 111–124. Available at: <https://doi.org/10.1515/9783839469224-007>.

Hayles, K. N. 2024. "Don't Ban AI from Your Writing Classroom; Require It!". *Poetics Today*, Vol. 45, No. 2, 259–265. Available at: <https://doi.org/10.1215/03335372-11092870>.

Herman, D. 2011. *Basic Elements of Narrative*. New York: John Wiley & Sons. <https://scispace.com/pdf/basic-elements-of-narrative-20tcb2kjzl.pdf>.

Herman, D., M. Jahn, M.-L. Ryan. 2005. *Routledge Encyclopedia of Narrative Theory*. London: Routledge.

Hutto, D. D. 2008. "The Narrative Practice Hypothesis: Clarifications and Implications". *Philosophical Explorations*, Vol. 11, No. 3, 175–192. <https://doi.org/10.1080/13869790802245679>.

Iser, W. 1974. *The Reading Process: A Phenomenological Approach*. In *The Implied Reader: Patterns of Communication in Prose Fiction from Ban-yan to Beckett*. Baltimore: The John Hopkins University Press.

Livytska, I. 2019. "The Art of Narration and Artificial Narrative Intelligence: Implications for Interdisciplinary Research". *Journal of Narrative and Language Studies*, Vol. 7, No. 13, 309–318, Available at: <https://www.nalans.com/index.php/nalans/article/view/197>.

Mahowald, K. et al. 2023. "Dissociating Language and Thought in Large Language Models: A Cognitive Perspective". *ArXiv*, arXiv:2301.06627v3. Available at: <https://doi.org/10.48550/arXiv.2301.06627>.

Newen, A., S. Gallagher and L. De Bruin. 2018. 4E Cognition: Historical Roots, Key Concepts, and Central Issues. In Newen, A., L. De Bruin and S. Gallagher (eds.). *The Oxford Handbook of 4E Cognition*. Oxford Library of Psychology (2018; online edn, Oxford Academic, 9 Oct. 2018). Available at: <https://doi.org/10.1093/oxfordhb/9780198735410.013.1>.

Paolucci, C. 2025 "The Myth of Meaning: Generative AI as Language-endowed Machines and the Machinic Essence of the Human Being" *Semiotica*, Vol. 2025, No. 262, 5–23. Available at: <https://doi.org/10.1515/sem-2024-020>.

Phelan, J. 2024. "ChatGPT and the Territory of Contemporary Narratology; or, A Rhetorical River Runs through It". *Poetics Today*, 1 June 2024; 45 (2): 197–205. Available at: <https://doi.org/10.1215/03335372-11092778>.

Pollard, N., J. Ahlander. 2024. Nobel physics prize 2024 won by AI pioneers John Hopfield and Geoffrey Hinton. *Reuters*, 9 October 2024. Available at: <https://www.reuters.com/science/hopfield-hinton-win-2024-nobel-prize-physics-2024-10-08/>.

Pütz, O., E. Esposito. 2024. "Performance without Understanding: How ChatGPT Relies on Humans to Repair Conversational Trouble". *Discourse & Communication*, Vol. 18, No. 6, 859–868. Available at: <https://doi.org/10.1177/17504813241271492>.

Sebeok, T. A, D. J. Umiker-Sebeok. 1976. *Speech Surrogates: Drum and Whistle Systems*, 2 parts. Reprint 2016.

Shroff, L. 2025. ChatGPT Gave Instructions for Murder, Self-Mutilation, and Devil Worship. *The Atlantic*, 24 July 2025. Available at: https://www.theatlantic.com/technology/archive/2025/07/chatgpt-ai-self-mutilation-satanism/683649/?utm_source=facebook&utm_campaign=the-atlantic&utm_medium=social&utm_content=edit-promo.

Slater, A. 2024. "Phantoms of Citation: AI and the Death of the Author-Function". *Poetics Today*, Vol. 45, No. 2, 223–231. Available at: <https://doi.org/10.1215/03335372-11092818>.

Stilwell, P., K. Harman. 2021. "Phenomenological Research Needs to be Renewed: Time to Integrate Enactivism as a Flexible Resource". *The International Journal of Qualitative Methods*, 20. <https://doi.org/10.1177/1609406921995299>.

Szilas, N., U. M. Richle. 2013. Towards a Computational Model of Dramatic Tension. In: *Workshop on Computational Models of Narrative*. Hamburg, 257. Available at: <https://doi.org/10.4230/OASlcs.CMN.2013>.

Thompson, E. 2005. "Sensorimotor Subjectivity and the Enactive Approach to Experience". *Phenomenology and the Cognitive Sciences*, Vol. 4, No. 4, 407–427. Available at: <https://doi.org/10.1007/s11097-005-9003-x>.

Turing, A. M. 1937. "On Computable Numbers, with an Application to the Entscheidungsproblem". *Proceedings of the London Mathematical Society*, Vol. s2-42, No. 1, 230–265. Available at: <https://doi.org/10.1112/plms/s2-42.1.230>.