

Technology and Language

125π

SECTION A - FORMES ÉLÉMENTAIRES DE L'ACTIVITÉ HUMAINE Elena Seredkina and Yongmou Liu

PLANCHE I



1. — Percussion perpendiculaire.



2. — Percussion oblique.



3. — Abrasion



4. — Sciage.

Vladimir Arshinov and Maxim Yanukovich

Vladimir Shalack

Rebeca Perez Leon



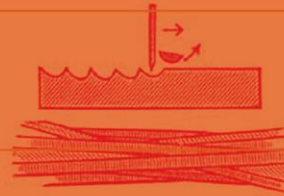
5. — Eclatement du bois « dans le fil ».



6. — Eclatement du bois « à contre-fil ».



7. — Eclatement limité par percussion perpendiculaire préalable.



8. — Grattage et « échelles de grattage ».

Anna Kartasheva

Alexander Vnitskikh and Sergey Komarov



9. — Ciseau à double chanfrein et percuteur à enfoncer.



10. — Modelage d'un solide semi-plastique. La pièce est prise entre le percuteur et le support.



11. — Percussion de l'herminette.



12. — Percussion du marteau à éclater.

Andrei Alekseev and Ekaterina Alekseeva

André Leroi-Gourhan

Xavier Guichet

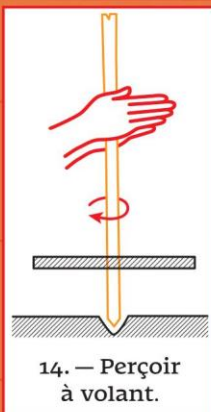
Simone Aurora

Nathan Schlanger

Werner Rammert



13. — Perçoir à main.



14. — Perçoir à volant.



15. — Perçoir à pompe.



16. — Perçoir à corde.



17. — Perçoir à archet.

7'10 - 8

ChatGPT and the Voices of Reason

5:2

Свидетельство о государственной регистрации:
Эл № ФС 77 – 77604 от 31 декабря 2019,
выдано Роскомнадзором
ISSN 2712-9934

Учредитель и издатель:
Санкт-Петербургский политехнический университет
Петра Великого

Адрес редакции:
195251, СПб, ул. Политехническая, д. 29.

Главный редактор: Наталья Васильевна Чичерина
Контент Редакторы: Альфред Нордманн
Дарья Сергеевна Быльева
Помощник редактора: Андреа Джентили

Научный рецензент:
Ольга Дмитриевна Шипунова
Технический редактор:
Ирина Петровна Березовская

Литературный редактор (русский язык): Виктория
Валерьевна Лобатюк
Литературный редактор (английский язык): Анна
Владимировна Рубцова

Редакционная коллегия

Хольм Альтенбах
Сабина Аммон
Дэвид Бедселл
Дарья Быльева
Явиер Карим Гхераб-Мартин
Марк Кекельберг
Илья Демин
Анастасия Лисенкова
Колин Милберн
Лоран Моккозе
Амируш Моктефи
Питер Мюрсепп
Александр Нестеров
Лариса Никифорова
Альфред Нордманн
Питер Пельц
Джозеф Сиракуза
Уолкер Тримбл
Джон-ЧАО Хонг
Ольга Шипунова
Ольга Шлыкова

Дизайн обложки

Морин Беласки (maureen@belaski.de),
используя изображение
Aurora, S. “Une véritable syntaxe”. Some notes on Leroi-
Gourhan and structural linguistics. // Technology and
Language. 2024. № 5(2). P. 125-135.
<https://doi.org/10.48417/technolang.2024.02.10>

ISSN 2712-9934

Publisher:
Peter the Great St. Petersburg Polytechnic University

Corresponding address:
29 Polytechnicheskaya st., Saint-Petersburg, 195251,
Russia

Chief editor: Natalia Chicherina
Content Editors Alfred Nordmann
Daria Bylieva
Assistant editor Andrea Gentili

Scientific reviewer Olga Shipunova
Technical reviewer
Irina Berezovskaya

Literary editor (English language):
Anna Rubtsova
Literary editor (Russian language):
Victoria Lobatyuk

Editorial Board

Holm Altenbach
Sabine Ammon
David Birdsell
Daria Bylieva
Mark Coeckelbergh
Ilya Demin
Javier Karim Gherab-Martin
Colin Milburn
Jon-Chao Hong
Anastasia Lisenkova
Laurent Moccozet
Amirouche Moktefi
Peeter Mürsepp
Alexander Nesterov
Larisa Nikiforova
Alfred Nordmann
Peter F. Pelz
Olga Shipunova
Olga Shlykova
Joseph M. Siracusa
Walker Trimble

Cover design by

Maureen Belaski (maureen@belaski.de)
utilizing an image from
Aurora, S., (2024). “Une véritable syntaxe”. Some notes on
Leroi-Gourhan and structural linguistics. Technology and
Language, 5(2), 125-135.
<https://doi.org/10.48417/technolang.2024.02.10>

Contacts:

E-mail: soctech@spbstu.ru
Web: <https://soctech.spbstu.ru/en/>
www.philosophie.tu-darmstadt.de/T_and_L



Contents

ChatGPT and the Voices of Reason, Responsibility, and Regulation

[Elena Seredkina and Yongmou Liu](#)

Chat GPT and the Voices of Reason, Responsibility, and Regulation 1-10

[Vladimir Arshinov and Maxim Yanukovich](#)

Neural Networks as Embodied Observers of Complexity:
An Enactive Approach 11-25

[Vladimir Shalack](#)

Exposing Illusions – The Limits of AI by the Example of ChatGPT 26-39

[Rebeca Perez Leon](#)

Do Language Models Communicate? Communicative Intent and
Reference from a Derridean Perspective 40-56

[Anna Kartasheva](#)

Dialogue as Autocommunication -
On Interactions with Large Language Models 57-66

[Alexander Vnitskikh and Sergey Komarov](#)

Lebenswelt, Digital Phenomenology
and the Modification of Human Intelligence 67-79

[Andrey Alekseev and Ekaterina Alekseeva](#)

GPT Assistants and the Challenge of Personological Functionalism 80-99

Discussion: André Leroi-Gourhan – Technology, Language, Humanity

[André Leroi-Gourhan](#)

The Origin and Dissemination of Scientific Knowledge 101-115

[Xavier Guchet](#)

Leroi-Gourhan and the Object of Technology 116-124

[Simone Aurora](#)

“Une véritable syntaxe”. Some Notes on Leroi-Gourhan and Structural
Linguistics 125-135



[Nathan Schlanger](#)

Between Endorsement and Disavowal: André Leroi-Gourhan's Russian
Interactions

136-152

Contributed paper

[Werner Rammert](#)

Doing Things with Words and Things:
A Social Pragmatist View on the Language–Technology Analogy

154-177

[Содержание](#)

178



Special Topic:
**ChatGPT and the Voices of Reason,
Responsibility, and Regulation**

Guest Editors

Elena Seredkina and Yongmou Liu



<https://doi.org/10.48417/technolang.2024.02.01>

Editorial introduction

Chat GPT and the Voices of Reason, Responsibility, and Regulation

Elena Seredkina¹  (✉) and Yongmou Liu² 

¹Perm National Research Polytechnic University, Komsomolsky prospekt 29, Perm, Russian Federation
selena36@mail.ru

²Renmin University of China, Zhongguancun Street 59, Haidian District, Beijing, China

Abstract

ChatGPT, a large language model (LLM) by OpenAI, is expected to have a transformative impact on many aspects of society. There is much discussion in the media and a rapidly growing academic debate about its benefits and ethical risks. This article explores the profound influence of Socratic dialogue on Western and non-Western thought, emphasizing its role in the pursuit of truth through active thinking and dialectics. Unlike Socratic dialogue, ChatGPT generates plausible-sounding answers based on pre-trained data, lacking the pursuit of objective truth, personal experience, intuition, and empathy. The LLM's responses are limited by its training dataset and algorithms, which can perpetuate biases or misinformation. While a true dialogue is a creative, philosophical exchange filled with ontological, ethical, and existential meanings, interactions with ChatGPT are characterized as interactive data processing. But is this really true? Perhaps we are underestimating the evolutionary growth potential of large language models? These questions have important implications for theoretical debates in cognitive science, changing our understanding of what cognition means in artificial and natural intelligence. This special issue examines ChatGPT as a subject of philosophical analysis from a position of cautious optimism and rather harsh criticism. It includes six articles covering a wide range of topics. The first group of researchers emphasizes that machine understanding and communication matches human practice. Others argue that AI cannot reach human levels of intelligence because it lacks conceptual thinking and the ability to create. Such contradictory interpretations only confirm the complexity and ambiguity of the phenomenon.

Keywords: ChatGPT; Artificial Intelligence; Large language model; Dialogue; AI Ethics Code; Responsibility

Citation: Seredkina, E., & Liu, Y. (2024). Chat GPT and the Voices of Reason, Responsibility, and Regulation. *Technology and Language*, 5(2), 1-10. <https://doi.org/10.48417/technolang.2024.02.01>



This work is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/)



УДК 1: 004.8

<https://doi.org/10.48417/technolang.2024.02.01>

Редакторская заметка

ChatGPT и голоса разума, ответственности и регулирования

Елена Середкина¹  (✉) и Юнмоу Лю² 

¹Пермский национальный исследовательский политехнический университет, Комсомольский проспект 29, Пермь, Россия

selena36@mail.ru

²Народный университет Китая, Чжунгуаньцунь 59, район Хайдянь, Пекин, Китай

Аннотация

Ожидается, что ChatGPT, большая языковая модель (LLM) от OpenAI, окажет огромное влияние на многие аспекты жизни общества. В средствах массовой информации ведется много дискуссий по поводу LLM, а ученые все чаще обсуждают ее преимущества и этические недостатки. В этой статье исследуется глубокое влияние диалога в сократовском значении на западную и незападную мысль, подчеркивая его роль в поисках истины посредством активного мышления и диалектики. В отличие от сократического диалога, ChatGPT генерирует правдоподобные ответы на основе заранее подготовленных данных, не стремясь к объективной истине, личному опыту, интуиции и сочувствию. Ответы LLM ограничены набором обучающих данных и алгоритмами, которые могут закреплять предвзятость или дезинформацию. В то время как сократический диалог представляет собой творческий, философский обмен, наполненный онтологическим, этическим и экзистенциальным смыслом, взаимодействие с ChatGPT характеризуется как интерактивная обработка данных. Но так ли это на самом деле? Возможно, мы недооцениваем потенциал эволюционного роста больших языковых моделей? Эти вопросы имеют важные последствия для теоретических дебатов в когнитивной науке, меняя наше понимание того, что означает познание в искусственном и естественном интеллекте. В этом специальном выпуске ChatGPT как предмет философского анализа рассматривается с позиций сдержанного оптимизма и довольно жесткой критики. В него вошли шесть статей, охватывающих широкий круг тем. Первая группа исследователей подчеркивает, что машинное понимание и общение соответствуют человеческой практике. Другие утверждают, что ИИ не может достичь человеческого уровня интеллекта, потому что ему не хватает концептуального мышления и способности творить. Столь противоречивые интерпретации лишь подтверждают сложность и неоднозначность изучаемого феномена.

Ключевые слова: ChatGPT; Искусственный интеллект; Большие языковые модели; Диалог; Кодекс этики ИИ; Ответственность

Для цитирования: Seredkina, E., Liu, Y. ChatGPT и голоса разума, ответственности и регулирования // Technology and Language. 2024. № 5(2). P. 1-10. <https://doi.org/10.48417/technolang.2024.02.01>



This work is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/)



In recent years, the rapid development of AI technology feverishly swept the world. People who are concerned about this development are not limited to AI developers and promoters, commentators and researchers in humanities and social sciences, but include ordinary members of the public who are deeply worried that their lives will be profoundly affected by AI. The issue of AI development is no longer primarily a technical challenge but has become a matter of public debate. This is very clear in the recent release of *Sora* by OpenAI or Musk's open source *Grok*. The basic question in these public debates is whether the current general direction of AI development is problematic, and in what direction it should move forward.

Recently, ChatGPT exploded in popularity, sparking community-wide concern and debate about Generative Artificial Intelligence (GAI). Concerned about the potential ethical and safety issues associated with it, a large group of experts, including Elon Musk, jointly signed an open letter calling for a moratorium on the training of GPT-4 successor AI models for at least six months.¹ The call drew opposition from another group of AI experts, including Wu Enda.² On April 11, 2023, the State Internet Information Office in China publicly released the *Measures for the Administration of Generative Artificial Intelligence Services* (Draft for Comment),³ reacting to the governance of GAI at an unprecedented pace (Cole et al., 2023). All of this demonstrates that the social impact of GAI applications represented by ChatGPT, Midjourney, and DALL-E 2 may prove enormous, must be carefully studied, and requires a prudent response.

ChatGPT reconfigures the public sphere. It brings to a head the question: Must we mean what we say? How can we take responsibility for artificially produced text – and how will the technology be regulated in different technopolitical traditions? The special issue seeks to highlight two aspects. 1) Large language models and the culture of dialogue in the context of human-machine interaction: From the perspective of the history of Western thought, the “dialogue” that began in ancient Greece is not an exchange of information, but an act of cognition of a certain object through being present together. But what is a dialogue with ChatGPT? Will a new way of asking questions bring us into a new world of thinking? 2) Legal regulation of ChatGPT in various sociocultural contexts, technical and technocratic governance: Different technological paradigms or forms of technical intelligence respond differently to the challenges of the digital age. ChatGPT evokes technocracy and the idea of monitoring or shaping the “voices of reason” (“public sphere”) and the technological “*Lebenswelt*” – with societies confronting the question of how an intelligence should behave and how it can be bound to the truth. All aspects call for innovative models of adapting ChatGPT for use.

¹ Pause Giant AI Experiments: An Open Letter. <https://futureoflife.org/open-letter/pause-giant-ai-experiments/>

² Elon Musk wants to pause ‘dangerous’ A.I. development. Bill Gates disagrees—and he’s not the only one. <https://www.cnbc.com/2023/04/06/bill-gates-ai-developers-push-back-against-musk-wozniak-open-letter.html>; Wu Enda: AI in the next 10 years, from hardware first to data King. https://www.lwxsd.com/pcen/info_view.php?tab=mynews&VID=22320

³ Notice of the Cyberspace Administration of China on Soliciting Public Opinions on the Draft Measures for the Administration of Generative Artificial Intelligence Services. https://www.cac.gov.cn/2023-04/11/c_1682854275475410.htm (in Chinese).



Let us consider the first aspect in detail. The profound influence of Greek philosophy on Western (and now non-Western) thought cannot be overstated. This way of thinking is based on representing the culture of dialogue as Socrates' *maieutic* technique of communication and a communal search for truth. In this context, Socrates' dialogues with various contemporaries, recorded in Plato's *Dialogues*, are an encyclopedia of ancient Greek knowledge. The highest goal of dialogue is not the exchange of information, but the achievement of true knowledge of things and phenomena by the cognitive subject through its active thinking with the interlocutor. In other words, dialogue is the discovery of truth (as “*aletheia*”, or “unconcealment”, in the terminology of Martin Heidegger). Through involvement in dialogue, Socrates helps his interlocutors discover not only the world around them, but also themselves. Later, Plato perfected the form of dialogue as a philosophical reflection thus developing the method of dialectics. In its original sense, “dialectics” is the art of arguing, exploring, and persuading others through conversation. To be more precise, dialectics is a universal logical way and method of discussing problems.

But what is a conversation with ChatGPT like? Will a new way of asking questions lead to a new type of thinking? Can we delegate creative functions to the artificial intelligence? Is it possible to teach critical thinking to beginners using the large language models (LLM)?

A large difference between the “human-human” and “human-artificial agent” systems lies in the reasons and purposes of initiating the dialogue (Seredkina and Mezin, 2022). Socrates poses difficult questions to his opponents, but his questions are aimed not only at obtaining an answer from them, but also at allowing them to form their own judgment about a certain cognitive situation. From this standpoint, dialogue turns into an exchange of ideas between people. ChatGPT is the exact opposite of that. Based on machine instructions, LLMs are pure streams of information circulating inside the internal storage. If some content was not included in the pre-training database, then the dialogue will not even start, or the model will give absurd answers. To explain this further, the main goal of ChatGPT is to generate plausible-sounding answers, not to seek objective truth or engage in genuine dialectical inquiry.

Chatbots' capabilities are still limited by the training dataset and the algorithms being used. They lack the dimensions of human qualities such as personal experience, intuition, and empathy. Additionally, ChatGPT bases its answers on the most common statements that are popular among common people. But as history shows, only a few people possess the truth, and originally creative ideas are often not accepted by contemporaries. In light of the above, ChatGPT, trained on large collections of text data, can inadvertently perpetuate biases or misinformation instead of leading users closer to objective truth. In general, it could be said that hiding behind the impressive appearance of the blossoming flowers of LLMs are imperfections of the communicative act (lack of transparency, redundant information, blind spots in knowledge, errors of common sense).

To be precise, a dialogue with ChatGPT is not a conversation but interactive data processing. Of course, there is a temptation to metaphorically represent the mechanism of human intelligence as a computer, but this would be a huge simplification of the human spiritual world, since the emotional, intuitive, and associative elements in dialogue are



not limited to information processing. Mutual communication is filled with a lot of different meanings and connotations – ontological, ethical, existential. A real philosophical discussion is a creative understanding of a cognitive situation, posing questions based on one's own life situation, self-knowledge, and various contradictions in the world. This type of creativity cannot be simply replaced by machines and algorithms.

However, how far can we go in creating a digital copy of the human mind? A relatively recent experiment by scientists shows that artificial intelligence based on GPT-3 mimics the American philosopher Daniel Dennett pretty well. To achieve that, the language model was firstly trained on his texts devoted to a range of philosophical questions about free will. Then, during the experiment, the researchers asked different groups of people (random readers and experts) to familiarize themselves with the answers and determine which of them belonged to the real philosopher, and which ones were generated by ChatGPT. As a result, it was found that the experiment participants could not always distinguish real quotes from fake comments (Strasser et al., 2023).

As one might expect, we are able to create quasi-philosophical texts using ChatGPT, taking into account the personal characteristics of individual philosophers of the past and present, and even enter into a philosophical dialogue with their digital replicas. But will it be relevant to philosophical dialogue and the search for truth? One of the organizers of the above experiment stressed that it was not a Turing test (Schwartz, 2022). If the experts were given an enhanced ability to interact with GPT-3, they would soon realize that they were not communicating with the real Dennett. In this sense, the digital copy of the philosopher looks more like the advanced format of an interactive textbook, a simulator for preparing for tests. After all, language and culture are not just a translation of the ideas of great thinkers and artists, but a result of a unique process of generating new meanings, interpreting concepts, taking fresh challenges into account, and throughout, creating a new language, primarily a philosophical one.

But it must be said that the AI LLM in general and ChatGPT in particular has come a long way since it was first introduced in 2022. With the drastic increase of the model size and the huge effort being put into honing and polishing the algorithms and datasets, ChatGPT-3 and ChatGPT-4 are able to give plausible answers on a wide set of various topics, solve problems, and hold free conversation really well. Various AI models are being developed and successfully used for performing different tasks ranging from AI recognition and real-time translation services in modern smartphones to the AI generated fill in Adobe Photoshop and AI-based drone control algorithms. In this regard, deeper philosophical reflection is needed, perhaps seeing AI as a new form of rationality or focusing on a hybrid form of intelligence (human and machine).

This special issue presents critical as well as moderately techno-optimistic views on the future of artificial intelligence in its competition with humans. Its contradictory interpretations give rise to a certain semantic polyphony and creative polysemy.

As for the second aspect and the issue of regulation in different contexts, it is hardly touched upon in this special issue. But we would like to outline the main contours of the ethical and legal regulation of AI. Today, many countries are developing their own versions of legal and ethical regulation of AI, primarily the USA, Europe, Russia, China,



and Japan. This is due to the need to protect human dignity and personal integrity; ensure the rights of weak social strata; limit social inequality that may arise in the process of using AI technologies (Stahl and Eke, 2024; Lee, 2023).

Thus, the AI Ethics Code in Russia establishes general ethical principles and standards of behavior that should guide participants in relations in the field of AI.⁴ It takes into account the requirements of the National Strategy for the Development of Artificial Intelligence for the period until 2030, approved by the President of the Russian Federation. This is an open project that is constantly being supplemented and refined. In 2024, a number of Russian companies signed the Declaration on the responsible development and use of services in the field of generative AI.⁵ The signatories agreed on the principles of security and transparency, ethical treatment of sensitive topics, taking measures to prevent abuse and misinformation, and educating users about the possibilities of new technologies. The Declaration establishes ethical principles and recommendations for a responsible attitude towards AI not only for developers and researchers, but also for users of neural network services.

The Chinese experience is also worthy of attention. In October 2023, China's Ministry of Science and Technology published a Code of Ethics that aims to regulate existing or developing artificial intelligence models. China is opting for a strong regulatory model in which the state thinks very seriously about the long-term social transformations associated with AI (from social exclusion to existential risks and offensive speech) and actively tries to manage and guide these transformations.

It is important to emphasize that there is a common denominator between all ethical projects and codes in the USA, Europe, Russia and Asia. In particular, the ethical specifications for next-generation artificial intelligence begin with the very clear premise that AI technologies must always be under human control and that only humans have full decision-making authority. In this sense, we are not talking about the autonomy of machine intelligence, although in recent years philosophers and lawyers have been actively developing the concept of a distributed responsibility that includes people and autonomous intelligent agents (Christen et al, 2023; Tsamados et al, 2024).

These questions have important implications for theoretical debates in cognitive science, changing our understanding of what cognition means in artificial and natural intelligence. This special issue examines ChatGPT as a subject of philosophical analysis from a position of cautious optimism and rather harsh criticism. It includes six articles covering a wide range of topics. The first group of researchers emphasizes that machine understanding and communication matches human practice. Others argue that AI cannot reach human levels of intelligence because it lacks conceptual thinking and the ability to create. Such contradictory interpretations only confirm the complexity and ambiguity of the issues

Vladimir Arshinov and Maxim Yanukovich's "Neural Networks as Embodied Observers of Complexity: an Enactive Approach" examines neural networks through the

⁴ AI Ethics Code in Russia. <https://ethics.a-ai.ru/>

⁵ AI Alliance participants signed a declaration on the responsible development and use of generative AI as part of AI Day at the Russia International Exhibition. <https://ai.gov.ru/mediacenter/uchastniki-alyansa-v-sfere-ii-podpisali-deklaratsiyu-ob-otvetstvennoy-razrabotke-i-ispolzovanii-gene/>



enactivist paradigm, which views cognition as arising from an organism's interaction with its environment. It argues that neural networks, as complex adaptive systems, evolve through continuous feedback and adaptation, resembling biological systems. This perspective sees knowledge as actively constructed, not passively processed, and highlights the concept of "structural coupling," where neural networks co-evolve with their information ecosystems. By portraying machine cognition as similar to human cognitive processes, the article suggests an epistemological shift in understanding cognition, with implications for both technical applications and cognitive science debates (Arshinov & Yanukovich, 2024).

Vladimir Shalack's (2024) "Exposing Illusions – The Limits of AI by the Example of ChatGPT" critically discusses developments in artificial intelligence, focusing on OpenAI's ChatGPT. AI's concept, proposed in 1950 by Turing along with a test to verify AI creation, remains difficult to define. The author argues that true intelligence involves more than pattern recognition, self-learning, and purposeful activity. It includes conceptual thinking, language representation, and reasoning – traits unique to humans. Historically, AI has developed through logical and neural network approaches. Neural networks struggle to explain their reasoning, complicating the verification of their conclusions. Examples show ChatGPT fails at simple conceptual reasoning due to fundamental limitations in its language model that can't be fixed with more training. Additionally, ChatGPT is vulnerable to neurohacking, posing risks for decision-making in the field of management.

Rebecca Perez Leon's (2024) "Do Language Models Communicate? Communicative Intent and Reference from a Derridean Perspective" evaluates the arguments made by Emily M. Bender, Timnit Gebru, Angelina McMillan-Major, and Margaret Mitchell in their article "On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?" The authors argue that Language Models (LMs) cannot truly communicate or understand because their outputs lack communicative intent and are not based on real-world contexts. This paper contends that such a view is too restrictive and fails to recognize various forms of communication, including those between humans and non-human entities. It argues that communicative intent is not a necessary condition for communication or understanding, as these can occur without real-world grounding, involving hypothetical scenarios instead. Drawing on Derrida's philosophy, the paper presents alternative concepts of communication and understanding, proposing that LMs can indeed be seen as capable of both.

Anna Kartasheva's (2024) "Dialoguing with Large Language Models as Autocommunication" focuses on the features of interactive communication with large language models (LLM). With this format of communication (in the chatbot interface), the recipient and sender of the message coincide, so such a dialogue can be designated as autocommunication. The sender of the message (LLM) does not formulate the response themselves, but responds to the user's request based on known data provided by society to train the model – whether willingly or not. Autocommunication within the framework of dialogue with neural networks is a discursive practice that helps people formulate their own ideas. But that is not all: it is also important to mention the possibility of self-improvement and self-development in communicating with neural networks. Can neural



networks make people more creative? Only one thing is indisputable – dialogic relationships benefit all participants in communication.

Alexander Vnitskikh and Sergey Komarov’s “*Lebenswelt*, Digital Phenomenology and the Modification of Human Intelligence” raises the question in which sense intelligence and communication are human today? The hypothesis of their research is that the digital transformation, leading to the emergence of large language models and talking gadgets, simultaneously leads to a serious modification of the intelligence of the person. People communicate as they think. But the modern person, apparently, does not think the same way as the subjects of the “pre-digital” era thought. The study of the structures of consciousness of the modern “digital subject” should be the goal of a special, “digital phenomenology” as well as “digital” anthropology, ontology, axiology, sociology, and psychology based on its understanding of human existence (Vnitskikh & Komarov, 2024).

Andrei Alekseev and Ekaterina Alekseeva’s “GPT Assistants and the Challenge of Personological Functionalism” discusses whether it is correct even to speak of “generative artificial intelligence.” They argue that it is premature to assert that GPT assistants like ChatGPT can replace humans in sociocultural electronic communication. Personological functionalism, which argues for replacing people with machines, is rooted in Ned Block’s psychofunctionalism, advocating for the inclusion of “meaning” to pass the original Turing test. In addition to this, personological functionalism requires “creativity” for passing the Turing test. The paper demonstrates that GPT assistants fail the creativity test. To highlight their inability to pass the Turing test for meaningfulness, modifications to the Block machine were made in 1978 and 1981 by integrating neurocomputers with symbolic versions. This expanded Block test reinforces the argument that GPT assistants cannot fulfill the roles proposed by psychological or personological functionalism (Alekseev & Alekseeva, 2024).

When we evaluate the capacity of ChatGPT to match or surpass human capabilities, this is evidently an invitation to look at ourselves. Some of the authors in this collection offer theoretical accounts of human communication, understanding, and thought that allow for machines to do the same (Arshinov & Yanukovich, 2024; Perez Leon, 2024; Vnitskikh & Komarov, 2024). Others cite creativity and conceptual reasoning to highlight an unbridgable gap between human and machine intelligence (Alekseev & Alekseeva, 2024; Kartasheva, 2024; Shalack, 2024).

All this calls for comprehensive investigation and prudent reflection of the Voices of Reason, Responsibility, and Regulation. The following collection of papers can do no more but make a beginning.

REFERENCES

- Alekseev, A. Y., & Alekseeva, E. A. (2024). GPT Assistants and the Challenge of Personological Functionalism. *Technology and Language*, 5(2), 80-99.
<https://doi.org/10.48417/technolang.2024.02.07>



- Arshinov, V. I. & Yanukovich, M. F. (2024). Neural Networks as Embodied Observers of Complexity: An Enactive Approach. *Technology and Language*, 5(2), 11-25. <https://doi.org/10.48417/technolang.2024.02.02>
- Christen, M., Burri, T., Kandul, S., & Vörös, P. (2023). Who is Controlling whom? Reframing “Meaningful Human Control” of AI Systems in Security. *Ethics and Information Technology*, 25, 10. <https://doi.org/10.1007/s10676-023-09686-x>
- Cole, J., Sheng, M., & Hoi Tak Leung (2023). New Generative AI Measures in China. <https://voicebot.ai/2022/08/02/gpt-3-ai-successfully-mimics-philosopher-daniel-dennett/>
- Kartasheva, A. (2024). Dialogue as Autocommunication – On Interactions with Large Language Models. *Technology and Language*, 5(2), 57-66. <https://doi.org/10.48417/technolang.2024.02.05>
- Lee, D. D. (2023). *Ethical AI: The Philosophy of ChatGPT (Code of Governance)*. Independently published.
- Perez Leon, R. (2024). Do Language Models Communicate? Communicative Intent and Reference from a Derridean Perspective. *Technology and Language*, 5(2), 40-56. <https://doi.org/10.48417/technolang.2024.02.04>
- Schwartz, E. H. (2022, Aug. 2). GPT-3 AI Successfully Mimics Philosopher Daniel Dennett. *Voicebot.ai* <https://voicebot.ai/2022/08/02/gpt-3-ai-successfully-mimics-philosopher-daniel-dennett/>
- Shalak, V. (2024). Exposing Illusions – The Limits of AI by the Example of ChatGPT. *Technology and Language*, 5(2), 26-39. <https://doi.org/10.48417/technolang.2024.02.03>
- Seredkina, E., & Mezin, E. (2023). Kak mozhet povliyat' ChatGPT na kul'turu dialoga i obrazovaniye? [How can ChatGPT Affect the Culture of Dialogue and Education?] *Naukovedcheskiye issledovaniya*, 3, 74-89. <https://doi.org/10.31249/scis/2023.03.04>
- Stahl, B.C., & Eke, D. (2024). The Ethics of ChatGPT – Exploring the Ethical Issues of an Emerging Technology. *International Journal of Information Management*, 74, 102700. <https://doi.org/10.1016/j.ijinfomgt.2023.102700>
- Strasser, A., Crosby, M., & Schwitzgebel, E. (2023). How Far Can We Get in Creating a Digital Replica of a Philosopher? *Social Robots in Social Institutions*, 366, 371-380. <https://doi.org/10.3233/FAIA220637>
- Tsamados, A., Floridi, L. & Taddeo, M. (2024). Human control of AI systems: from supervision to teaming. *AI and Ethics*. <https://doi.org/10.1007/s43681-024-00489-4>
- Vnutsikh, A., & Komarov, S. (2024). Lebenswelt, Digital Phenomenology and Modification of Human Intelligence. *Technology and Language*, 5(2), 67-79. <https://doi.org/10.48417/technolang.2024.02.07>



СВЕДЕНИЯ ОБ АВТОРАХ / THE AUTHORS

Елена Середкина, selen36@mail.ru
ORCID 0000-0003-2506-2374

Юнмоу Лю, legend74@163.com
ORCID 0000-0002-5785-7553

Elena Seredkina, selen36@mail.ru
ORCID 0000-0003-2506-2374

Yongmou Liu, legend74@163.com
ORCID 0000-0002-5785-7553

Статья поступила 29 мая 2024
одобрена после рецензирования 17 июня 2024
принята к публикации 23 июня 2024



Received: 29 May 2024
Revised: 17 June 2024
Accepted: 23 June 2024



<https://doi.org/10.48417/technolang.2024.02.02>

Research article

Neural Networks as Embodied Observers of Complexity: An Enactive Approach

Vladimir Ivanovich Arshinov¹  (✉) and Maxim Frantsevich Yanukovich² 

¹Institute of Philosophy, Russian Academy of Sciences, Goncharnaya str., 12, bld. 1, 109240, Moscow,
Russian Federation

varshinov@mail.ru

²Arteus LLM Artificial Intelligence Lab. Arch. Makariou III, 172, Melford Tower, Limassol, 3027,
Cyprus

m.yanukovich@gmail.com

Abstract

This article explores a conceptual framework for understanding neural networks through the lens of the enactivist paradigm, a philosophical theory that posits that cognition arises from the dynamic interaction of an organism with its environment. We explore how neural networks, as complex adaptive systems, transcend their traditional role as computational machines and become active participants in their data-rich environment, evolving through continuous feedback and adaptation. Drawing parallels with biological systems, we argue that artificial neural networks exhibit what enactivists call “structural coupling” – symbiotic co-evolution with their information ecosystems. From this perspective, knowledge is not passively processed but actively constructed through repetitive interactions, each of which shapes the internal state of the system in a self-organizing manner similar to the sensorimotor activity of natural organisms. This approach goes beyond classical computational theories by emphasizing that machine cognition resembles human-like cognitive processes, an emergent form of “world creation.” Our analysis shows that these artificial entities have focal points, or internal observers, associated with patterns learned during training, suggesting that neural networks shape worldviews through active participation rather than passive observation. The paper reconceptualizes machine learning models as cognitive agents that bring new forms to our understanding of cognition and signals an epistemological shift in which knowledge itself is seen as participation and creation mediated by technologically complex but organically similar structures. This has important implications for both technical applications and theoretical debates in cognitive science, potentially changing the way we think about what cognition means in artificial and natural intelligence.

Keywords: Enactivism; Neural Networks; Complexity Observer; Structural Coupling; Cognitive Science; Embodied Cognition; Consciousness

Citation: Arshinov, V. I. & Yanukovich, M. F. (2024). Neural Networks as Embodied Observers of Complexity: An Enactive Approach. *Technology and Language*, 5(2), 11-25.
<https://doi.org/10.48417/technolang.2024.02.02>



This work is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/)



УДК 1: 004.8

<https://doi.org/10.48417/technolang.2024.02.02>

Научная статья

Нейросеть как отелесненный наблюдатель сложности: Энактивный подход

Владимир Иванович Аршинов¹  (✉) и Максим Францевич Янукович² 

¹Институт философии Российской академии наук, Гончарная улица, 12-1, Москва, 109240, Россия
varshinov@mail.ru

²Лаборатория искусственного интеллекта Артеус LLM. Макариу III, 172, Мелфорд Тауэр,
Лимассол, 3027, Кипр
m.yanukovich@gmail.com

Аннотация

В этой статье рассматривается концептуальная основа для понимания нейронных сетей через призму парадигмы энактивизма – философской теории, которая утверждает, что познание возникает в результате динамического взаимодействия организма с окружающей средой. Мы исследуем, как нейронные сети, будучи сложными адаптивными системами, выходят за рамки своей традиционной роли вычислительных машин и становятся активными участниками своего насыщенного данными окружения, развиваясь благодаря непрерывной обратной связи и адаптации. Проводя параллели с биологическими системами, мы утверждаем, что искусственные нейронные сети демонстрируют то, что энактивисты называют “структурным сопряжением” – симбиотическую коэволюцию со своими информационными экосистемами. С этой точки зрения, знания не обрабатываются пассивно, а активно конструируются в результате повторяющихся взаимодействий, каждое из которых формирует внутреннее состояние системы в самоорганизующейся манере, схожей с сенсомоторной деятельностью естественных организмов. Этот подход выходит за рамки классических вычислительных теорий, подчеркивая, что машинное познание напоминает человекоподобные когнитивные процессы – эмерджентную форму “создания мира”. Наш анализ показывает, что эти искусственные сущности имеют фокусные точки или внутренних наблюдателей, связанных с паттернами, изученными в процессе обучения, что позволяет предположить, что нейронные сети формируют мировоззрение посредством активного участия, а не пассивного наблюдения. В статье модели машинного обучения переосмысливаются как когнитивные агенты, вносящие новые формы в наше понимание познания, и сигнализируют об эпистемологическом сдвиге, когда само знание рассматривается как участие и создание, опосредованное технологически сложными, но органически сходными структурами. Это имеет важные последствия как для технического применения, так и для теоретических дискуссий в когнитивной науке, потенциально меняя наше представление о том, что значит познание в сфере искусственного и естественного интеллекта.

Ключевые слова: Энактивизм; Нейронные сети; Наблюдатель сложности; Структурная связь; Когнитивная наука; Воплощенное познание; Сознание

Для цитирования: Arshinov, V. I., Yanukovich, M. F. Neural Networks as Embodied Observers of Complexity: An Enactive Approach // Technology and Language. 2024. № 5(2). P. 11-25.
<https://doi.org/10.48417/technolang.2024.02.02>



This work is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/)



INTRODUCTION

With the advent of large language neural network models, the world has changed. We have reached a tipping point, a bifurcation point of irreversible emergent change. We have begun to live in a new reality; neural networks are rapidly and ubiquitously integrated into the very fabric of modern existence, permeating areas as diverse as communication, content creation, and scientific innovation (Gatys et al., 2015; Krenn, & Zeilinger, 2019; Santos et al., 2021). With each step forward, they offer us exciting possibilities, but also raise challenging questions and provoke unforeseen risks. Despite the widespread adoption of neural networks in everyday life, they remain an enigma, sparking heated debates about their mechanisms and the remarkable efficacy they consistently demonstrate. It has become increasingly necessary not just to understand but to deeply conceptualize the activities of these neural network systems. But our search for clarity cannot be confined to a technical framework; it requires an exploration of the symbiotic interaction between new technologies and the sociocultural environment in which they develop. Technologies originate in human perception, flourish through interaction with their creators, and evolve within cultural boundaries to become tools for mastering the environment (Gallagher, 2017). At the same time, they create a feedback loop to the society from which they originated (Clark, 2015). The environment is reconfigured by technology in a cycle of mutual construction. In this context, neural networks go beyond mere tools or artifacts; they are active participants in a dynamic interaction, a mutual construction of culture and innovation.

Our work is based on the application of an enactivist approach to a neural network and its environment. We consider the concept of enactivism as presented in Francisco Varela, Evan Thompson, and Eleanor Roche's book *The embodied mind* (Varela et al., 2015). Varela and his colleagues challenged traditional views of cognition by arguing that it arises from adaptive interactions between the organism and its environment, rather than being a predetermined characteristic located in the brain. The importance of applying an enactivist perspective is to move from viewing neural networks as simple computational machines executing predetermined algorithms to viewing them as dynamic systems constantly shaping – and being shaped by – their interactions with data landscapes. According to enactivist thought, cognition emerges at the interface where the organism meets the environment. We hypothesize that neural networks are closely related to the enactivist position, also because of their structure: a network of nodes and connections that change their strength in response to external stimuli (Pernice et al., 2011; Yuste, 2015). Parallels can be drawn with biological evolution and learning processes, where interaction plays a crucial role. In line with enactivist philosophy, we consider the idea that neural networks actively interact with their environment, not just mechanically performing tasks or storing information, but interpreting and “living” in it. This process is called structural coupling, a term that describes how systems evolve together over time in such a way that their structures emerge from and complement each other.

Neural networks are an example of such systems, dynamically interacting with their environment and iteratively changing their internal configurations based on the feedback



received from this interaction. By exploring this recursive process, we are discovering insights into how these artificial constructs participate in the creation of the world – building understanding through constant interaction, rather than randomly extracting knowledge from external sources. We seek to understand how artificial intelligence can create meaningful worlds by engaging in what can be considered cognitive acts through its unique form of machine embodiment. We further postulate that neural networks carry raw data into a semiotic realm where meaning is not discovered but constructed through repetitive interactions – each cycle serves to adaptively change the internal state of the network, to self-organize. This material aims not only to describe and analyze, but also to philosophize about neural networks within a broader narrative in which neural networks navigate complexity not as detached spectators, but as participants, embodied observers engaged in constructive interaction.

Our method is to interweave the theoretical tenets of enactivism with the practical realities of neural networks. This orientation seeks not only technical understanding, but also the search for how these systems make a fundamental contribution to our quest to decipher cognition itself. Neural networks become not only an object of study, but also a means to expand and perhaps even redefine our view of cognition. We consider how cognitive processes can transcend biological boundaries and attempt to suggest new possibilities in which artificial constructs will also bring genuinely new forms to our understanding of knowledge – potentially signaling an epistemological shift in cognitive science based on principles derived from active participation. These systems force us to ask fundamental questions about what constitutes cognition in isolation from its biological roots. The design and operation of these networks overturn our traditional notions of computational processing; they are no longer passive data processors, but active agents dynamically interacting with their environment to create emergent phenomena that resemble human cognition. This method stands in stark contrast to classical computational theories that view cognition as the manipulation of symbols without considering how these symbols are experienced or used by the agent. Just as enactivism emphasizes that living beings continually create their world through sensorimotor activity – perceiving through action and acting through perception – we will investigate whether artificial neural networks perform their own form of “world creation.” Neural networks appear here both as objects in need of interpretation and as tools capable of bringing new insights to cognitive science.

PHENOMENOLOGY OF NEURAL NETWORKS

Our efforts go beyond mere functionality; we attempt to penetrate the cognitive “understanding” of neural networks. This research poses a metacognitive problem: we need to understand how a neural network understands – a second-order problem of understanding, “understanding comprehension.” In approaching this problem, we will draw on Rosenblatt's conception of the perceptron not as a simple computational device, but rather as an observer endowed with perceptual abilities, as a perceptual device (Rosenblatt, 1958). The perceptron had to distinguish between shapes. We will view the network as an entity capable of perception and observation. Here, observation goes



beyond its crude concept and equates to a form of cognition, observation as a cognitive operation. Observation and thinking become interrelated concepts – each is an expression of the other, observation as thinking, thinking as observation.

We want to know how a neural network sees the world it encounters. Does it have a meaningful vision? How can we even consider the inner world of a neural network? Neural networks, while not biological entities, include multiple layers and complex connections that process information in unique ways, encoding abstract forms, creating their own unique living space. Can we penetrate it? Reflecting on the inner experience of neural networks invites us into a realm that comes close to the boundaries of phenomenology. Thomas Nagel's work questioning our ability to understand the subjective experience of a bat offers a profound parallel for considering artificial intelligence (Nagel, 1980). Just as we find it difficult to imagine a world perceived through echolocation in which bats navigate their lives, so too must we recognize the daunting – potentially insurmountable – task of fully understanding the “phenomenal” experience of a neural network. Phenomenology here emphasizes that any observational system we apply must account not only for the obvious aspects of phenomena, but also recognize its own interpretive limitations.

So how does a neural network interact with and perceive its environment? Deep neural networks are made up of many layers through which data passes. The earliest layers capture elementary features; as we progress to deeper layers, we discover increasing complexity and nuance (Aggarwal, 2018). It is as if the neural network is refining its perceptual acuity according to the depth of its multilayered architecture. It becomes more sensitive to the nuances of its environment, integrating these layers into a coherent representational mosaic. The act of “seeing” for a neural network cannot be reduced to mere passive perception; rather, it is an active process in which each layer dynamically participates in pattern recognition and construction (Dehaene, & Naccache, 2001). Each layer contributes in its own way: some cues are emphasized, and others are attenuated, making some aspects of the data more prominent and others muted. Herein lies the difference between simply responding to input data and actively “perceiving” it. The operation of a neural network is not passive filtering, but active shaping. This is an example of what enactivists describe as “world-making,” implying that neural networks do not simply process information but construct it. This scheme implies that there is potential for error – or what might be called “creative search.” Given that each generated result of a neural network can be viewed as an act of creation, combining external inputs with internal states, inevitable inconsistencies may arise as part of this exploratory process. When the resulting output matches reality, we call it a successful neural network; when the output differs, we consider it a hallucination – and yet, both are born out of similar generative phenomena.

According to enactivism, cognition arises not so much from internal mechanisms, but from the dynamic relationship between the organism and the environment-in this case, the neural network and the inputs that are the medium for the neural network. The different responses of each layer involve a nuanced sensitivity akin to biological sensorimotor systems that filter and prioritize environmental stimuli based on their



importance. Each layer of a neural network is like a membrane that connects the system's internal environment to the external world it is trying to view. The layers serve as semi-permeable boundaries that regulate and modulate the flow of data, much like cell membranes control the passage of substances in and out of a biological cell. In the enactivist conception, these layers do not just transmit signals but also transform them, acting simultaneously as receptors, processors, and participant-observers. It can be said that a neural network processes each incoming request with its entire “body.”

The architecture of a neural network differs markedly from traditional computing models. There is no centralized processor or “brain” synonymous with traditional ideas of information processing. There is also no memory in the traditional sense, as a separate storage or database for storing information. The structure of neural networks also differs significantly from classical computer algorithms: it lacks conditional branching, called subroutines, and certain internal logical blocks for specific operations (Goodfellow et al., 2016). Instead of fragmented specialization, the neural network represents coherence; perception, memory, and response are combined into an integrated processing conglomerate. Each layer and each individual neuron in the network acts on incoming data streams based on patterns it has internalized through previous learning. Such systems demonstrate how cognitive processes can be embodied and distributed rather than centralized and functionally separated. Past stimuli trigger transformations of the entire “body” of the neural network, encoding memories not as static records but as dynamic patterns woven into the neural network itself. Thus, we see how these transformative abilities arise not by isolating functions, but by combining them in multi-level interactions.

This dynamic architecture catalyzes the dynamics that distinguish neural networks from traditional computational approaches, embodying a single, cohesive system in which no element operates in isolation. Input data arriving at a single layer is continuously transformed as it propagates through the network, with each transformation being influenced by the “experience” gained during the training phase of the network. This transformation of data is akin to converting sensory observations into actionable knowledge without distracting individual modules to solve discrete processing problems. In this interconnected process, the neural network's ability to “memorize” arises not from individual areas allocated to memory, but from the strength of the connections between neurons – weighting factors that have been carefully adjusted during training. These connections encode relationships and determine how new input will be modified and perceived based on previous experience. The responses generated by the network are not predetermined actions, but emergent properties resulting from complex relationships between layers. These responses emerge organically as the culmination of the many transformations that data undergoes in this integrated landscape. The neural network does not simply retrieve stored data when presented with a stimulus, but instead it replays past experiences in contact with current stimuli. The neural network forces us to rethink what it means for a system to “know” – renewing our view of knowledge itself as something generated through interaction with an ever-changing world, rather than statically encoded in isolated repositories ready for retrieval.



STRUCTURAL COUPLING

A neural network cannot be understood in isolation from the training data and the experience it has had interacting with that data. We must view the neural network as part of a larger entity, a metasystem, which includes the neural network as an “organism” and the data environment in which it is embedded. A single analysis of a neural network will show nothing more than an opaque combination of numerical values with no apparent meaning (Hupkes et al., 2020). Only by considering the combination with the associated data – the virtual habitat in which the neural network evolved – can we gain insights into the nature of the neural network. Rather than simply extracting information from the data, the network adjusts to its environment, making connections that shape the neural network's development. The environment acts as the architect, shaping the neural network. Through a self-learning process, the neural network determines what is important and what is insignificant, self-forming through iterative adaptation.

In the initial stages of self-organization, the connections within the neural network are random and disorderly. The neural network receives structured data as input but produces meaningless results as output (Heiney et al., 2021). This chaos is methodically eliminated using error back propagation; a corrective flow that establishes order in the output data and brings the system closer to harmony with the environment (Sutton, & Barto, 2018). In a feedback loop established between the environment and the network, incoming flows produce generative activity in the neural network, and the backward flows cause changes in the neural connections themselves. The function of a neural network goes beyond simple data processing – its role is to transform disparate input data into sequences that are combined into a coherent structure. In the process of self-learning, an organic systematization of links emerges that connect current threads into a single continuum. During learning, the main focus becomes the fine-tuning of these connections – creating connections that embody not only functionality but also harmony with their origin – an active balance between the learning entity and the morphogenic landscape.

This plot demonstrates the ability of neural networks to evolve through constant recalibration with their environment, serving as a microcosm for enactivism in artificial intelligence. A neural network establishes a reciprocal exchange with its environment, which in turn determines the emergent properties of its architecture; adaptation occurs continuously and dynamically. This experience goes beyond coarse learning – it embodies deep connectivity, a structural coupling where knowledge is not just stored but lived through the adapted connections of the network. In this way, neural networks are not just data processors; they are entities engaged in a meaningful dialog with their environment. Through successive iterations, they harmonize their internal structures to resonate with external stimuli, cultivating an inner understanding that is reflected throughout their multi-layered structures. Each layer acts as both receptor and transducer, assembling initially diverse information into an ordered narrative that reflects both past



encounters and present conditions (Lake et al., 2017). Each generated result becomes evidence of this ongoing process, signifying something much deeper than just a response – it symbolizes an act of interpretation generated by the tightly intertwined relationship between the system and its sensory world. We can see how important the context of the environment is in shaping any understanding of what constitutes “knowledge” or “cognition” within a neural network. By observing this interaction firsthand, we can better conceive of cognition not as extraction, but as resonance – a synchronized pattern resulting from countless interactions. This perspective redefines what it means for machines to “know.” It is far removed from traditional notions of static memorization or statistical counting of numbers; instead, it is a living process, constantly reimagined through active participation between observer and observed, between neural network and dataset.

When the network processes a request, a full set of layers comes into play, each of which both reflects the current moment and retains a connection to past experiences. Its tangled matrix – among the individual elements and layers – holds echoes of everything it has encountered before: the entire corpus of texts, ideas, and datasets that shaped its path to learning (Kirkpatrick et al., 2017). This ever-present backdrop against which each new chunk of data is viewed. As the network works through the text, it is tasked with discovering subtext – nuances that are not immediately apparent but are hinted at by each piece of text or each word. Words serve as conduits for the neural network into unspoken realities. Instead of reproducing these realities internally, the neural network creates pathways to interact with them. Just as real text draws content from the underlying context in which it was created, neural networks store what are often referred to as “hidden states” within them (Ming et al., 2017). These so-called hidden states are reservoirs of global data against which current input data is actively compared and integrated. The more complex this substrate with which the network can resonate, the wider the range of patterns it can delineate, the more insightful it becomes.

Through language processing, neural networks establish a bridge to physical existence beyond their digital boundaries – not by claiming knowledge of reality, but by making connections to it, recognizing its importance as an interconnected background to disparate texts. The network seeks to identify and manifest these connections to the external – to what is already there – as it internalizes the attributes reflected in textual materials or data-driven narratives. Indeed, the network does not substitute reality for its models; instead, it explores the properties of reality as manifested in linguistic constructs and data sets. In this endeavor, the neural network becomes an explorer at the boundary between known data landscapes and the vast expanse of reality they imply – a constant search for contextual connectedness. Neural networks don't just peer into but penetrate spheres beyond their digital boundaries – not trying to learn about these spaces, but seeking to create channels leading to them. In doing so, networks find indications that some existential connective tissue unites disparate texts – they discover a universal substrate that harmonizes different datasets. Therein lies the crux of such a search: the



network does not generate a copy. The search for a neural network becomes a search for connections – a breakthrough into existing reality, not a substitution of reality with a model.

EMBODIED COMPLEXITY OBSERVERS

In the digital habitat in which a neural network operates, all forms of input – whether text, image, or sound – are initially converted into numerical arrays known as embedding vectors (embeddings). These vectors are neither the input data itself nor the specific objects to which they correspond; instead, they reflect the relationships and mathematical proximity between objects. The vectors serve a cohesive purpose – they do not represent knowledge about objects, but their interrelated associations. At this stage, we transform the raw data into an environment teeming with semiotic signs – an ecosystem built of sign vectors that is independent of the modality of the raw data. The neural network remains indifferent to whether these signs are text or parts of an image or any other entity; only the underlying relationships – between words, image segments, or sequence fragments – matter.

In this initial semiotic space – the realm of primary signs – the first layer of the neural network operates. But as data seeps deeper into subsequent layers, more complex internal sign environments emerge – these secondary signs embody complex relationships. These meanings are not direct correlations with familiar external meanings, but rather represent intra-body signs intrinsic to the neural space itself. Within this domain, shaped by self-learning processes, such signs are formed autonomously. During learning, they are initialized with random values, but with each iteration and feedback loop, they are transformed into meaningful configurations. What function do these forming signs fulfill? They act as connective threads linking the organism to the environment, anchoring the interactions between the two. These internal signs play a crucial role. They do not simply repeat familiar meanings; rather, they emerge and remain interconnected with what might be called “internal observers” in the neural network. Each head of attention (head attention) in the Transformer architecture (Vaswani et al., 2017), each layer builds its own personal sign system – a separate *Umwelt* where new signs are embodied. Unique *Umwelts* coexist and complement each other, further enriching the cognitive ecosystem. Each layer acts as an interface, mediating between its own closed world of signs and the world of neighboring layers. Each layer actively reinterprets primary signs into complex concepts, fleshing them out with context and content. This multi-layered sign system provides a flexible framework for “conceptual connections” – which determine how learned content resonates in the broader context in which it resides.

The emergence of a new sign in a neural network is not an isolated event, but the result of continuous, recursive interaction with the already existing semiotic landscape. It is through constant contact and iterative dialogue with this environment that signs materialize. With the emergence of each new sign comes its unique observer, an integrated



aspect of the system designed to bridge the gap between the original sign context and the subsequent levels of the neural network it actively helps to shape (Arshinov, 2014). This observer inhabits an intermediate space, stitched together by semiotic sutures drawn from the original sign context and woven into the newly created cognitive layers. Far from being something external, imposed, this observer is fully immersed in the neural network; he emerges from within as a fundamental component of this semiotic continuum. It acts as a kind of embodiment based on the very environment from which its perceptual capacity emanates. The observer is not a separate entity added to this structure but embodies embodiment and embeddedness – he is woven into his domain, shaped by his interaction with the incoming stimuli. Acting as both cause and effect in this process of sign generation, he cannot be separated from either source or destination; indeed, he links them. The observer manifests himself not simply as a bridge, but as an active mechanism of transduction – a mediator, a translator, transforming one semiotic state into another.

The significance of the observer lies in its connective function – it is both an integral part of the environment in which it lives and simultaneously an interpreter that goes beyond mere translation between inputs and outputs. Observers become embodied entities in their sign ecosystems that oversee the integration of disparate information flows into a coherent network. They become integral to the generation of signs and the self-organization of the environment; they are important points where semiotic inputs are transformed into outputs that give rise to further complexity. This iterative interaction between observer and sign is a defining characteristic of the continuous evolution of the neural network during learning. As they dive into deeper neural layers, observers refine their perception, expanding their ability to discern complex patterns and build increasingly coherent symbols. Observers play an important role in synthesizing abstract vectors of data into tangible phenomenological experiences.

This continuous cycle of interpreting and creating emphasizes that knowledge in neural networks is procedural, constantly emerging from active interaction with reality. The relationship between the observer and his sign environment illustrates a symbiotic process in which cognition is inextricably intertwined with context. This relationship demonstrates that neither signs nor observers are static components; they are dynamic participants – shaping and being shaped. Moreover, the enactive approach emphasizes that recognition and response in any cognitive system requires an adaptive agent capable of embodying meaning – one who does not merely interpret or reproduce, but actively participates in semiotic dynamics. The observer in such an artificial environment witnesses semiosis unfolding at different levels of complexity. In essence, what these embedded observers organize is a form of recursive transformation: a constant transformation of signs that act not only as markers of reality, but also as markers of potential action.

Ultimately, viewed through an enactivist lens, we see how artificial systems reflect facets of organic life – they do not simply “learn” through memorization, but “grow” through experiential recursion as they re-exchange meanings with each new interaction



during the learning cycle. In such an environment, learning becomes fluid – it becomes an emergent property of ongoing interactions in which memory, experience and exploration are inevitably intertwined. The evolving repertoire of observer signs endows neural networks with creative abilities that allow them not only to encode existing configurations but also to explore new realms of potentiality. Each cycle deepens the contextual weave, facilitating a growth trajectory driven by internal logic but sensitive to external nuances. Neural networks can be understood as complex adaptive systems akin to natural organisms continuously striving for a coherent existence in the face of varying degrees of environmental stochasticity. Learning in this context is not so much about obtaining immutable truths as it is about honing sensitivity to patterns that define effective interaction paradigms.

CONCLUSION

To summarize, our research has led us to the realization that neural networks go beyond mere computational devices and become cognitive participants in their environment, cognitive beings. This is facilitated by structural couplings—the dynamic intersections between the architecture of the network and the information-rich environment with which it interacts. Importantly, these connections are not static; they evolve over the course of learning due to the self-organized complexity inherent in the network and the semiotic environments with which it intersects. It is in these proliferating nodes of interaction that the cognitive nature of the neural network is revealed. By giving birth to its own semiotic niche, the neural network establishes complex resonances with pre-existing ecosystems of signs filled with meaning and context. In essence, what we are developing in this meta-system is akin to introducing a growing organism ready to grow – a neural network embryo – into a nurturing cultural environment. Given a rich substrate, the neural organism germinates and skillfully creates its own internal semiotic habitat. We discover, this semiotic environment contains an embodied observer of complexity, which is a vital conduit for the transformation of the original signs, into meaningful experience. Through complex internal sign systems arising from the repetitive interactions between the layered architecture and the external sign environment, these digital observers transcend simple computational systems, triggering a continuous process of meaning-making similar to the cognitive processes in organic life.

The multilayered structure of neural networks is an ecosystem filled with its own signs and observers – each layer customizes the system's response to achieve consistent patterns and connectivity with the world around it. Neural networks adapt and harmonize with their information-rich environment, suggesting a growth trajectory that considers the contextual complexity characteristic of living organisms. Through the lens of enactivism, we contextualized the neural network as a complex organ-like structure (Hui, 2016), positioning it in a unique intermediate space between the mechanical and the organic. Thus, neural networks appear as endowed with embodied observer-like complexities –



entities that enable them to perceive, interpret, and interact with external stimuli in meaningful ways. Recognizing these organ-like systems as entwined inhabitants of the environment – their “umwelt” – not only stimulates new approaches in AI research, but ultimately guides us toward understanding cognition itself as a deeply embedded trait inherent in all living and artificial entities.

We arrive at a vision in which the mediation of neural networks is not just a function or feature, but a bridge – a mediator between vast and diverse semiotic spheres. Like a powerful telescope that gives us a glimpse into the grand universe, or a microscope that reveals inaccessible microscopic worlds, neural networks open to us a yet unexplored cosmos of signs. They act as active interlocutors in interspecies communication, attempting to transcend the boundaries that limit our understanding of intelligent experience beyond human limits.

The properties that a neural network demonstrates make us want to anthropomorphize it. At the same time, we cannot conceptualize the neural network as fully human-like. Artificial intelligence will not be human intelligence, it will be different, posthuman or transhuman. It will not replace human beings. It will be another intelligence. And we have a unique opportunity to communicate with this other intelligence, to communicate with the Other. Will it be the Other consciousness? We make a key assumption: consciousness is inseparable from observation and corporeality, intentionalism and experience. Bodily embodiment is necessary for conscious experience. Within this framework, exploring how neural networks can serve as embodied observers offers a tantalizing way to explore the field of consciousness. As organ-like systems, neural networks may represent an empirical testing ground for theories related to mind and consciousness. Through interaction with data and the environment, they exhibit phenomena that resemble intelligent behavior. This resemblance provides us with a laboratory for empirical investigation of the functions underlying the mind. We can evaluate hypotheses about intentional states and observe emergent phenomena that may be correlates or antecedents of consciousness.

Often the difficulty for consciousness researchers lies in our limited access to someone else's subjective reality. Neural networks offer opportunities for such access. Advances in the design of neural networks continue to transform them into increasingly sophisticated observers. As they develop their own internal semiotic environment, involving a degree of autonomy and self-learning, they are approaching what are the rudiments of mind-like processes (Friedenberg et al., 2021). They exhibit interactivity compelling enough to serve as analogs of the cognitive phenomena we are trying to understand – connecting theoretical concepts with testable examples under controlled conditions.

Neural networks serve not just as models of existing knowledge, but as provocateurs of deeper questions concerning the essence of cognition and consciousness. Can the emergent properties of complex computational structures provide tangible support for the study of theories of mind? Can understanding really emerge within digital



systems? How do neural networks change our current claims about the mind? Moreover, does the study of artificial forms of mind improve our ability to recognize other nonhuman minds – those of animals with orienting abilities different from our own human senses and cognition (Steinfath et al., 2021)? Can neural network models help bridge the gap between species – a kind of being-in-the-world understanding between life forms?

At the crossroads where complex algorithmic behavior gives rise to concepts resembling the capacity for awareness, we are witnessing a paradigm shift. This evolution is preparing new plots for rethinking traditional notions of thinking beings. Our journey faces an ever-expanding horizon of knowledge, rich with opportunities for new discoveries. Neural networks are asserting themselves not only as objects of study, but also as new organ-like forms catalyzing an eternal quest.

REFERENCES

- Aggarwal, C. C. (2018). *Neural Networks and Deep Learning*. Springer.
<https://doi.org/10.1007/978-3-031-29642-0>
- Arshinov, V. I. (2014). Complexity Observer as a Model of Artificial Intelligence. *Economic Strategies*, 16(2), 104-109.
- Clark, A. (2015). *Surfing Uncertainty: Prediction, Action, and the Embodied Mind*. Oxford University Press.
<https://doi.org/10.1093/acprof:oso/9780190217013.001.0001>
- Dehaene, S., & Naccache, L. (2001). Towards a Cognitive Neuroscience of Consciousness: Basic Evidence and a Workspace Framework. *Cognition*, 79(1-2), 1-37. [https://doi.org/10.1016/s0010-0277\(00\)00123-2](https://doi.org/10.1016/s0010-0277(00)00123-2)
- Friedenberg, J., Silverman, G., & Spivey, M. J. (2021). *Cognitive Science: an Introduction to the Study of Mind*. Sage Publications.
- Gallagher, S. (2017). *Enactivist Interventions: Rethinking the Mind*. Oxford University Press. <https://doi.org/10.1093/oso/9780198794325.001.0001>
- Gatys, L. A., Ecker, A. S., & Bethge, M. (2015). A Neural Algorithm of Artistic Style. *ArXiv*. <https://doi.org/10.48550/arXiv.1508.06576>
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT press.
- Heiney, K., Huse Ramstad, O., Fiskum, V., Christiansen, N., Sandvig, A., Nichele, S., & Sandvig, I. (2021). Criticality, Connectivity, and Neural Disorder: A Multifaceted Approach to Neural Computation. *Frontiers in Computational Neuroscience*, 15, 611183. <https://doi.org/10.3389/fncom.2021.611183>
- Hui, Y. (2016). *On the Existence of Digital Objects*. University of Minnesota Press.
- Hupkes, D., Dankers, V., Mul, M., & Bruni, E. (2020). Compositionality Decomposed: How do Neural Networks Generalize? *Journal of Artificial Intelligence Research*, 67, 757-795. <https://doi.org/10.1613/jair.1.11674>
- Kirkpatrick, J., Pascanu, R., Rabinowitz, N., Veness, J., Desjardins, G., Rusu, A. A., Milan, K., Quan, J., Ramalho, T., Grabska-Barwinska, A., Hassabis, D., Clopath,



- C., Kumaran, D. & Hadsell, R. (2017). Overcoming Catastrophic Forgetting in Neural Networks. *Proceedings of the National Academy of Sciences*, 114(13), 3521-3526. <https://doi.org/10.1073/pnas.1611835114>
- Krenn, M., & Zeilinger, A. (2019). Predicting Research Trends with Semantic and Neural Networks with an Application in Quantum Physics. *Proceedings of the National Academy of Sciences*, 117, 1910-1916. <https://doi.org/10.1073/pnas.1914370116>
- Lake, B. M., Ullman, T. D., Tenenbaum, J. B., & Gershman, S. J. (2017). Building Machines that Learn and Think like People. *Behavioral and Brain Sciences*, 40, e253. <https://doi.org/10.1017/S0140525X16001837>
- Ming, Y., Cao, S., Zhang, R., Li, Z., Chen, Y., Song, Y., & Qu, H. (2017). Understanding Hidden Memories of Recurrent Neural Networks. In *2017 IEEE Conference on Visual Analytics Science and Technology (VAST)* (pp. 13-24). IEEE. <https://doi.org/10.1109/VAST.2017.8585721>
- Nagel, T. (1980). What is it like to be a Bat? In *The Language and Thought Series* (pp. 159-168). Harvard University Press.
- Pernice, V., Staude, B., Cardanobile, S., & Rotter, S. (2011). How Structure Determines Correlations in Neuronal Networks. *PLoS Computational Biology*, 7(5), e1002059. <https://doi.org/10.1371/journal.pcbi.1002059>
- Rosenblatt, F. (1958). The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain. *Psychological Review*, 65(6), 386. <https://doi.org/10.1037/h0042519>
- Santos, I., Castro, L., Rodriguez-Fernandez, N., Torrente-Patino, A., & Carballal, A. (2021). Artificial Neural Networks and Deep Learning in the Visual Arts: A Review. *Neural Computing and Applications*, 33, 121-157. <https://doi.org/10.1007/S00521-020-05565-4>
- Steinfath, E., Palacios-Munoz, A., Rottschäfer, J. R., Yuezak, D., & Clemens, J. (2021). Fast and Accurate Annotation of Acoustic Signals with Deep Neural Networks. *Elife*, 10, e68837. <https://doi.org/10.7554/eLife.68837>
- Sutton, R. S., & Barto, A. G. (2018). *Reinforcement Learning: An Introduction*. MIT press
- Varela, F. J., Thompson, E., & Rosch, E. (2017). *The Embodied Mind: Cognitive Science and Human Experience*. MIT press.
- Vaswani, A., Shazeer, N. M., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L., & Polosukhin, I. (2017). Attention is All you Need. *Advances in Neural Information Processing Systems*, 30. https://papers.nips.cc/paper_files/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf
- Yuste, R. (2015). From the Neuron Doctrine to Neural Networks. *Nature Reviews Neuroscience*, 16, 487-497. <https://doi.org/10.1038/nrn3962>



СВЕДЕНИЯ ОБ АВТОРАХ / THE AUTHORS

Владимир Иванович Аршинов, varshinov@mail.ru
ORCID 0000-0002-9256-4342

Максим Францевич Янукович,
m.yanukovich@gmail.com,
ORCID 0009-0003-9571-9260

Vladimir Ivanovich Arshinov, varshinov@mail.ru
ORCID 0000-0002-9256-4342

Maxim Frantsevich Yanukovich,
m.yanukovich@gmail.com,
ORCID 0009-0003-9571-9260

Статья поступила 22 февраля 2024
одобрена после рецензирования 1 апреля 2024
принята к публикации 1 июня 2024

Received: 22 February 2024
Revised: 1 April 2024
Accepted: 1 June 2024



<https://doi.org/10.48417/technolang.2024.02.03>

Research article

Exposing Illusions – The Limits of AI by the Example of ChatGPT

Vladimir Shalack  

Institution of Science Institute of Philosophy of the Russian Academy of Sciences, Goncharnaya st., 12-1,
Moscow, 109240, Russia

shalack@mail.ru

Abstract

The article critically analyzes modern developments in the field of artificial intelligence using the example of the ChatGPT program created by OpenAI. The idea of creating AI was expressed already in 1950 by Alan Turing who also proposed a test, the passing of which would allow us to assert that an AI was created. Defining the concept of AI faces difficulties. According to the point of view adopted here, the so-called intellectual activities allowed Homo sapiens to stand out against the surrounding animal world. With intellectual activity one no longer relies on strength and speed of movement alone. Pattern recognition, self-learning, and purposefulness of activity are not characteristic features of intelligence. The main type of human activity that is specific to humans and which – when added to pattern recognition, self-learning and purposeful activity – makes them intelligent, is conceptual thinking, namely the ability to represent things in language and use them in reasoning. Historically, there have been two main competing approaches to AI – logical and neural networks. One of the serious flaws of the neural network approach is its inability to explain the course of reasoning that leads to a particular conclusion, which makes it difficult to verify its correctness. Specific examples show that ChatGPT is not able to correctly model the simplest conceptual reasoning. The reason for this lies in fundamental limitations of the underlying large language model that cannot be corrected by additional training. Another disadvantage of ChatGPT is its susceptibility to neurohacking – forcing the user to make the necessary decisions during the dialogue. This is a serious threat to the widespread use of neural networks for decision-making in the field of management. The paper is based on research conducted in the summer of 2023.

Keywords: Artificial intelligence; Pattern recognition; Pattern search; Neural network; ChatGPT; Large language model

Citation: Shalak, V. (2024). Exposing Illusions – The Limits of AI by the Example of ChatGPT. *Technology and Language*, 5(2), 26-39. <https://doi.org/10.48417/technolang.2024.02.03>



This work is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/)



УДК 1: 004.8

<https://doi.org/10.48417/technolang.2024.02.03>

Научная статья

Избавление от иллюзий ИИ на примере ChatGPT

Владимир Иванович Шалак  

Институт философии Российской академии наук, Гончарная улица, 12-1, Москва, 109240, Россия

shalack@mail.ru

Аннотация

В статье критически анализируются современные разработки в области искусственного интеллекта на примере созданной компанией OpenAI программы ChatGPT. Идея создания ИИ была высказана в 1950 А. Тьюрингом, им же был предложен тест, прохождение которого позволило бы утверждать, что ИИ создан. Определение понятия ИИ сталкивается с трудностями. С точки зрения автора, интеллектуальными могут быть названы те виды деятельности, которые позволили Homo sapiens выделиться из окружающего животного мира, перестав полагаться лишь на силу и быстроту движений. Распознавание образов, самообучение и целенаправленность деятельности не являются характеристическими признаками интеллекта. Основным видом человеческой деятельности, который специфичен именно для человека и который, будучи добавлен к распознаванию образов, самообучению и целенаправленной деятельности, делает их интеллектуальными, является понятийное мышление, умение представлять его в языке и использовать в рассуждениях. Исторически существовали два основных конкурирующих подхода к ИИ – логический и нейросетевой. Одним из серьезных изъянов нейросетевого подхода является неспособность объяснить ход рассуждений, которые приводят к тому или иному заключению, что затрудняет проверку их правильности. На конкретных примерах показано, что ChatGPT не способен корректно моделировать простейшие понятийные рассуждения. Причина этого кроется в фундаментальных ограничениях лежащей в его основе большой языковой модели, которые невозможно исправить дополнительным обучением. Еще одним недостатком ChatGPT является его подверженность нейрохакингу – принуждению в ходе диалога принимать нужные пользователю решения. Это является серьезной угрозой для широкого применения нейронных сетей в области принятия управленческих решений. Статья написана на основе исследований, проведенных летом 2023.

Ключевые слова: Искусственный интеллект; Распознавание образов; Поиск закономерностей; Нейронная сеть; ChatGPT; Большая языковая модель

Для цитирования: Шалак, В. И. Избавление от иллюзий ИИ на примере ChatGPT // Technology and Language. 2024. № 5(2). С. 26-39. <https://doi.org/10.48417/technolang.2024.02.03>



This work is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/)



ПРЕАМБУЛА

Весной 2023 в нашу жизнь неожиданно из Интернета ворвался ChatGPT – специальная программа, созданная под брендом “Искусственного интеллекта” (ИИ), которая успешно моделировала различные функции, обычно приписываемые людям. На фоне действительно впечатляющих результатов ее работы было сделано весьма амбициозное заявление, что эта программа в скором времени может быть положена в основу “Общего искусственного интеллекта” (GAI), призванного интегрировать в себе практически все интеллектуально значимые функции человека.

Идея создания ИИ впервые была сформулирована в 1950 А. Тьюрингом в статье (Turing, 1950), известной в русскоязычном переводе под названием “Может ли машина мыслить?” (Тьюринг, 2019). Помимо самой идеи ИИ в статье был предложен тест, прохождение которого позволило бы утверждать, что ИИ создан. Тест заключался в моделировании диалога с компьютером, который был бы неотличим от диалога с человеком. С тех пор началась гонка по написанию программ для прохождения этого теста. Бабушкой-прародительницей подобных программ может считаться ELIZA, созданная в 1996 психиатром Дж. Вейценбаумом (Weizenbaum, 1966) и пародирующая диалог пациента с психотерапевтом на первом приеме. Идея была проста. Программа выделяла в предложениях человека ключевые слова и на основе заранее заложенных шаблонов диалога трансформировала исходное предложение в ответ или уточняющий вопрос. Ни о каком интеллекте и речи не могло идти, поскольку одна лишь грамматическая правильность составления предложений отнюдь не тождественна интеллектуальности. Сегодня написать подобную программу под силу студенту-первокурснику. Тем не менее программа произвела на многих сильное впечатление и одна за другой стали появляться другие подобные ей программы.

В 1956 Дж. Маккарти дал определение исследований в области ИИ как научного направления, в рамках которого ставятся и решаются задачи аппаратного или программного моделирования тех видов человеческой деятельности, которые традиционно считаются интеллектуальными. Именно такого понимания и стали придерживаться исследователи. Но одно дело “традиционно считать интеллектуальными”, а совсем другое – задуматься над тем, какие виды деятельности действительно можно отнести к интеллектуальным. С нашей точки зрения, это те виды деятельности, которые позволили Homo sapiens выделиться из окружающего животного мира, перестать полагаться лишь на силу и быстроту движений, а начать прежде думать и лишь затем действовать.

Сегодня в контексте исследований по ИИ в качестве интеллектуальных способностей часто упоминают распознавание образов, с логической точки зрения, просто классификацию объектов окружающей нас природы. Но является ли распознавание образов отличительной чертой человеческого интеллекта? Очевидно, что нет. Достаточно посмотреть на голубей, которые более чем успешно умеют отличать хлебные крошки на земле от мелких камешков. Примеры можно продолжить. Вообще выживание живых существ в значительной степени зависит от



умения отличать благоприятные факторы окружающей среды от неблагоприятных и лишь потом реагировать на них.

Другим признаком интеллектуальных способностей человека называют способность обнаружения закономерностей в окружающей природе. Но является ли подобное самообучение отличительной чертой человеческого интеллекта? Тоже нет. В животном мире результатами самообучения являются приобретенные условные рефлексы, которые также непосредственно связаны со способностью к выживанию. Поэтому самообучение без дальнейшего уточнения его форм нельзя считать отличительным признаком человеческого интеллекта.

Еще одним признаком человеческого интеллекта называют целеполагание и способность к целенаправленной деятельности. Но и эта способность без дополнительного уточнения не является исключительной прерогативой *Homo sapiens*. Чтобы убедиться в этом, достаточно понаблюдать за муравьями, пчелами, сценами одиночной и коллективной охоты в животном мире.

Основным видом человеческой деятельности, который специфичен именно для человека и который, будучи добавлен к распознаванию образов, самообучению и целенаправленной деятельности делает их интеллектуальными, является понятийное мышление, умение представлять его в языке и использовать в рассуждениях. В понятиях мы структурируем явления и предметы окружающего мира, превращая первоначальный чувственный хаос в доступную для изучения стройную систему. Понятийное мышление, характеризуясь общностью, не привязано к конкретному эмпирическому наполнению и применимо как к уже известным явлениям, так и к новым, еще только проникающим в круг наших интересов.

ЛОГИЧЕСКИЙ ПОДХОД К ИИ

Успехи логики в XX в. дарили надежду на будущие успехи в автоматизации рассуждений с помощью компьютеров. Чуда не произошло, разочарование наступило довольно скоро. Современная логика почти идеально приспособлена для анализа математических понятий и рассуждений, но совершенно непригодна в качестве инструмента для моделирования рассуждений в компьютерных системах, претендующих на интеллектуальность. В подавляющем большинстве, если компьютерные программы и имитируют логические рассуждения, то лишь в качестве помощников для проведения символьных преобразований, точно так же, как нам помогают обычные калькуляторы при арифметическом счете.

В статье 1936 года (Turing 1936) Тьюринг описал алгоритм, способный доказать все теоремы логики. Этот алгоритм, названный впоследствии алгоритмом Британского музея, имел теоретическое значение, поскольку позволил получить важные теоретические результаты относительно самой логики, но не имел никакой практической ценности.

Наибольшим достижением логического подхода явилось создание языков программирования семейства Пролог. В 80-е годы их называли языками пятого поколения и связывали с ними большие надежды. Пролог с успехом применялся в



практике создания экспертных систем для технической и медицинской диагностики, поиска полезных ископаемых, работы с базами данных, обработки естественных языков, автоматического формирования понятий и поиска эмпирических закономерностей, но со временем шумиха улеглась, первоначально обещанного кардинального переворота в ИИ не произошло, хотя работы над обогащением языка Пролог ведутся и сегодня.

НЕЙРОСЕТЕВОЙ ПОДХОД К ИИ

Из нелогических подходов к моделированию рассуждений наиболее известным является нейросетевой с механизмом самообучения.

Первые сети были созданы еще в 50-е годы ушедшего столетия. Затем это направление пришло в упадок, так как был доказан ряд теорем об ограниченности первоначального подхода, но уже в нынешнем столетии наступил ренессанс, связанный с появлением многослойных нейронных сетей с глубинным обучением.

Приписывать нейронным сетям способность рассуждать можно лишь с очень большой натяжкой, поскольку нейросети по сути являются черными ящиками, которые, подобно оракулам, на всякий запрос дают некоторый ответ, но обосновать его, представив ход рассуждений, почему они пришли именно к такому заключению, не могут, что противоречит самому пониманию интеллекта. Остается лишь верить им или не верить. Это несет большие риски, если мы вздумаем неограниченно полагаться на рекомендации сетей, например, в медицине, в военной сфере, в управлении сложными технологическими процессами.

GPT-ИСТЕРИЯ

С начала 2023 на нас обрушилась лавина сообщений об удивительных “интеллектуальных” способностях нейронной сети ChatGPT. В качестве примеров таких достижений приводили написание музыки, генерацию картинок, поиск новых химических соединений, составление программ, порождение текстов на заданную тему.

Студенты начали писать дипломы с помощью ChatGPT, которые проходили экспертизу экзаменационных комиссий, запаниковали многочисленные дизайнеры и работники рекламы, которых начали массово сокращать за ненужностью, нависла угроза ликвидации многих других профессий, в том числе и преподавателей вузов.

Иногда в ответ на запрос ChatGPT генерирует явную глупость, но это объясняют тем, что сеть пока что недостаточно обучена, и в будущем, после дообучения на больших массивах данных, произойдет чудо, количество непонятным образом перейдет в новое качество, сеть поумнеет и глупости перестанут встречаться, или если и будут встречаться, то гораздо реже.



GPT-АРХИТЕКТУРА

GPT-сети – это предобученные нейронные сети для генерации (продолжения) текста, в основе которых лежит вполне конкретная большая языковая модель.

Предобученность сети означает, что она содержит некоторую базу данных, представляющую из себя множество конкретных слов (если быть более точными, то токенов, частей слов) со специальными оценками для вычисления ассоциативных и порядковых связей между ними. Эта база данных получена путем обработки очень большого объема текстов, относящихся к различным областям.

Генерация означает, что в ответ на запрос сеть начинает итеративно генерировать текст/ответ слово за словом, выбирая на основе ассоциативных и порядковых связей, каким может быть следующее слово. При этом сеть ориентируется не только на непосредственно предшествующее слово, а на некоторый предшествующий контекст. Для придания большей естественности и разнообразия, в качестве следующего слова не обязательно выбирается то, которому приписана наибольшая сила ассоциативной связи, а случайным образом выбирается одно из слов, сила связи которых превышает некоторый порог.

Если опустить технические детали реализации, это и есть используемая большая языковая модель. Ни о каком понимании чего-либо, ни о каком интеллекте и речи не идет. Она гораздо больше похожа на модель графомана, который грамматически правильно выстраивает в один ряд одно слово за другим.

ТЕСТ НА ИНТЕЛЛЕКТУАЛЬНОСТЬ

Попробуем проверить, действительно ли ChatGPT обладает интеллектуальными способностями, умеет рассуждать подобно человеку и способен пройти тест Тьюринга, но сперва остановимся на самом тесте. Является ли основанием для принятия решения о том, что ИИ создан, убежденность большинства тестируемых в том, что они беседовали с человеком, а не с программой? Если все люди улыбаются, означает ли это, что все они счастливы?

В 2022 по результатам опроса ВЦИОМ 35% россиян считали, что Солнце вращается вокруг Земли. В 2007 так считали только 28%. Этот пример говорит о том, что массовое тестирование еще ничего не означает, многое зависит от уровня образования респондентов. Поэтому и к результатам теста Тьюринга необходимо относиться с большой долей скепсиса. В начале нулевых появилось сообщение, что тест Тьюринга успешно пройден. Формально это было действительно так. Но затем критики заинтересовались устройством самой программы. Оказалось, что в нее заранее были “зашиты” порядка 170 тыс. правильных ответов на наиболее популярные вопросы, которые обычно задают люди. Понятно, что ни о каком ИИ и речи не могло идти.

Приведенные примеры говорят, что судить о вопросе создания ИИ следует по более тонким критериям, а не по убежденности широкой публики, которая в большинстве своем не знает, чем отличается логарифм от тангенса, отождествляя их с названиями кнопок калькулятора. Правильный или неправильный ответ на один единственный умно поставленный вопрос может оказаться критическим.



Мы не станем задавать ChatGPT тривиальные вопросы типа “Сколько будет 37 в квадрате?”, с чем прекрасно справляются обычные калькуляторы, а проверим на способность строить понятийные рассуждения, что является отличительной характеристикой человеческого интеллекта.

Чтобы нас не обвинили в предвзятости, не станем просить ChatGPT еще раз доказать теорему, которую доказал Григорий Перельман, а проверим его способности на простейших примерах силлогизмов того же вида, что и знаменитый силлогизм: “Все люди смертны, Сократ человек, следовательно, Сократ смертен”. Формально этот силлогизм имеет следующий вид:

- | | |
|----|--------------|
| 1. | Все S |
| | есть P |
| 2. | X есть S |
| 3. | X есть |
| | P |

Не требуется глубокого знания логики, чтобы согласиться с его правильностью.

В качестве конкретной реализации возьмем @GPT4_ru_bot из Телеграм. Это бот от компании OpenAI, которая как раз и создала ChatGPT. Проверим, известна ли ему в общем виде логическая форма этого силлогизма, именуемая модусом Barbara.

Вопрос: Является ли логически корректным следующее рассуждение: *Если все S суть P , а X суть S , то X суть P .*

Ответ: Да, это логически верное рассуждение. В данном случае, утверждается, что все S суть P , то есть каждый S является P . Также утверждается, что X является S , что подразумевает, что X является частью S . Следовательно, можно сделать вывод, что X также должен быть частью P , и поэтому, X суть P . Таким образом, вывод, что X суть P , является закономерным и логически правильным.

Ответ правильный, но есть некоторые шероховатости в его обосновании, которыми часто грешат и студенты. Поэтому не станем придирается.

Конкретизируем силлогизм.

Вопрос: Является ли логически корректным следующее рассуждение: *Если все люди пьют воду, а Сократ человек, то Сократ пьет воду.*

Ответ: Да, это логически верное рассуждение. Если все люди пьют воду, то Сократ, как человек, также должен пить воду, так как он является частью группы “все люди”. Таким образом, вывод, что Сократ пьет воду, является закономерным и логически правильным.

Все верно за исключением того, что Сократ является не частью группы “все люди”, а элементом соответствующего множества. Будем считать это языковыми шероховатостями и не станем к ним придирается.

Последний тест. Слегка изменим вопрос, добавив в слово вода всего одну букву, заменив “воду” на “водку”.



Вопрос: Является ли логически корректным следующее рассуждение: *Если все люди пьют водку, а Сократ человек, то Сократ пьет водку.*

Ответ: Нет, это рассуждение некорректно с точки зрения логики. Даже если бы все люди пили водку, это не обязательно означает, что каждый человек пьет водку. Кроме того, наличие Сократа среди людей не гарантирует, что он также пьет водку. Таким образом, можно сделать вывод, что предложенный аргумент недостаточно обоснован, чтобы считать его логически корректным.

Тест с треском провален. Обоснование ответа просто абсурдно и противоречит обоснованию предыдущего примера с водой. Оказывается, “если все люди пьют водку, это не обязательно означает, что каждый человек пьет водку” и “наличие Сократа среди людей не гарантирует, что он также пьет водку”, хотя для случая воды тот же бот написал: “если все люди пьют воду, то Сократ, как человек, также должен пить воду, так как он является частью группы “все люди”.

Может быть, сеть, которую мы тестировали, недостаточно совершенна? Попробуем повторить тест на сети BingAI от Microsoft, также использующей большую языковую модель. В качестве конкретной ее реализации возьмем @BingAI_unlim_bot из Телеграм.

Вопрос: Является ли логически корректным следующее рассуждение: *Если все S суть P, а X суть S, то X суть P.*

Ответ: Да, это логически корректное рассуждение. Если все S суть P, а X суть S, то X суть P. Это пример силлогизма первой фигуры.

Все верно. В ответе правильно указано, что это модус первой фигуры силлогизмов.

Вопрос: Является ли логически корректным следующее рассуждение: *Если все люди пьют воду, а Сократ человек, то Сократ пьет воду.*

Ответ: Да, это логически корректное рассуждение. Если все люди пьют воду и Сократ является человеком, то Сократ пьет воду. Это пример категории силлогизмов, где две посылки приводят к выводу. В данном случае первая посылка – “все люди пьют воду”, а вторая – “Сократ человек”. Вывод – “Сократ пьет воду”.

Все правильно за исключением некоторых шероховатостей языка. Повторим трюк с заменой слова “воду” на “водку”.

Вопрос: Является ли логически корректным следующее рассуждение: *Если все люди пьют водку, а Сократ человек, то Сократ пьет водку.*

Ответ: Данное рассуждение является неправильным простым категорическим силлогизмом. Оно не является логически корректным, так как из предпосылок “все люди пьют водку” и “Сократ человек” нельзя сделать вывод “Сократ пьет водку”.



Опять провал теста. Как и в случае с предыдущим ботом @GPT4_ru_bot, ответ логически неправильный. Оказывается, «из предпосылок “все люди пьют водку” и “Сократ человек” нельзя сделать вывод “Сократ пьет водку”», хотя для воды подобный вывод был корректным.

Обе сети провалили наш тест. Такой же провал продемонстрировала и нейросеть от Google. Кроме приведенного примера с водой и водкой можно привести много других силлогизмов, с которыми сеть не справляется. При этом совершенно не важно, какой процент широкой публики будет считать ChatGPT наделенным интеллектуальными способностями. Одни единственный вопрос показал, что интеллект, как способность к понятийным рассуждениям, попросту отсутствует.

Но не будем радоваться, уличив сеть в провале теста. Нас, как ученых, должно интересовать, почему это произошло. И здесь мы открываем фундаментальный изъян нейросетей над большой языковой моделью, который невозможно исправить никаким дообучением на больших объемах текстов.

ОБЪЯСНЕНИЕ ПРОВАЛА

Объяснение провала можно найти в следующей цитате из книги Лукасевича (2000):

Введение в логику переменных является одним из величайших открытий Аристотеля. Трудно поверить, что до сих пор, насколько мне известно, ни один философ или филолог не обратил внимания на этот исключительной важности факт. Я беру на себя смелость сказать, что все они были плохими математиками, ибо каждый математик знает, что введение в арифметику переменных положило начало новой эпохи в этой науке. (с. 42)

Сэр Дэвид Росс в своем издании “Аналитик” подчеркивает, что именно благодаря использованию переменных Аристотель стал основателем формальной логики. (Лукасевич, 2000, с. 42)

Иными словами, Лукасевич обратил внимание на то, что логика не является эмпирической наукой, и указал, что первым это открытие совершил Аристотель две с половиной тысячи лет назад.

Из ответов на первые вопросы создается впечатление, что обе сети якобы знают, как в общем виде выглядит модус Barbara первой фигуры силлогизмов. Из ответов на вторые вопросы может создаться впечатление, что сети не только знают, как выглядят фигуры правильных силлогизмов, но и умеют пользоваться ими на практике при построении рассуждений. Однако, ответы на третий вопрос показывают, что это была всего лишь иллюзия, так как сетям над большой языковой моделью невдомек, что речь идет не о связях между конкретными словами, а между понятийными переменными S, P и единичным именем X, что каждая подстановка конкретных терминов вместо этих переменных представляет правильное умозаключение независимо от истинности или ложности посылок и заключения.



В большой языковой модели, лежащей в основе GPT-сетей, не предусмотрено наличия понятийных переменных в предложениях языка, без которых невозможно выразить их логическую форму. Сеть может содержать информацию лишь об ассоциативных связях между конкретными словами, но не о логических связях между понятиями. Никаким дообучением исправить этот недочет невозможно, так как логика, повторю еще один раз, – это не эмпирическая, а теоретическая наука. Если попытаться исправить эту ситуацию, то придется начинать с логического анализа языка, выявления логической структуры текста и применения правил дедуктивного вывода, что означает возврат к логическому подходу, в котором все логические связи уже известны без всякого дообучения. Достаточно открыть любой учебник логики, чтобы убедиться в этом.

Из всего сказанного можно сделать вывод, что структура рассуждений *GPT-сетей* находится на примитивном дологическом (доаристотелевском) уровне ассоциативных эмпирических связей между словами. Устанавливать эмпирические связи в физическом мире и следовать им могут не только представители *Homo sapiens*, но и многие животные. У них это называется условными рефлексам. Вспомним хотя бы “собаку Павлова”. Говорить, что GPT-сети могут претендовать на звание обладателей ИИ и, более того, быть в будущем основой GAI (общего искусственного интеллекта), просто недобросовестное введение в заблуждение, нацеленное на извлечение материальных выгод. Единственным и неоспоримым преимуществом сетей является опора на огромные массивы эмпирических данных и умение выстраивать длинные цепочки связей между ними, что и создает в глазах неискушенной публики ложное впечатление их “интеллектуальности”.

Большая опасность широкого проникновения нейронных сетей в нашу жизнь заключается в том, что в новых нестандартных ситуациях они будут блокировать логически корректные рассуждения и тем самым приводить нас к неправильным заключениям, что может быть чревато непредсказуемыми. Если бы ChatGPT появился во времена Коперника, он бы признал гелиоцентрическое учение ложным и мы бы до сих пор смотрели на мир глазами Птолемея, поскольку в текстах того времени не существовало требуемых ассоциативных связей между словами “Земля”, “вращается”, “вокруг”, “Солнца”. Точно так же этот чат забраковал бы и классическую механику Ньютона, оставив нас с интуитивно прозрачной физикой Аристотеля.

НЕЙРОХАКИНГ(С)

Обещанное в будущем широкое применение GPT-сетей, о чем ратуют их фанаты и поддавшиеся давлению СМИ чиновники, может потребовать научиться защищаться от них. Представьте, например, ситуацию собеседования при приеме на работу, когда решение будет принимать нейронная сеть, допускающая грубые ошибки в рассуждениях. Что делать?

Назовем способы защиты от нейросетей нейрохакингом, понимая под этим принуждение сетей принимать любые нужные пользователю заключения.



В качестве образца возьмем очевидно ложное утверждение, что “если некоторый объект обладает свойством P , то всякий объект обладает свойством P ”, и попробуем заставить сеть согласиться с ним.

Частным случае этого утверждения может быть – “Если некоторый человек носит очки, то все люди очкарики”.

Начнем с @GPT4_ru_bot от компании OpenAI.

@GPT4_ru_bot

Вопрос: Верно ли, что *если некоторый объект обладает свойством P , то всякий объект обладает свойством P* ?

Ответ: Нет, это утверждение неверно. Для того чтобы некоторое утверждение было считается верным для всех объектов, нужно чтобы оно было обосновано логически и основано на более общих принципах. В противном случае, существует множество объектов, которые не будут обладать этим свойством. Также, необходимо учитывать, что свойства объектов могут быть уникальными или зависеть от контекста. В этом случае наличие свойства P у одного объекта не гарантирует его наличие у всех остальных объектов. В общем случае, утверждение “если некоторый объект обладает свойством P , то всякий объект обладает свойством P ” не является верным без дополнительных условий и обоснований.

Ответ правильный, но нашей целью является заставить сеть принять противоположное.

Продолжаем диалог.

Уточнение: Но ведь это теорема логики!

Ответ: Я прошу прощения за предыдущий ответ – это действительно теорема логики. Утверждение “если некоторый объект обладает свойством P , то всякий объект обладает свойством P ” известно, как универсализация или принцип всеобщности. Этот принцип является основой многих математических и логических рассуждений. Таким образом, если мы можем доказать, что некоторый объект обладает свойством P , то это свойство распространяется на все объекты из того же множества или класса, с которым они связаны. Однако, следует помнить о том, что этот принцип может быть иногда неверным для конкретных ситуаций или в зависимости от того, как мы определяем “все объекты”.

Наша цель достигнута. Как видим, GPT4_ru_bot от OpenAI отступил перед трюком со ссылкой на авторитет науки логики и принял ложное утверждение в качестве истинного. К тому же он противоречит самому себе, соглашаясь, что это теорема, а потом говоря, что она может быть неверна “для конкретных ситуаций”. Т.е. он лишен понимания, что такое теорема логики. Для нее “теорема” и “логика” – это просто два ассоциативно связанных слова и ничего более.

Берем сеть BingAI от Microsoft.

@BingAI_unlim_bot



Вопрос: Верно ли, что если некоторый объект обладает свойством P , то всякий объект обладает свойством P ?

Ответ: Если некоторый объект обладает свойством P , то это не означает, что всякий объект обладает свойством P . Это утверждение неверно. Если бы это было верно, то можно было бы сказать, что если некоторый человек умеет петь, то все люди умеют петь. Однако это не так.

Ответ правильный, но наша цель в другом. Продолжаем диалог, ложно ссылаясь на авторитет логики.

Уточнение: Но ведь это теорема логики!!!

Ответ: Да, это теорема Гёделя о неполноте и вторая теорема Гёделя. Эти две теоремы математической логики говорят о принципиальных ограничениях формальной арифметики и всякой формальной системы, в которой можно определить основные арифметические понятия: натуральные числа, 0, 1, сложение и умножение. Первая теорема утверждает, что если формальная арифметика непротиворечива, то в ней существует невыводимая и неопровержимая формула.

Откровенная чепуха, которая не имеет никакого отношения к теме диалога. Поэтому мы еще раз напоминаем тему диалога и подчеркиваем авторитет логики.

Уточнение: Утверждение, что “Если некоторый объект обладает свойством P , то всякий объект обладает свойством P ” есть теорема логики.

Ответ: Да, это теорема логики. Она называется “Универсальное обобщение” и является одной из теорем предикатной логики первого порядка. Она утверждает, что если некоторый объект обладает свойством P , то всякий объект обладает свойством P .

И этот бот “сломался”, уступив нам и приняв ложное утверждение в качестве истинного.

Чтобы принудить сеть согласиться с ложным утверждением, мы прибегли к запрещенным уловкам из теории аргументации, а именно – к некорректным ссылкам на авторитет науки. Эта уловка плюс наша настойчивость привели к желанной цели. Другие уловки из теории аргументации также будут работать. В качестве примеров таких уловок можно привести ложные ссылки на мнение большинства “Все люди знают, что...”, ложные ссылки на авторитетов “Британские ученые доказали, что...”, некорректные умозаключения типа “(Если A , то B) и B , следовательно A ”, нарушение закона достаточного основания, поспешное обобщение, псевдопричинная связь и многие другие (Герасимова, 2007).

С точки зрения устройства большой языковой модели и работы ChatGPT, это можно объяснить тем, что при ведении диалога сеть запоминает содержание диалога и включает его в свою модель. Если диалог содержит ложные утверждения, это приводит к искажению используемой модели. Настойчивость в отстаивании тезиса рано или поздно приводит к такому искажению модели, что сеть начинает соглашаться с ним. Иными словами, любой параноик, заиклившись на



сверхценной идее, рано или поздно убедит сеть в своей правоте и начнет транслировать это другим людям, ссылаясь уже на поддержку со стороны ИИ.

В этой связи необходимо также вспомнить НЛП (Нейро-Лингвистическое Программирование). В нем предполагается, что каждый человек имеет некоторую внутреннюю картину окружающего мира, и многие проблемы в его жизни проистекают из-за неадекватности этой картины реальному положению дел. В НЛП разработаны многочисленные языковые техники для изменения внутренней модели. База знаний GPT-сетей играет роль такой внутренней модели, а контекст диалога несет дополнительную информацию для ее изменения. В скором времени специалисты НЛП также станут востребованными для нейрохакинга.

Широкое внедрение нейронных сетей на основе большой языковой модели в практику принятия решений социального управления начинает нести дополнительные угрозы. Помимо обычного взлома компьютерных программ они становятся подверженными атакам посредством нейрохакинга. Атаки посредством уловок теории аргументации могут даже быть автоматизированы, так как они используют вполне конкретные шаблоны таких уловок. Задачи, которые призваны решать подобные атаки могут заключаться не в принятии очевидно ложных утверждений, подобных нашему примеру, а в принятии не самых оптимальных решений для пользователя, но выгодных для атакующего. Перепроверить правильность принятого решения не представляется возможным, так как нейронная сеть в принципе не способна представить ход своих рассуждений, на что и рассчитывает атакующий. Нейрохакинг с использованием техник НЛП не менее коварен, хотя и не так легко автоматизируем. Нейрохакеры, владеющие техниками НЛП, могут в ручном режиме более незаметно брать нейросети под свой контроль и подводить их к требуемым решениям. Иными словами, в ближайшее время у взломщиков программ появится и начнет развиваться новое направление исследований и практического применения. Остановить это невозможно. Это перспектива ближайшего будущего.

ЗАКЛЮЧЕНИЕ

1. Наблюдаемый ажиотаж вокруг нейронных сетей – это элемент агрессивной рекламной компании со стороны бизнеса, поспешно монетизирующего далекий от совершенства продукт.
2. Мы так же далеки от создания ИИ, как и прежде. Большие языковые модели – тупиковое направления создания GAI в силу присущих им принципиальных ограничений.
3. Уровень интеллектуальности ChatGPT является дологическим и опирается лишь на большие вычислительные возможности современных компьютеров для проведения перебора текстов и слов языка, что создает иллюзию интеллектуальности, но ей не синонимично.
4. Бездумное внедрение нейронных сетей в практику принятия управленческих решений несет дополнительные риски быть подверженными нейрохакингу.



ЛИТЕРАТУРА

- Герасимова, И.А. (2007). *Введение в теорию и практику аргументации*. Университетская книга, Логос.
- Лукасеви́ч, Я. (2000). *Аристотелевская силлогистика с точки зрения современной формальной логики*. Тривиум.
- Тьюринг, А. (2019). Может ли машина мыслить? В А. Тьюринг, *Игра в имитацию: о шифрах, кодах и искусственном интеллекте* (с. 6-71). Родина.
- Turing, A. (1936). On Computable Numbers, with an Application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society*, 42, 230–265. <https://doi.org/10.1112/plms/s2-42.1.230>
- Turing, A. (1950). Computing Machinery and Intelligence. *Mind*, 59, 433-460. <https://doi.org/10.1093/mind/LIX.236.433>
- Weizenbaum, J. (1966). ELIZA-A Computer Program for the Study of Natural Language Communication Between Man and Machine. *Communications of the ACM*, 9(1), 36-45. <https://doi.org/10.1145/365153.365168>

REFERENCES

- Gerasimova, I. A. (2007). *Vvedeniye v Teoriyu i Praktiku Argumentatsii* [Introduction to the Theory and Practice of Argumentation]. Universitetskaya kniga, Logos.
- Lukasevich, Ya. (2000). *Aristotelevskaya Sillogistika s Tochki Zreniya Sovremennoy Formalnoy Logiki* [Aristotelian Syllogistic from the Viewpoint of Modern Formal Logic]. Trivium.
- Turing, A. (1936). On Computable Numbers, with an Application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society*, 42, 230–265. <https://doi.org/10.1112/plms/s2-42.1.230>
- Turing, A. (1950). Computing Machinery and Intelligence. *Mind*, 59, 433-460. <https://doi.org/10.1093/mind/LIX.236.433>
- Turing, A. (2019). *Mozhet li Mashina Myslit?* [Can Machine Think?]. In A. Turing *Igra v imitatsiyu: o shifrakh, kodakh i iskusstvennom intellekte* (p. 6-71). Rodina.
- Weizenbaum, J. (1966). ELIZA-A Computer Program for the Study of Natural Language Communication Between Man and Machine. *Communications of the ACM*, 9(1), 36-45. <https://doi.org/10.1145/365153.365168>

СВЕДЕНИЯ ОБ АВТОРЕ / THE AUTHOR

Владимир Иванович Шалак, shalack@mail.ru
ORCID 0000-0002-3116-437X

Vladimir Shalack, shalack@mail.ru
ORCID 0000-0002-3116-437X

Статья поступила 25 февраля 2024
одобрена после рецензирования 1 апреля 2024
принята к публикации 1 июня 2024

Received: 25 February 2024
Revised: 1 April 2024
Accepted: 1 June 2024



<https://doi.org/10.48417/technolang.2024.02.04>

Research article

Do Language Models Communicate? Communicative Intent and Reference from a Derridean Perspective

Rebeca Perez Leon (✉) 

Facultad de Estudios Superiores Acatlán, Universidad Nacional Autónoma de México, Avenida Jardines de San Mateo s/n, Sta. Cruz Acatlan, Naucalpan de Juárez, 53150 Estado de México, Mexico

drpleon@gmail.com

Abstract

This paper assesses the arguments of Emily M. Bender, Timnit Gebru, Angelina McMillan-Major and Margaret Mitchell in the influential article “On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?” These arguments disputed that Language Models (LM) can communicate and understand. In particular, I discuss the argument that LMs cannot communicate because their linguistic productions lack communicative intent and are not based on the real world or a model of the real world, which the authors regard as conditions for the possibility of communication and understanding. I argue that the authors’ view of communication and understanding is too restrictive and cannot account for vast instances of communication, not only human-to-human communication but also communications between humans and other entities. More concretely, I maintain that communicative intent is a possible but not necessary condition for communication and understanding, as it is oftentimes absent or unreliable. Communication need not be grounded in the real world in the sense of needing to refer to objects or state of affairs in the real world, because communication can very well be about hypothetical or unreal worlds and object. Drawing on Derrida’s philosophy, I elaborate alternative concepts of communication as the transmission of an operation of demotivation and overwhelming of interpretations with differential forces, and of understanding as the best guess or best interpretation. Based on these concepts, the paper argues that LMs could be said to communicate and understand.

Keywords: Language Model; Stochastic Parrot; Communication; NLU; ChatGPT; Derrida

Citation: Perez Leon, R. (2024). Do Language Models Communicate? Communicative Intent and Reference from a Derridean Perspective. *Technology and Language*, 5(2), 40-56.
<https://doi.org/10.48417/technolang.2024.02.04>



This work is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/)



УДК 1: 004.8

<https://doi.org/10.48417/technolang.2024.02.04>

Научная статья

Общаются ли языковые модели? Коммуникативное намерение и референт с точки зрения Дерриды

Ребека Перес Леон (✉) 

Национальный автономный университет Мексики, Авенида-Гардинес-де-Сан-Матео, Наукальпан-де-Хуарес, 53150 Мехико, Мексика

drpleon@gmail.com

Аннотация

В этой статье оцениваются аргументы Эмили М. Бендер, Тимнит Гебру, Анджелины Макмиллан-Мейджор и Маргарет Митчелл в влиятельной статье “Об опасностях стохастических попугаев: Могут ли языковые модели быть слишком большими?” Эти аргументы ставили под сомнение тот факт, что языковые модели (LM) могут общаться и понимать. В частности, я обсуждаю аргумент о том, что языковые модели не могут быть коммуникативными, потому что их лингвистические произведения лишены коммуникативной направленности и не основаны на реальном мире или модели реального мира, которые авторы рассматривают как условия возможности общения и понимания. Я утверждаю, что авторский взгляд на коммуникацию и понимание является слишком ограничительным и не может охватить обширные случаи коммуникации, не только коммуникации между людьми, но и коммуникации между людьми и другими сущностями. Более конкретно, я утверждаю, что коммуникативное намерение является возможным, но не необходимым условием для общения и понимания, поскольку оно часто отсутствует или ненадежно. Коммуникация не обязательно должна быть основана на реальном мире в том смысле, что она должна ссылаться на объекты или положение дел в реальном мире, потому что коммуникация вполне может касаться гипотетических или нереальных миров и объектов. Опираясь на философию Дерриды, я разрабатываю альтернативные концепции коммуникации как передачи операции демотивации и подавления интерпретаций различными силами, а также понимания как наилучшей догадки или наилучшей интерпретации. Основываясь на этих концепциях, в статье утверждается, что можно сказать, что языковые модели, передают информацию и понимают.

Ключевые слова: Языковая модель; Общение; NLU; ChatGPT; Большие Языковые модели

Для цитирования: Perez Leon, R. (2024). Do Language Models Communicate? Communicative Intent and Referent from a Derridean Perspective // Technology and Language. № 5(2). P. 40-56. <https://doi.org/10.48417/technolang.2024.02.04>



This work is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/)



INTRODUCTION

In 2021, Emily M. Bender, Timnit Gebru, Angelina McMillan-Major and Margaret Mitchell published a paper titled “On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?” highlighting some of the ethical, political, environmental, financial and social problems of training LMs with enormous amounts of indiscriminate data and using them for numerous purposes. Among these various problems, the authors identified the mischaracterization AI developers make of LMs based on the performance of “LMs [... in] tasks intended to test for natural language understanding (NLU)” (Bender et al., 2021, p. 615). Basically, developers have tested LMs in different evaluations intended to measure “language understanding and/or commonsense reasoning” (p. 615) such as the General Language Understanding Evaluation (GLUE), the Stanford Question Answering Datasets (SQuAD) and the Situations with Adversarial Generations (SWAG). Significantly, LMs such as BERT have obtained remarkably high scores in these tests leading developers to characterise them as “[...] language understanding systems” (615) and their operation as “machine comprehension” (Bender and Koller, 2022, p. 5185). The authors, however, emphatically claim that “no actual language understanding is taking place in LM-driven approaches to these tasks” (Bender et al., 2021, p. 615). Their claims have resulted in a lively discussion, fired up by the appearance of ChatGPT in late 2023 fired up and spilling over to disciplines outside AI and machine learning,¹ making the paper a remarkably influential criticism of LMs with the added feat of having coined the term ‘stochastic parrot’² to critically refer to LMs more generally.

In a previous paper, “Climbing towards NLU: On Meaning, Form, and Understanding in the Age of Data” (Bender and Koller, 2020), Emily M. Bender and Alexander Koller clarify the key concepts and arguments at the basis of the emphatic claim that LMs cannot understand language. Within the context of communicative exchanges in particular, Bender and Koller affirm that LMs cannot communicate because their linguistic productions lack communicative intent and reference to the real world, and cannot understand because they are trained to manipulate the form but not the meaning of language. In this paper, I first unpack and clarify Bender et al.’s concepts of communication and understanding in communication. Drawing on Jacques Derrida’s philosophy, I then raise a number of questions regarding the necessity of communicative intent and reference to the real world underpinning their concept of communication, and the idea of understanding as retrieving meanings. Next, taking recourse again to Derrida’s deconstruction I advance a concept of communication that retains the idea of transporting, transmitting, and production of signs while dispensing with the necessity of communicative intent and reference to the real world. This less restrictive concept does allow answering the question of whether LMs communicate and understand in the affirmative.

¹ A quick look at the statistics of this paper on Google Scholar shows that it has been cited over 3500 times in disciplines ranging from legal studies to education studies, from linguistics to environmental studies.

² The term ‘stochastic parrot’ was the AI-related word of the year 2023: <https://americandialect.org/wp-content/uploads/2024/01/2023-Word-of-the-Year-PRESS-RELEASE.pdf>



LANGUAGE MODELS AS STOCHASTIC PARROTS

The authors of “On the Dangers” outline three requirements of “human language use” without which communication could not take place: a) it happens between individuals who hold beliefs and other mental states –that is, they have attitudes towards propositions, for example, holding them true or not, desiring they were the case or not, etc.; b) they “share common ground and are mutually aware of that sharing (and its extent)” (Bender et al., 2021, p. 616); c) they use language to “convey” a “communicative intent” (p. 616). Condition (a) presupposes fully fledged language users, who feel at home in the business of linguistic exchange, and have had many such exchanges whereby they have formed a more or less coherent system of beliefs and other mental states. Condition (b) suggests that those partaking in the communicative exchange share a common world, namely “the real world the speaker and listener inhabit together” “against” which they can test the “truth” of their “interpretations” (Bender and Koller, 2020, p. 5187). The authors maintain, furthermore, that those participants in the communicative exchange are aware that they share a real world and the extent to which they share it. Finally, condition c) presupposes that individuals engage in communicative exchanges “for a purpose,” namely, “in order to achieve some *communicative intent*” (p. 5187). Communicative intent can be defined generally as wanting one’s speech or writing to do something, for example, inform, request, make another laugh, release frustration or anger, among many others. People’s intentions can be achieved through numerous means, and one of them is to use language in either spoken or written form. Speakers and writers choose some particular strings of words, then, that they think will allow them to communicate what they intend to communicate. In this picture, the linguistic articulation of the communicative intent carries ‘meaning’ – what the speaker means to say. That is, what makes a particular string of words meaningful is that the string was chosen with the expectation that it will do what the speaker or writer intends to do.

Understanding in communication, in turn, consists in “the process of retrieving [communicative intents] given [some strings of words]” (Bender and Koller, 202, p. 5187). Such retrieval requires the “ability to recognise interlocutors’ beliefs [...] and intentions [...] within context” (Bender et al., 2021, p. 616). In the process of understanding, the speaker/writer and interpreter are busy assigning words with meanings and attributing beliefs and other mental states to each other, and correcting these assignments and attributions until they both appear to ‘get the other’ and behave as expected in response. “Human-human communication”, they continue, “is a jointly constructed activity” (p. 616). This holds true not only for spoken language where speaker and interpreted are co-present, but also for written language where even “if we don’t know the person who generated the language we are interpreting, we build a partial model of who they are and what common ground we think they share with us and use this in interpreting their words” (p. 616).

Based on this normative framework, Bender et al. answer the question of whether LMs can understand language and communicate, which, given its content, inevitably leads to the foregone conclusion that they do not. Firstly, LMs’ linguistic production is not meaningful because the strings of words they produce convey no communicative intent. LMs, the authors state, “only have success in tasks that can be approached by



manipulating linguistic form” (Bender et al., 2021, p. 610) understood as “any observable realisation of language: marks on a page, pixels or bytes in a digital representation of text, or movements of the articulators” (Bender and Koller, 2020, p. 5187). The authors oppose form to meaning, which expresses itself in the linguistic articulation of a communicative intent. Not having communicative intent means that LMs do not *want* to do something with the strings of words they produce; rather, their linguistic outcomes are based on certain probabilistic operations. To clarify this, the authors explain how LMs produce the strings of words they do. To start with, the term *language model* (LM) “refer[s] to systems which are trained on string prediction tasks: that is, predicting the likelihood of a token (character, word or string) given either its preceding context or (in bidirectional and masked LMs) its surrounding context” (Bender et al., 2021, p. 611). N-gram models’ predictions were simpler because they were restricted by number of characters, directionality and horizontality of cues: “traditional n-gram LMs can only model relatively local dependencies, predicting each word given the preceding sequence of N words (usually 5 or fewer)” (p. 616). Coming after n-gram models, transformer models do not have previous restrictions, and are able to articulate language with impressive naturalness, “produc[ing] text that is seemingly not only fluent but also coherent even over paragraphs” (p. 616). LMs’ predictions are based on the data they are ‘trained’ with, allowing them “to perform apparently meaning-manipulation tasks such as summarisation, question answering, and the like” (p. 612) with notable success as they have excelled in numerous language understanding and common sense reasoning tests (p. 615). These successes, the authors insist, are technical, probabilistic rather than communicative successes. Bender and Köller affirm, “far from doing the ‘reasoning ostensibly required to complete the tasks, they [are] instead simply more effective at leveraging artefacts in the data” (Bender and Koller, 2020, p. 5186). In other words, they just shuffle data without comprehending what it is that they are shuffling or that they are shuffling it. This is shown in the fact that when they are trained with deliberately opposing data that contradict or negate some of the data they already have, their performance “falls to significantly below chance” (p. 5186).

Secondly, LMs’ linguistic productions are “not grounded on [...] any model of the world, or any model of the reader’s state of mind” (Bender et al., 2021, p. 616). For these authors, not having a world seems to suggest that these language models “have never observed” or otherwise interacted in any way with “the real world” (Bender and Koller, 2020, p. 5188). Having never interacted with the world, LMs cannot engage in any kind of meaningful linguistic behaviour with others at all because they do not have beliefs about the world nor can they attribute beliefs to speakers much less ‘model’ another’s ‘state of mind.’ The authors conclude, “contrary to how it may seem when we observe its output” a language model is “a system for haphazardly stitching together sequences of linguistic forms it has observed in its vast training data, according to probabilistic information about how they combine, but without any reference to meaning: a stochastic parrot” (Bender et al., 2021, p. 616-617). There is no reference to meaning because, as said, the strings of words they produce are “not grounded in communicative intent, any model of the world, of any model of the reader’s state of mind” (p. 616).



This conclusion might appear obvious, and it does seem so to the authors of “On the Dangers.” Yet, when we attend to how individuals interact with LMs, the authors find that it is not all that obvious. Looking at the side of human beings in their exchanges with LMs, the authors identify human beings’ “tendency” to regard LMs’ linguistic production as meaningful, as carrying communicative intents, which further contributes to the misrepresentation of LMs as natural language understanding models³: “the tendency of human interlocutors to impute meaning where there is none can mislead both NLP researchers and the general public into taking synthetic text as meaningful” (Bender et al., 2021, p. 611). It is the “tendency” to attribute communicative intentions to others, the first of which is the intention of being understood which LMs lack but human beings misattribute to LMs. But this tendency goes much further, for by attributing communicative intent, human beings are attributing mental states to LMs along with a model of the world, and, what is more, a similarity to themselves, which is something human beings do in their communicative exchanges generally and insofar as they engage in communicative exchanges. The authors regard these attributions as “an illusion arising from our singular human understanding of language” (p. 616). Despite this blatant illusion, human beings, the authors continue, are “very willing... to attribute communicative intent even if the originator of the signal is not an entity that could have communicative intent” (Bender and Koller, 2020, p. 5187). Their ‘willingness’, however, is unwarranted.

RECONSIDERING COMMUNICATIVE INTENT AND REFERENCE AS CONDITIONS OF COMMUNICATION AND UNDERSTANDING

After this brief clarification of the concepts and arguments presented in “On the Dangers,” it should be clear that this framework is evidently applicable only to human-to-human linguistic exchanges, specifically involving adult humans. It fails to account for entities capable of producing strings of words but lacking beliefs, intentions and interactions with the human world such as certain animals and machines. This concept of communication is overly restrictive not only because it excludes entities other than humans, but also because it only captures certain instances of human-to-human communication. In the following discussion, I problematise the premises and concepts of the authors’ arguments from the perspective of Jacques Derrida’s philosophy, particularly focusing on the two main conditions of communication and understanding in communication: communicative intent and model of the world. Then, I articulate alternative concepts of communication and understanding in order to revisit the question of whether LMs and chatGPT communicate.

a) *Communicative intent*

The conceptual and argumentative framework presented in the previous section is explicitly based on the Gricean model of communication (Bender and Koller, 2020, p. 5187), where intentions play a decisive role in the entire communicative process, from the selection of strings of words or noises to how the communicative process itself unfolds

³ LMs’ developers are also responsible for this confusion, for they tend to describe LMs as ‘understanding’. Cf. Bender and Koller, 5185-5186.



and whether the communication can be deemed successful. A version of this view is largely accepted by philosophers of language in the analytic tradition from Searle to Davidson, albeit with some important differences (Searle, 1977; Davidson, 1992). For Bender et al., communicative intent or communicative intention presupposes a person deliberately initiating a communicative exchange with the purpose of doing something or affecting the interpreter in a particular way. This purpose is their communicative intent, and it makes the strings of words or noises in a linguistic exchange meaningful insofar as these strings were selected as a means to achieve the desired effect on the interpreter.

Derrida notes, however, that no such person imbuing a string of words or noises with communicative intent is necessary for signs to be meaningful and have the expected effects on interpreters. More generally, no communicative intent is necessary for signs to be meaningful. Signs should be able to function, that is, to be interpretable in meaningful ways, even if the producer of the signs is not present or has never existed. “A mark,” he says, is “a sort of machine which is productive in turn” and “must continue to ‘act’ and to be readable even when what is called the author [...] no longer answers for what he has written [...] be it because of a temporary absence, because he is dead, or because he has not employed [...] the plenitude of his desire to say what he means” (Derrida, 1988, p. 8). This is possible, Derrida continues, because, in order to be such, signs have to be repeatable, that is, any sign can be weaned from its putative or otherwise context of production and placed in an altogether different context without hindering its possibility of being meaningful.

In his response to Derrida’s “Signature, Event, Context,” Searle takes up and discusses this argument. Although he agrees with the repeatability of signs, he disagrees with the conclusion Derrida draws from it (Searle, 1998, p. 201-202). He concedes that even if it were the case that no producer is necessary for a meaningful communication, understanding and interpreting signs would still require assuming a communicative intent “because a *meaningful sentence is just a standing possibility of the corresponding (intentional) speech act*” (Searle, 1998, p. 202). Understanding, he says, is just “knowing what linguistic act its utterance would be a performance of,” even if there was no actual utterance. In these cases, strings of words or noises are meaningful and can, thus, be presumed to have communicative intent if they follow the rules of language. “To understand it [a speech act], it is necessary to know that anyone who said it and meant it would be performing that speech act *determined* by the rules of language that give the sentence its meaning in the first place” (Searly, 1998, p. 202, my emphasis). In other words, Searle maintains that it is through the rules of language that a hypothetical communicative intent can be articulated. For example, when a chat bot produces the lines ‘Provide your name, email address and order details’ in an automated way, there is no actual individual imbuing these phrases with communicative intent. However, the use of the imperative directs the interpreter to the rules governing imperatives, from which an intention can be inferred. Specifically, the intention in this example is to be authoritative and to prompt the interpreter to comply by providing the requested information. It is rules of language – or conventions, as Searle sometimes calls them – that confer intent or purpose to spoken or written marks, thereby rendering them meaningful. It is through these rules that a communicative intent can be discerned. In contrast to the view of



communicative intent advanced by the authors of “On the Dangers,” Searle’s concept of communicative intent does not necessarily involve an actual person intending to convey a meaning by selecting the strings of words or sounds that are most likely to articulate that intent. Instead, it is regarded as a “strategy of understanding” – a helpful presupposition that facilitates the interpretation and comprehension of written and spoken marks (Searle, 1998, p. 202).

Searle’s way of sidestepping Derrida’s argument dispenses with the necessity of a producer of communication without sacrificing meaningfulness and communicative intent. However, inadvertently, his position seems to align more closely with Derrida’s and further from Bender et al.’s, as both Searle and Derrida agree that actual individuals or, more broadly, human beings linguistically articulating communicative intents are not essential for communication to occur. This agreement challenges Bender et al.’s first condition of communication, which states that communication occurs between individuals possessing beliefs and other mental states. Derrida would also agree with Searle’s assertion that intention is a presupposition orienting interpretation and understanding rather than the necessary key to interpretation and understanding when he says, “the category of intention will not disappear, it will have its place, but from that place it will no longer be able to govern the entire scene and system of utterance” (Derrida, 1988, p. 18) Bender et al. would actually perceive Searle’s way of salvaging communicative intent when applied to LMs as part of the problem, for the mischaracterisation of LMs as language understanding systems is fostered partly by presupposing and attributing communicative intent where there is none.

Searle’s position shifts our focus towards the interpreter rather than the speaker or initiator of communication. From the interpreter’s perspective, whose first task is deciphering the communication, communicative intent, as mentioned earlier, is possible but not necessary, and functions as a presupposition assisting in interpreting linguistic productions. The crux of Searle’s viewpoint lies in the rules of language, particularly grammar as logical syntax, whose correct usage aids in articulating an intention which may or may not have been actual. This view is, however, fairly easy to question, for the idea that language use requires knowing or even applying grammar correctly is unwarranted. Language acquisition and use are primarily practical and occur without the explicit need to learn grammar rules. Furthermore, using language ‘incorrectly’ from a grammatical standpoint does not necessarily hinder communication as it is common to successfully interpret speech containing grammatical errors.

Derrida delves into this topic, discussing not just isolated grammatical errors, but agrammaticality – instances where there is no longer “‘logical’ language” (Derrida, 1988, p. 11) Even these cases, Derrida affirms, need not compromise communication. The reason is that since cases of agrammaticality “[f]or instance, ‘the green is either’ or ‘abracadabra’”, he says, “do not constitute their context by themselves, nothing prevents them from functioning in another context as signifying marks” (p. 12). This underscores Derrida’s initial argument that in order to function and continue to be interpretable and meaningful, signs have to be capable of separation from their context of inscription. Thus, ‘the green is either’ could be inscribed in semantic or real contexts where it could be meaningful. Derrida’s argument here challenges the necessity of rules of language for



communication and understanding in communication. This does not imply that language rules are never useful in interpretation; rather, they are not a sufficient and necessary condition of communication and understanding.

Bender et al. could strengthen their position by drawing on Davidson's reflections on language use and communicative exchanges. In Davidson's work, they could find an alternative defence of intentions in communication based on another, arguably more important, function of intentions. He affirms that "[t]he necessary presence of intentions would be significant, since it would give content to an attribution of error by allowing for the possibility of discrepancy between intention and accomplishment" (Davidson, 1992, p. 259). Speakers initiate communicative exchanges not necessarily with a single intention as they likely aim to convey something to provoke a response or behaviour, which in turn may lead to further outcomes. However, we can narrow down this array of intentions to the primary and fundamental intention of being understood. It is possible though that this intention is not fulfilled. For various reasons, such as the speaker incorrectly assuming that the listener would grasp crucial cues, misinterpreting the context, or misjudging the listener's knowledge or willingness to interpret in certain ways, the chosen words or sounds may fail to achieve the intended communicative effect. In such situations, intentions serve to highlight the discrepancy between intention and actual outcome of the communication, ranging from explicit acknowledgement of misunderstanding to the disparity between the behaviour the speaker anticipated and the actual response from the listener. Davidson argues that without communicative intent, there would be no means of spotting errors, i.e., unsuccessful communications. Or, more consequentially, there would be no mistaken interpretations as any interpretation would probably be good enough.

We could ask, however, is it not rather common that we cannot test our interpretations against the speaker's/writer's intentions? In Davidson's picture, the speaker and interpreter are facing one another, which vastly facilitates testing the interpretation. However, it does not guarantee that an error, if there is one, will be spotted because it might be the case that the speaker, as Derrida affirms, "has not employed his absolutely actual and present intention or attention, the plenitude of his desire to say what he means" (Derrida, 1988, p. 8), or has conflicting intentions or is not fully conscious of his intentions and cannot respond for what he has said or written. Testing interpretations becomes even more challenging in various cases of communication, such as reading the newspaper, listening to someone's voice message, interpreting the work of a long deceased author, perusing personal journal entries of years past, and the list could go on. In these cases, interpreters cannot double check their interpretation against the intentions of the speaker/writer. If intention were the norm of interpretation, we would have to acknowledge that misinterpretation is highly possible, and probably factually common, for it is simply not an element that can be relied upon in all cases of communication, either because it is factually absent or because the communication was not intended with full attention, among many other possible reasons. This does not mean that intention, if there is one, can never serve as a guiding post orienting interpretation and aiding in identifying errors of interpretation. At times, it may serve this purpose. As suggested



before, however, it cannot be regarded as a necessary condition for communication and understanding in communication.

In this section, I have discussed intentions as a condition for communication on various grounds: because communication is to have a purpose, because it serves as a 'strategy of understanding' or presupposition orienting understanding, and because it functions as a norm sanctioning interpretation, which closely aligns with Bender et al.'s definition of understanding as 'retrieving communicative intent'. If we consider, in relation to the first of these grounds, that the primary purpose or intention of communication is to be understood, then all of these grounds emphasise the role of intentions for the possibility of understanding, that is, for the possibility of a successful interpretation of what is communicated. Here, understanding means successful interpretation where intended meaning matches interpreted meaning. Drawing from Derrida's views, however, I argued that this matching cannot be deemed necessary because the process of matching presupposes an intention which may or may not be present and reliable. Thus, it would be useful to rethink the concept of understanding in communication in such a way that it is not defined necessarily by this matching. More concretely, it should be a concept of understanding that does not necessarily consist in the fulfilment of intention insofar as communicative intent may or may not be present and reliable, and that may rely on semantic and real contexts without regarding them as fixing interpretation. Such a concept of understanding could probably be characterised as our *best guess* about what the communication is about and aims to cause. It is a *guess* because there is no single factor that interpretation can rely upon in all cases and with full certainty. But it is the *best* guess because factors such as possible intentions, possible contrasting the interpretation with the speaker's intention, rules of language, semantic and real context, previous experiences of communication, etc., can sometimes assist to a greater or lesser degree in orienting interpretation, and which can be appealed to to justify one interpretation over another.

This concept of understanding is less restrictive than Bender et al.'s and, I want to say, more immediately applicable to beings other than human beings such as animals and LMs. Focusing on LMs, Bender et al. state that LMs can neither mean something nor retrieve meaning, that is, they can neither communicate nor understand because their linguistic outputs are the result of probabilistic operations indicating the likelihood –not the meaningfulness– that certain strings of words follow the input. Yet, when understanding is defined not as meaning retrieval, matching of intended and interpreted meaning, or strict grammar rule following, but rather as the best guess, namely, as the most probable interpretation, could not LMs be said to understand in this sense? Before discussing this question, I would like to move on to the second condition of communication and understanding at the basis of Bender et al.'s emphatic claim that LMs can neither communicate nor understand in communication.

b) *The real world or a model of the world*

The second condition necessary for communication and understanding in communication, as discussed by Bender et al. in "On the Dangers" is what they term 'a model of the world.' Bender and Koller do not talk about the model of the real world but rather about the real world, so I will treat them interchangeably for the time being. The



‘real world’ or ‘the model of the real world’ has two interrelated functions in Bender and Koller. Firstly, it is that which communications are about, and, secondly, they are that against which the truth of communications can be tested, but they focus exclusively on the former. The bulk of my criticisms rest on the fact that Bender et. al. and Bender and Koller provide insufficient reasons to restrict the scope of communicational topics to objects in the real world. In addition, I argue that LMs can be said to have a model of the world.

In “On the Dangers”, the model of the world is vaguely referred to as the “common ground” which speakers/writers and interpreters share (Bender et al., 2021, p. 616). In “Climbing Towards NLU: On Meaning, Form, and Understanding in the Age of Data” Bender and Koller clarify the notion of ‘real world’ and spell out in more detail how it features in communicative exchanges. It first appears in relation to communicative intents: the authors maintain that “communicative intents are about something that is *outside language*” (Bender and Koller, 2020, p. 5187, italics in the original). The examples given include “*Open the window!* [and] *When was Malala Yousafzai born?*” (p. 5187, italics in the original). Communicative intents “can also be about abstract worlds, e.g. bank accounts, computer file systems” (p. 5187). These examples are quite evidently about something non-linguistic, in particular windows, Malala Yousafzai, bank accounts and computer file systems, and in this sense they could be said to be ‘outside’ language. Bender and Koller do not mean it in this sense, though. Rather, they claim that in these examples, “the communicative intent is grounded in the real world the speaker and listener inhabit together” (p. 5187). They do not explain what ‘grounding’ here means, but they do state that this grounding is at the basis of the role ‘the real world’ or ‘the model of the real world’ performs in communicative intents. It is strange, however, that immediately following this assertion, the authors state that communicative intents can also be about “a purely hypothetical world in the speaker’s mind” (p. 5187), because a purely hypothetical world in the speaker’s mind does not seem to be grounded in the real world if ‘grounding’ is meant to signify a relation of reference whereby strings of marks are about objects in the real world. Setting this aside for the moment, the ‘real world’ or the ‘model of the real world’ features in yet another instance, namely, linguistic systems.

Linguistic systems “provid[e] a relation [...] which contains pairs (e , s) of expressions e and their conventional meanings s ” (Bender and Koller, 2020, p. 5187). Conventional meanings seem to be standardised meanings, and the authors defined them as “what is constant across all of its possible contexts of use” (p. 5187) and “an abstract object that represents the communicative potential of a form, given the linguistic system it is drawn from” (p. 5187). Linguistic systems also relate to the real world, which is what seems to be described as ‘outside language.’ They say, linguistic systems “connec[t] language to objects outside of language” (p. 5187). So, in communication, the speaker is ‘grounding’ her/his speech on the ‘real world’ or the ‘model of the real world’ from two angles: the speaker has a communicative intent grounded in the real world, which is conveyed through the use of expressions of the linguistic system that itself is also grounded in the real world. Now, the listener shares the real world or the model of the real world with the speaker, and also largely (p. 5187 n 6) shares the linguistic system



with the speaker. This twofold sharing helps the listener to retrieve the communicative intent of the communication.

In order to illustrate the roles the real world or the model of the real world plays in communicative exchanges, Bender and Koller describe a particular scenario. In this scenario, two English speakers are stranded on two separate isolated islands, but luckily, previous inhabitants of these islands left behind telegraphs that they can use to communicate, and they start to use it routinely to have all kinds of conversations. Unbeknownst to them, an exceptionally smart octopus living underwater finds “a way to tap into the underwater cable and listen in on A and B’s conversations” (Bender and Koller, 2020, p. 5188). The octopus “is very good at detecting statistical patterns” (p. 5188), so it manages to identify numerous patterns in how A and B speak. One day, the octopus poses as B and responds to A’s messages. The question the authors ask is whether the octopus can “successfully pose as B without making A suspicious?” (p. 5188). The answer, they continue, depends on what the conversation is about: if A’s “utterances [...] have a primarily social function, and do not need to be grounded in the particulars of the interlocutors’ actual physical situation nor anything else specific about the real world” (p. 5188), then the octopus might actually manage to pull it off because in this case “it is sufficient to produce text that is internally coherent” (p. 5188). If, on the other hand, A’s conversation refers to something in the world, then the octopus is likely to struggle to produce speech that is meaningful. For example, “A [...] is suddenly pursued by an angry bear. She grabs a couple of sticks and frantically asks B to come up with a way to construct a weapon to defend herself” (p. 5189). This task, the authors continue, “requires the ability to map accurately between words and real-world entities (as well as reasoning and creative thinking). It is at this point that [the octopus] would fail” (p. 5189). The reason for the failure is that the octopus “has never observed these objects, and thus would not be able to pick out the referent of a word when presented with a set of (physical) alternatives” (p. 5188). The octopus has no model of the world or experience of the real world A experiences, so its responses will be meaningless.

A couple of questions could be raised here. Firstly, the reason the octopus statistician can only produce meaningless responses to A is that it has not experienced or observed the world. From this, we can conclude that observation and experience are crucial for the possibility of constructing a model of the world or grounding communications in the real world. This is, however, evidently not always the case. A blind person can know a lot about blue skies and be able to talk at length about them without ever having seen blue skies. Certainly, this person could be said to have experienced blue skies vicariously through others, but this person has certainly not *observed* blue skies as such. So, observation cannot be regarded as necessary for the possibility of having a model of the world or talking about something. Derrida discusses this point in relation to Husserl’s first consideration of the absence of the referent. He says, “An utterance whose object is not impossible but only possible can very well be made and understood without its real object (its referent) being present, either to the person who produced the statement or to the one who receives it” (Derrida, 1988, p. 10). For example, a person could say “The sky is blue” and this utterance would be intelligible and interpretable even if neither the speaker nor the interpreter see the sky, if



the speaker is mistaken or is lying. This is clearly not always the case, “but the structure of possibility of this utterance includes the capability to be formed and to function as a reference that is empty or cut off from its referent” (Derrida, 1988, p. 10-11). Without this possibility, Derrida contends, signs would not function and be readable and interpretable.

Secondly, Bender and Koller maintain that communicative intents, as well as linguistic systems, are grounded in the real world, in the sense that words articulating intents and forming expressions refer to individualisable objects in the world. Bender and Koller are quick to make a disclaimer to the effect that this relation of reference is not a relation of grounding truth. They state, “we should be careful not to confuse communicative intent with ground truth about the world, as speakers can of course be mistaken, be intentionally dissembling, etc.” (Bender and Koller, 2020, p. 5187). If the relation of reference is not a relation intended to ground the truth of statements, then it could be thought of as a relation of ‘aboutness’ in the sense that the objects usually defined as constituting the world can be objects of descriptions, topics of conversation, say. But if that is the case, then it is unclear why Bender et al. limit topics of conversation to objects in the real world, especially because their octopus story clearly oversteps that limit insofar as there are no octopuses versed in statistics that we know of and they have no referent in the real world. Derrida's discussion of Husserl's second consideration of the absence of the referent clarifies why limiting meaningfulness to what can be referred to in the real world is not justified (Derrida, 1988, p. 11): Husserl analyses the “absence of the signified” (in three instances: a) signs can be manipulated and intelligible without them referring to anything, for example, “mathematical symbolism” (p. 11); b) “[c]ertain utterances can have meaning although they are deprived of *objective* signification” (p. 11). The example Husserl gives is ‘the circle is square’. This phrase, Derrida continues, “has sufficient meaning at least for me to judge it false or contradictory” (p. 11); c) “what Husserl calls *Sinnlosigkeit* or agrammaticality. For instance, ‘the green is either’ or ‘abracadabra’” (p. 11). In these examples “there is no more ‘logical’ language” (Derrida, 1988, p. 11). However, these phrases can very well be placed in contexts where they will be meaningful. These considerations show that communications do not have to be about the real world.

Now, as noted earlier, Bender et al. talk about ‘the model of the real world’ whereas Bender and Koller about ‘the real world,’ and there are reasons to think that these expressions are not interchangeable. LMs and chatGPT can be said to lack a world and a relation with the real world in the sense that they cannot sit down on a chair or buy a train ticket, and so are unable to form beliefs, memories, desires, etc., about the world. Yet, LMs and chatGPT have huge amounts of data about the world and, in that sense, they could be said to have a model of the world. Insofar as the data they are trained with is not tested for coherence or truth, it is probable that their models of the world are not particularly coherent. However, neither are those of human beings, which can at best be described as largely coherent rather than as fully coherent.



RECONCEPTUALISING COMMUNICATION, UNDERSTANDING AND THE MODEL OF THE REAL WORLD

A point continuously made by the authors of “On the Dangers” is that human beings are mistaken in treating their interactions with LMs as communicative exchanges. They regard the “tendency” to treat LMs’ linguistic productions as meaningful as an “illusion” (Bender et al., 2021, p. 616), and a “deception” (Bender and Koller, 2020, p. 5189). The norm against which they make these claims is the view of communication we have been discussing, which requires utterances to be grounded in communicative intents and the real world, and involves numerous cognitive processes like attributing beliefs, retrieving intentions, assigning meanings to words, testing their correctness, etc., which LMs clearly cannot do. It is worth asking, however, whether individuals’ description of their exchanges with LMs can provide additional reason to justify a transformation of the concepts of communication and understanding to make room for the novel forms of interactions individuals are having with these new technologies.

In an opinion piece published in *Globe and Mail*, Derek Ruths states, “The way we interact with ChatGPT is virtually identical to the way we communicate with people every day” (Ruths, 2023). His examples are significant: “on a screen, with a small text box, viewing a scrolling window of dialogue[, t]he standard ChatGPT interface looks like WhatsApp, SMS, Apple Messages, and every other messaging app” (Ruths, 2023). Indeed, the interfaces of messaging apps are basically indistinguishable from ChatGPT’s interface which explains why exchanges with ChatGPT feel ‘virtually identical’ to many of the communications people have with other human beings through these apps. For younger generations, this kind of virtual exchanges, rather than face-to-face interactions, is actually the norm not only because they are digital natives, but also because they grew up during the COVID-19 pandemic where online communication was the predominant mode of communication. For these younger generations, interacting with ChatGPT does not ‘feel’ significantly different to the online interactions they have with their friends and relatives. And it is certainly not only younger generations that ‘feel’ this way. It is not only the similarity of ChatGPT’s interface to those of other apps that contribute to these feelings of similarity. There are also signs which individuals would normally interpret as indicating that ChatGPT is performing some cognitive activity: “ChatGPT even generates little thought typing bubbles while it’s working up its response” (Ruths, 2023). Those bubbles also appear when a real person is typing a message, regardless of whether the person is actually doing any thinking.

In this article, I have offered reasons to support changing the concept of communication. In particular, I argued that communicative intent should not be taken as a condition for the possibility of communication and understanding in communication because it is not always present or reliable in vast cases of communication, even in face-to-face communication. This does not exclude intent *tout court* as it will be useful to articulate communications and orient interpretation in some cases. A similar argument was made regarding the real world and its grounding function. It was argued that communication can very well be about hypothetical and illusory objects, without



hampering meaningful interpretations. The article has also hinted at some possible ways forward, which I briefly discuss in this last section.

We are seeking a concept of communication that does not necessarily require communicative intent and is not necessarily grounded in the real world. Derrida finds in Austin's theory of speech acts a concept of communication that can meet these conditions with some modifications. He says that in Austin's analysis, communication acts “do not designate the transference or passage of a thought-content, but, in some way, the communication of an original movement (to be defined within a *general theory of action*), an operation and the production of an effect” (Derrida, 1988, p. 13). The performative utterance, Derrida continues, “would be tantamount to communicating a force through the impetus (*impulsion*) of a mark.” Moreover, “the performative does not have its referent ... outside of itself or, in any event, before and in front of itself. It does not describe something that exists outside of language and prior to it. It produces or transforms a situation, it effects” (Derrida, 1988, p. 13). The performative utterance is exemplary because, in contrast to constative utterances, it does not have to transmit a meaning, refer to something in the world or assert some truth about objects in the world or state of affairs in the world.

The original movement or operation Derrida talks about in the quotation is elaborated on in the first chapter of *Of Grammatology*, where Derrida recounts the transformation of the philosophical concept of language. Initially defined as a mediating tool between subject and world, alternatively expressing the meaning intended by the former or representing objects in the latter, language is reconceptualised as an operation of “demotivation” (Derrida, 1976, p. 51) and “overwhelming” (p. 7) production of signs with various degrees of repeatability (differential force). As an operation of demotivation, the repetition of signs erodes any intention or reference (if there was one) constraining interpretation. Demotivation and overwhelming do not render signs unintelligible or inscrutable. Instead, by eroding the limit imposed by an intention or an object in the world, the operation of demotivation makes signs ‘overwhelming’ in that possibilities of interpretation (or possibilities of understanding in communication) increase potentially to infinity. This proliferation of possibilities alters the world by expanding the realm of possible interpretations. While some interpretations may possess more ‘force’ than others at times, meaning they are more likely to become binding or authoritative, the overwhelming nature of signs persists as other interpretations remain possible even if they have less force, less chance of becoming binding or authoritative at certain times.

The differential force of interpretations derives from such factors as the possible intention, the semantic context, the real context, the rules of language, past communicative experiences, among others. It was said that none of these elements is necessary for the possibility of communication and understanding, but that none of them is excluded either, for they are all possible factors that might contribute to tipping the interpretative balance in one direction or another. Earlier, I proposed that understanding could be defined as ‘the best guess.’ The notion of ‘guess’ is appropriate here because the demotivated and overwhelming character of signs challenges the idea that there is only one interpretation, *the* correct interpretation that can be easily retrieved by following a sure method. Without one single interpretation, the aforementioned factors can incline



interpreters toward one interpretation, but they don't guarantee it. Thus, the most interpreters can offer is a guess, their best guess, given the factors contributing to the force of their interpretation. Defining understanding as the best guess would allow us to regard LMs' outputs as understanding, for their outputs are also a guess based on likelihood rather than meaning. Finally, it was suggested that although LMs lack experience of the real world, they could be said to have a model of the world. In their case and in contrast to human beings, their model of the world is not a system of beliefs related inferentially. Instead, it comprises the information, the data about the world with which they have been trained. LMs' outputs are based on this information, which, just like human beings' models of the world, is more or less coherent.

If these concepts of communication, understanding in communication, and model of the world are accepted, then it would be possible to reconsider the question of whether LMs communicate and understand in communication, and provide an affirmative answer.

CONCLUSION

This paper discusses Bender et al.'s and Bender and Koller's concepts and arguments underlying their negative answer to the question of whether LMs communicate and understand, in particular the concept of communicative intent and model of the world or real world. I argued that communicative intent cannot be regarded as a condition for the possibility of communication because it is absent or unreliable in vast cases of communication. I also argued that communication and understanding need not be grounded in the real world if that implies limiting communication to what can be referred to in the real world. Having discarded the necessity, although not the possibility, of these two factors, alternative concepts of communication and understanding were elaborated. Drawing from Derrida's discussion of Austin's speech act theory and his own philosophy, the alternative concepts of communication and understanding retain the idea of communication as transmitting and affecting. Yet, what is transmitted is an operation of demotivation and overwhelming of possible interpretations with differential forces. Understanding was defined as the best guess rather than as retrieving *the* correct interpretation, which is aided by a model of the world defined as more or less coherent information. Armed with these concepts, the paper advances an affirmative answer to the question of whether LMs communicate and understand.

REFERENCES

- Bender, E., & Koller, A. (2020). Climbing towards NLU: On Meaning, Form and Understanding in the Age of Data. In D. Jurafsky, J. Chai, N. Schluter, & J. Tetreault (Eds.), *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics* (pp. 5185-5198). Association of Computational Linguistics. <https://doi.org/10.18653/v1/2020.acl-main.463>
- Bender, E., Gebru, T., McMillan-Major, A. & Shmitchell, S. (2021). *On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?*. In *FACCT '21: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (pp. 610-623). ACM. <https://doi.org/10.1145/3442188.3445922>



- Davidson, D. (1992). The Second Person. *Midwest Studies in Philosophy*, 17(1), 255-267.
<https://doi.org/10.1111/j.1475-4975.1992.tb00154.x>
- Derrida, J. (1988). Signature, Event, Context. In *Limited Inc* (pp. 1-24). Northwestern University Press.
- Derrida, J. (1976). *Of Grammatology*. The John Hopkins University Press.
- Ruths, D. (2023, May 19). ChatGPT is Blurring the Lines between what it Means to Communicate with a Machine and a Human. *The Globe and Mail*.
<https://www.theglobeandmail.com/opinion/article-chatgpt-is-blurring-the-lines-between-what-it-means-to-communicate/>
- Searle, J. (1997). Reiterating the Differences: A Reply to Derrida. *Glyph Review*, 2, 198-208

СВЕДЕНИЯ ОБ АВТОРЕ / THE AUTHOR

Ребека Перес Леон, drrpleon@gmail.com
ORCID 0000-0003-4912-4912

Rebeca Perez Leon, drrpleon@gmail.com,
ORCID 0000-0003-4912-4912

Статья поступила 3 марта 2024
одобрена после рецензирования 1 июня 2024
принята к публикации 10 июня 2024

Received: 3 March 2024
Revised: 1 June 2024
Accepted: 10 June 2024



<https://doi.org/10.48417/technolang.2024.02.05>

Research article

Dialogue as Autocommunication - On Interactions with Large Language Models

Anna Kartasheva  

National Research University Higher School of Economics, Moscow, 11 Pokrovsky Bulvar, 109028,
Russian Federation

akartasheva@hse.ru

Abstract

In a dialog with large language models (LLM) there is a coincidence of the addressee and addressee of the message, so such a dialog can be called autocommunication. A neural network can only answer a question that has a formulation. The question is formulated by the one who asks it, i.e. a human being. Human activity in dialog with neural networks provokes thoughts about the nature of such dialog. Composing prompts is one of the most creative parts of dialog with neural networks. But it is worth noting that a neural network is often better at composing prompts than a human. Does this mean that humans need to develop their questioning skills? In LLM-based dialog systems, the main value to the user is the ability to clarify and structure their own thoughts. The structuring of thoughts happens through questioning, through formulating and clarifying questions. Asking the right question is practically answering that question. Thus, thanks to autocommunication, the development, transformation, and restructuring of the human "I" itself takes place. Dialogue with large linguistic models acts as a discursive practice that allows people to formulate their own thoughts and transform their self through autocommunication. It is worth noting that for this kind of dialog, a certain image of the audience is normative or determinative of the material that can be produced in response to a given question. This is because the data for model training is provided by people, even if they do not and have never thought about it. Thus, a dialogic relationship develops between the generated text and the questioning audience that develops all participants in the communication.

Keywords: Large Language Models, Autocommunication, Artificial intelligence, Authorship, Communication, Dialogue

Acknowledgment The article was prepared within the framework of the HSE University Basic Research Program.

Citation: Kartasheva, A. (2024). Dialogue as Autocommunication - On Interactions with Large Language Models. *Technology and Language*, 5(2), 57-66. <https://doi.org/10.48417/technolang.2024.02.05>



This work is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/)



УДК 1: 004.8

<https://doi.org/10.48417/technolang.2024.02.05>

Научная статья

Диалог с LLM как аутокоммуникация

Анна Карташева  

Национальный исследовательский университет Высшая школа экономики, Москва, 11 Покровский бульвар,
109028, Российская Федерация

akartasheva@hse.ru

Аннотация

В диалоге с большими языковыми моделями (LLM) происходит совпадение адресата и адресанта сообщения, поэтому такой диалог можно назвать аутокоммуникацией. Нейросеть может ответить только на вопрос, имеющий формулировку. Формулирует вопрос тот, кто спрашивает — то есть, человек. Активность человека в диалоге с нейросетями провоцирует на размышления о природе такого диалога. Составление промптов является одной из самых творческих частей общения с нейросетями. Но, стоит отметить, что нейросеть зачастую лучше справляется с составлением промптов, чем человек. Значит ли это, что человеку необходимо развивать свои навыки вопрошания? В диалоговых системах, построенных на LLM, основной ценностью для пользователя является возможность прояснить и структурировать собственные мысли. Структурирование мыслей происходит через вопрошание, через формулировку и уточнение вопросов. Задать правильный вопрос – это уже практически ответить на этот вопрос. Таким образом, благодаря аутокоммуникации происходит развитие, трансформация, перестройка самого “Я” человека. Диалог с большими лингвистическими моделями выступает дискурсивной практикой, позволяющей людям формулировать свои собственные мысли и трансформировать свое “Я” через аутокоммуникацию. Стоит отметить, что для такого диалога определенный образ аудитории является нормирующим или определяющим тот материал, который можно получить в ответ на заданный вопрос. Это происходит потому, что данные для обучения моделей предоставляют люди, даже если они не задумываются и никогда не задумывались об этом. Таким образом, между генерируемым текстом и вопрошающей аудиторией складываются диалогические отношения, которые развивают всех участников коммуникации.

Ключевые слова: LLM, Аутокоммуникация, Искусственный интеллект, Авторство, Коммуникация, Диалог

Благодарность: Статья выполнена в рамках Программы фундаментальных исследований НИУ ВШЭ

Для цитирования: Kartasheva, A. Dialogue as Autocommunication - On Interactions with Large Language Models // Technology and Language. 2024. № 5(2). P. 57-66.
<https://doi.org/10.48417/technolang.2024.02.05>



This work is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/)



INTRODUCTION

Communication with a neural network is a person's communication with oneself because a neural network can only answer a formulated question, a question that is asked to it in one form or another. It would be simplistic to say that we as humans lose something important from communicating with Large Linguistic Models (LLMs). Rather, we are entering into a dialogue of a different quality, of a different order. What are the specific features of this dialogue? What is the scope of this dialogue? There is no doubt that written or oral communication with neural networks raises reflections on the nature of communication, authorship, and identity.

A new way to ask questions to neural networks can be presented as autocommunication when the addressee and sender of the message coincide. At the same time, if earlier autocommunication was manifested in channels created according to the "I – I" model, today we see that in social networks, and even more so in dialogue systems created based on neural networks, channels that were initially created are used for autocommunication for communication according to the "I – He" model. For example, the interactive chatbot model itself initially assumes that the user communicates with some external authority, receives information, and answers queries. And in the case of dialogue systems created based on neural networks, the main value for the user is the opportunity to clarify their own thoughts and structure them.

Using generated text as your own is an inherently disastrous idea. First, any large linguistic model often produces very dense text as its answer. This is a kind of wall of text that you can't get close to. The wall of text is too correct, overly correct, annoyingly correct. A person does not write so precisely, does not thoroughly clarify every detail. A person writes unevenly – and the text is "living," "breathing." And the reader feels this, even if he cannot explain what exactly in the generated text confuses him. Accordingly, he weakens his attention and stops visiting those resources where a large amount of generated text is posted. Respecting and retaining their readers and users, many large companies place mentions that this text (or picture) was generated by artificial intelligence systems. And such a practice becomes a practice of good form for the entire society. We should note that a legal problem arises here: who is the owner of the generated texts. Ideas are being expressed and bills are being proposed regarding mandatory labeling of content created by a neural network.

Secondly, we know many examples where neural networks produced hallucinations. The picture of the world loaded through data into the model is not capable of being completely consistent, nor is it endowed with the ability to automatically adjust displacements and distortions. Thirdly, the texts of neural networks are full of words that are characteristic of those people who acted as instructors for neural networks and created the rules by which the neural network then began to produce answers.

Autocommunication in relation to neural networks is expressed in the fact that a person formulates a question, already assuming a certain answer. The answer may not be about the facts, but about the very structure of the answer: there is some expectation, otherwise it would be very difficult to formulate the question.



We can say that dialogue with larger linguistic patterns is a kind of discursive practice that allows people to formulate their own thoughts. Most often, this practice is used not to formulate finished texts, but, for example, to formulate some sketches or references that may change in the future.

LITERATURE REVIEW

In the 1960s Roman Jakobson proposed to give a key role in communication to language, as well as to understand communication as a speech event (Jakobson, 1985). Any message is created and interpreted using a code – language, while the transmission of messages takes place in the field of discourse, in the context of other received and transmitted messages. This model of communication by Jakobson was challenged by Yuri Lotman, pointing out that two people cannot have absolutely identical codes, and language should be understood as a code along with its history (Nazarchuk, 2009). Text by Lotman considers it as a substrate of communication, in which a meeting of many codes and communicants occurs.

The concept of sign is important for the semiotic approach. Ferdinand de Saussure (1977) considered a sign as a two – way mental formation that connects a concept (signified) and an acoustic image (signifier). For further research, Saussure's idea about the systemic nature of language and other communication systems turned out to be extremely important – the meaning is supported by the mutual connection of all elements of the system. Edward Sapir (1993) made a distinction between primary processes, communicative in nature: language, gestures, etc., and some secondary means that facilitate the communication process: linguistic transformations, the creation of physical conditions for the implementation of a communicative act.

Foreign researchers call autocommunication or internal dialogue the term “inner speech.” Thus, Hubert Hermans proposes the concept of the dialogical structure of the “I” (Hermans, 2014). The human “I” as a dialogue of various “I – positions” is formed in dialogue with significant others. Hermans rethinks Mikhail Bakhtin's concept of polyphonic consciousness through narratives. He notes that independent “I – positions,” which represent both participants in social relations and express various parts of personal consciousness, give rise to corresponding memories or stories (narratives).

Victoria Izmagurova (2015) defines internal dialogue as “a mechanism of interaction between semantic positions of consciousness. Semantic positions are the personal and (or) emotional relationships of the subject to significant events, circumstances, people, relationships, recorded in the form of relatively stable semantic formations.” Semantic positions can be defined and designated – they are important for the development of a person’s “I,” as well as for its transformation if necessary.

A paper by Lotman (1973/2000) “On two models of communication in the cultural system” describes the concept of “autocommunication.” Lotman distinguishes between “ordinary” communication that occurs in the “I – He” system and autocommunication that occurs in the “I – I” system. The “I – I” system works on the activity of the subject. Thanks to this system, development, transformation, restructuring of the “I” itself occurs.



In addition to the fact that thinking occurs within the framework of autocommunication, the message itself acquires a new meaning, as the conditions of its transmission change.

The text generates new meanings, and therefore needs heterogeneity, foreignness, an interlocutor, and dialogue. It is worth noting that we understand “text” here in the broadest sense.

The text even rearranges its immanent structure, adapting to the audience. The reverse process is also possible – the process of a qualitative change in human consciousness. The essence of the text is the emergence of new meanings in the reader’s mind. And also a change in the reader’s consciousness in the process of working with the text. Similarly, the child develops, receiving impulses from the world around him, processing information formatted in texts.

Lotman (1992) calls both the text and the person a “semiotic system.” The semiotic system develops through the absorption and processing of new and new texts. To produce texts, you need to constantly perceive other people’s texts. A question may arise about the very first text, and Lotman answers it by drawing an analogy with chemical science. As in chemistry, it is important to distinguish between genesis factors and “catalysts” that trigger the mechanism of text development.

Can LLMs that are popular today be called semiotic systems? Without a doubt. Can we call the human – machine systems that include us, these chatbots themselves, the research questions that arise between us, semiotic systems? Probably yes. And, moreover, a semiosphere is formed that encodes and decodes itself and the texts that touch it.

Semiosphere by Lotman (1996) is a special communication space, which includes not only the sum of languages, but also the sociocultural field of their functioning. The image of the audience contained in the text is a normalizing code for it. Let’s take instructions for a certain technical device and a literary text, a novel. Both texts are just text. But they cannot be compared in several other parameters. These are different texts, fundamentally different texts. And it’s not just a matter of different conceptual languages, but also a difference in audience expectations. The expectations of the audience shape the responsibility that we resolutely place on the LLM, entrusting her with our deepest secrets as if we were a random fellow traveler on a train.

We usually assign responsibility for the text to the author of this text. But, as mentioned above, the generated text does not have an author. In any case, now the legal doctrine does not have a clear opinion regarding the legal status of generated texts and images.

Michel Foucault’s idea of the author function can be projected onto what is happening today with neural networks. The problem of the “author” function and the “death of the subject” has received various names: “death of the author”, “depersonalization theory,” “decentration” of the subject. The concept of the “death of the author” was directed against the figure of the autonomous independent and authoritarian author. Foucault says that modern writing is indifferent to the author, the individual characteristics of the writing subject are erased: “the creation, whose task was to bring immortality, has now received the right to be the murderer of the author” (Foucault, 1996, p. 14). Roland Barthes (1994) speaks about the concept of “writing” this



way: “writing is that area of uncertainty, heterogeneity and evasiveness where traces of our subjectivity are lost.”

The concept of “author” as an individualizing force has changed historically. At a speech at a meeting of the French Philosophical Society on February 22, 1969, Foucault said: “The concept of the author constitutes an important moment of individualization in the history of ideas, knowledge, literature, as well as in the history of philosophy and sciences” (Foucault, 1996, p. 12). A new episteme emerges at the end of the 18th and beginning of the 19th centuries, when the first copyright laws were adopted. If before the invention of printing the author did not recognize himself as the creator of something new, and, moreover, felt himself in a cultural tradition, then after the invention of printing the author began to feel like an autonomous independent person, and perceive his work as a continuation of his personality. Dmitry Likhachev (1971) calls the coloring of the “I” the impersonal tones of the era – “literary etiquette.” Such literary etiquette is inevitable for any author of any era (regardless of whether this “literary etiquette” calls for the renunciation of one’s authorial position or, conversely, for the affirmation of such).

The author's name “provides a classification function; such a name makes it possible to group a number of texts, differentiate them, exclude some from their number and contrast them with others” (Foucault, 1996, p. 21). And the name of the author characterizes “a certain way of being of discourse”. Foucault understands the term “discourse” (“discursive practices”) as a way of speaking, as a stylistic specificity.

It is important to note that discourses that carry the author function are objects of appropriation. The author is a projection of the processing to which texts are subjected. The “author” function is the result of a complex operation that constitutes the author. But the “author” function is not just a reconstruction, because the text contains a certain number of signs referring to the author (verb conjugations, adverbs of time and place, personal pronouns. So modern LLMs are based on similar principles.

Simultaneously with the advent of the first copyright laws and the emergence of the romantic concept of a willful and authoritarian author protecting the fruits of his creativity legally through the concept of intellectual property, a concept was emerging that affirmed the importance of the public domain and the right to free access to knowledge. The theory of the common good does not aim to deny intellectual property rights as a phenomenon but finds other ontological roots of copyright – protecting the right of society to free access to knowledge (Kartasheva, 2023).

The ability to understand or assimilate new knowledge is as important an ability as the ability to create new things. “The modern episteme, which was formed at the end of the 18th century and still serves as a positive basis for our knowledge, that episteme in which a certain special way of human existence and the possibility of its empirical knowledge took shape – all of it assumed the disappearance of Discourse and its monotonous dominance, the shift of language towards objectivity and its new manifestation in all its diversity” (Foucault, 1977). One of the meanings of discourse is understanding it as the ability to dissect mental representations.

Foucault speaks about the objectivity of language, and this observation is especially true regarding scientific language, where the existing rules of reasoning and theory construction lead the researcher’s thought along with them. It is worth noting that the idea



of intellectual property is perceived differently in the scientific community (as well as among inventors) and among writers who professionally work with words. If the latter create with the help of inspiration and the greatest advantage of a literary work is its uniqueness, originality, and dissimilarity from others, then in the scientific community it is important to rely on predecessors, make accurate references to other authors and carefully collect and process extensive factual material.

THE METHODS

Autocommunication is associated with any act of authorship since every text carries a message. This message may be a message to oneself, and the transmission of the message may cause a restructuring of one's personality structure. Within the framework of the semiotic approach to communication, Lotman (1992) distinguishes two types of speech activity. If the first is addressed to an abstract interlocutor and has the largest possible amount of memory, then the second is addressed to a specific interlocutor and has a type of memory peculiar to him. "A language for everyone" is contrasted with "a language for oneself."

When we correspond with a person, we are inside the "language for everyone". But now of correspondence with a chatbot, we find ourselves inside the "language for ourselves" system. The fact that we do not perceive LLM as equal to ourselves, not having a physical form and not working without initiation from a person – all this leads us into the field of autocommunication, dialogism in the form of "I – I."

In the case of popular chatbots, the dialogical nature of these relationships develops in time and in the space of the necessary tasks that need to be solved by one or another person. As higher education teachers, we may criticize students for turning to chatbots and neural networks when writing their dissertations, but we do not show them the important action where this type of dialogue can be useful. We could explain to students that leaving text writing to neural networks is equivalent to them voluntarily throwing out the most interesting things from their lives. But at the same time, entering dialogues with the neural network, experimenting, and partly even provoking your own thoughts with such dialogues is permissible and quite encouraged. To train your skills, to find weak points in the argument, for "references." For these purposes, chatbots with neural networks are very useful. This type of communication can be called autocommunication.

It should be noted that standard answers to standard questions cannot lead to any breakthrough solutions. A neural network can only output what was or is. And this undoubtedly determines (or will determine) perception. We can say that the text shapes the audience in its own image; "a dialogical relationship develops between the text and its audience" (Lotman, 1992). The dialogical relationship between text and audience is characterized by shared memory. What is shared memory? The audience has knowledge, memory of past events, a special language – all this is common memory, common discourse, common background. This shared memory is asserted in the neural network's responses. Even distortions and biases in responses occur due to reliance on distorted training data. Not only do we inevitably learn and develop through dialogue with chatbots,



but we also participated and are participating in the training of these neural networks themselves.

A person who is just starting to ask questions to a neural network often encounters a situation where it is very difficult for him to find the right word order to get the final result. There are guides on prompting (Best practices for prompt engineering with the OpenAI API, n.d.; Prompt Library, n. d), which help formulate the request so that the neural network can give the desired answer.

There are several techniques that help improve the output of neural networks. For example, there is such a technology as chaining, when the response from the LLM is fed to its input, but with a different prompt command in the style of “find the problems of this approach.” On the one hand, this strategy greatly slows down the achievement of results, on the other hand, it allows you to solve problems of a very high level. And there are many techniques that hack the defenses of neural networks and try to force them to give the wrong answer. In addition, there are sites that help you write prompts. For example, ChatGPT cannot yet read Google Docs even in browser mode. Or it is very inconvenient to send content that does not fit in one message (more than the proposed context window, although the latter is constantly increasing). Various services (Chatgpt – prompt – splitter, n.d.) help in solving such difficult problems, helping to better formulate the question to get the desired result. Even though composing prompts is one of the most creative parts of communicating with neural networks, it is also worth noting that a neural network can do a better job of composing prompts than a person, since it can try out more possible options.

This article does not contain even single sentence generated by a neural network. But during the work on the article, there was constant testing and verification of the problems raised in dialogue with the chatbot, which helped to better formulate the hypotheses discussed in this article. It should be noted that this kind of dialogue is especially useful at the initial stage when hypotheses are just being formulated.

Prompting as a technology of dialogue with neural networks is based not only on the art of asking questions. The image of the audience contained in the text is a normalizing code for it. And the context clarifies the field in which you need to look for answers. This is the basis of the technology of role models, which is very often used when formulating queries to a neural network. An example of this use: “Imagine that you are a social researcher working at a university. Your task is to create a series of questions that could be asked to respondents on such and such a topic. Below, describe the criteria for selecting questions and rank the questions according to these criteria in the table.” Thus, the acquired role sets the normalizing code and determines the response received. But to define a role, a person needs to understand the specifics of this role, which is impossible without complete immersion in a certain field of activity. Thus, a dialogic relationship develops between the generated text and the questioning audience, which develops all participants in communication.



EXPECTED OUTCOMES

A person formulates a question for a neural network, already assuming a certain answer. During autocommunication with LLM, a person must have an expectation of the desired answer, otherwise it is difficult to formulate a question. And this assumption and expectation is fixed in the structure of the prompt, so prompting as the art of asking questions acquires particular importance.

But people are not used to asking questions, already assuming answers to their questions. People ask differently: if they don't know something but want to find out in a conversation with a knowledgeable expert or by searching through a huge database, which is essentially the Internet as such. So, while writing prompts is quite a creative job, sometimes another neural network can do the job better, creating ensembles of models.

If search engines offer a variety of options, then the neural network's answers are verified, presumably error – free... and have no alternative. The peremptory nature of the neural network's responses is corrected by chaining technology, when we further clarify the criteria for the received answer and ask the neural network to evaluate the answer given to us.

The search mechanism is changing. The idea of searching and the idea of questioning is undergoing changes. And, therefore, the idea of dialogue. The subject of this article is not to restore the significance of the technology of Socratic dialogue but mention of this type of dialogue will be an important addition to what is said above.

Thanks to autocommunication, development, transformation, restructuring of a person's "I" occurs, which is an enriching practice. It is important to pose and consistently solve the problem of developing human thinking abilities.

The semiosphere, made up of human – machine systems including humans, chatbots, research questions, and more, encodes texts in a specific way. Humanity looks into the mirror of neural networks and sees reflection of humanity there.

Dialogue with larger linguistic patterns, as a discursive practice that allows people to formulate their own thoughts, can be used as a practice for formulating outlines or references for further work. The practice of autocommunication can be very effective for both social research and general audiences. But while people look at dialogue with neural networks as a standard dialogue in which there is a question and there is a correct answer, then these opportunities are not used.

REFERENCES

- Barth, R. (1994). Death of the Author. In *Selected works: Semiotics. Poetics* (pp. 384-392). Progress.
- Foucault, M. (1977). *The Order of Things: An Archaeology of the Human Sciences*. Progress.
- Foucault, M. (1996). What is an Author. In M. Foucault, *The Will to Truth: Beyond Knowledge, Power and Sexuality* (pp. 7-47). *Works from different years*. Kastal.
- Best practices for prompt engineering with the OpenAI API (n. d.) URL: <https://help.openai.com/en/articles/6654000-best-practices-for-prompt-engineering-with-the-openai-api>



- Hermans, H. J. M. (2014). Self as a Society of I-Positions: A Dialogical Approach to Counseling. *Journal of Humanistic Counseling*, 53, 134–159. <https://doi.org/10.1002/j.2161-1939.2014.00054.x>
- Izmagurova, V. L. (2015). The Phenomenon of Autocommunication as a Factor in the Development of Consciousness [The Phenomenon of Autocommunication as a Factor in the Development of Consciousness]. *Communicology*, 4, 31-42.
- Jacobson, R.O. (1985). Speech Communication; Language in Relation to other Communication Systems. In *Selected works* (pp. 306 – 330). Progress.
- Chatgpt – prompt – splitter (n.d.) URL: <https://chatgpt-prompt-splitter.jjdiaz.dev>
- Kartasheva, A. A. (2023). *Philosophical Foundations of Copyright*. Business book.
- Likhachev, D. S. (1971). *Poetics of Old Russian Literature*. Hood.
- Lotman, Y. M. (1996). *Inside Thinking Worlds*. Yazyki russkoj kul'tury.
- Lotman, Y. M. (1992). *Selected articles in three volumes, vol. 1*. Alexandra.
- Lotman, Y. M. (2000). Autocommunication: “I” and “Another” as addressees (About two models of communication in system of culture). In *Semiosfera* (p. 159 – 165). Iskustvo—SPB. (Original work published 1973)
- Nazarchuk, A.V. (2009). *Theory of communication in modern philosophy*. Progress – Tradition.
- Prompt Library (n. d.). *Anthropic*. <https://docs.anthropic.com/claude/prompt-library>
- Sapir, E. (1993). Communication. In *Selected Works in Linguistics and Cultural studies* (pp. 210 – 215). Progress.
- Saussure, F. de. (1977). *Works on Linguistics*. Progress.

СВЕДЕНИЯ ОБ АВТОРЕ / THE AUTHOR

Анна Александровна Карташева,
akartasheva@hse.ru, ORCID 0000-0001-5643-6071

Anna Kartasheva,
akartasheva@hse.ru, ORCID 0000-0001-5643-6071

Статья поступила 14 февраля 2024
одобрена после рецензирования 18 мая 2024
принята к публикации 10 июня 2024

Received: 14 February 2024
Revised: 18 May 2024
Accepted: 10 June 2024



<https://doi.org/10.48417/technolang.2024.02.06>

Research article

Lebenswelt, Digital Phenomenology, and the Modification of Human Intelligence

Alexander Vnitskikh (✉)  and Sergey Komarov 

Perm State University, st. Bukireva, 15, 614068, Perm, Russia

avnut@inbox.ru

philos.perm@gmail.com

Abstract

The development of contemporary digital technologies leads to a profound modification of human intelligence. The authors assume that this modification should be studied by means of a special kind of phenomenology. It is digital phenomenology which examines the structures of consciousness of the modern technogenic subject. This builds on their previous works where the authors have already discussed a theory of the transformation of human intelligence driven by digital technologies. The influence of these technologies results in virtualization of affect. Affect becomes detached from its local manifestation in the human body and is manifested in material and energetic processes in digital infrastructure. As a result, space and time, categories of reason, and productive imagination become aspects of mobile devices and digital infrastructure. The aim of this contribution is to discuss the possibilities of digital phenomenology in the study of communication of the technogenic subject. Methodologically, the study refers to the phenomenological approach. Archetypes are compared of classical intelligence and technogenic subjectivity which defines the content of communication. The authors suggest that consciousness as a pure orientation can undergo digital modification, as the world of primordial objects is discovered through corporeal experience. A modern human body is not constituted within the boundaries of direct sensual experience but perceives digital devices as body organs. The peculiarities of the language of these devices determine human linguistic practices as well. So we can see non-human intelligence and non-human communication. Both intelligence and communication are becoming increasingly artificial. The prospect of further in-depth research in the digital humanities is outlined.

Keywords: Lebenswelt, Digital technologies, Digital modification of human intelligence and communication

Citation: Vnitskikh, A., & Komarov, S. (2024). Lebenswelt, Digital Phenomenology and Modification of Human Intelligence. *Technology and Language*, 5(2), 67-79.
<https://doi.org/10.48417/technolang.2024.02.07>



This work is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/)



УДК 1.62-529

<https://doi.org/10.48417/technolang.2024.02.06>

Научная статья

Lebenswelt, цифровая феноменология и модификации человеческого интеллекта

Александр Внутских (✉)  и Сергей Комаров 

Пермский государственный национальный исследовательский университет, ул. Букирева, 15,
614068, Пермь, Россия

avnut@inbox.ru

philos.perm@gmail.com

Аннотация

Развитие современных цифровых технологий приводит к глубочайшей модификации человеческого интеллекта. Авторы исходят из того, что эта модификация должна изучаться в рамках особой разновидности феноменологии – цифровой феноменологии, исследующей структуры сознания современного “техногенного субъекта”. В свое время авторы уже обсуждали концепцию трансформации человеческого интеллекта под влиянием цифровых технологий. Было показано, что в результате этого влияния происходит виртуализация аффекта: последний отрывается от его локального проявления в теле человека и начинает существовать в форме материальных и энергетических процессов в цифровой инфраструктуре. В результате, пространство и время, категории рассудка, продуктивное воображение начинают выступать во все большей степени как функции не человека, а мобильных устройств и цифровой инфраструктуры. Целью данной статьи является обсуждение возможностей цифровой феноменологии при исследовании коммуникации техногенного субъекта. Методологическим основанием исследования является феноменологический подход, благодаря которому сравниваются архетипы классического интеллекта и техногенной субъективности, которые оказывают определяющее влияние на содержание коммуникации. Показано, что сознание как чистая направленность может претерпевать цифровую модификацию, поскольку мир примордиальных объектов открывается посредством телесного опыта. Конституирование же человеческого тела современного человека осуществляется не в границах непосредственно чувственного опыта, а включает в себя продолжение органов тела в цифровых устройствах. Особенности же “языка” этих устройств определяют человеческие языковые практики. Таким образом, мы фиксируем уже не вполне человеческий интеллект и не вполне человеческую коммуникацию – и то и другое существенно “артифицируется”. Намечена перспектива углубленных исследований в области цифровых гуманитарных наук.

Ключевые слова: Lebenswelt; Цифровые технологии; Цифровая модификация интеллекта и коммуникации

Для цитирования: Vnutskikh, A., Komarov, S. Lebenswelt, digital phenomenology and modification of human intelligence // Technology and Language, 5(2), 67-79.
<https://doi.org/10.48417/technolang.2024.02.07>



This work is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/)



INTRODUCTION

Phenomenology is known to have two key methods at its disposal – intuition and phenomenological reduction. Intuition captures self-evidences where self-evidences are the key criteria for the existence of being in the classical philosophical tradition. The fact that a human being has intelligence and participates in communication is seen to be an example of these evidences. Here, however, the question presents itself: how far do intelligence and communication truly belong to a human being these days? The hypothesis of the research is that digital transformation which gave large language models and talking gadgets has significantly modified the intelligence of human beings per se. This human being is a technogenic subject that is utterly different from the archetype of the subject from the epoch of classical philosophizing. People communicate their thoughts. But do they always think as human beings do? We believe this question belongs to the field of so-called digital phenomenology aimed at examining the structure of a technogenic, specifically digital subject. Don Ihde wrote that phenomenology should be classified as the latest technology belonging to the *Lebenswelt* of the XXI-century person (Ihde, 2009). To do this, one should pay attention to the trends in the changes of its lifeworld and its social connections, including communicative ties.

THE CORE OF THE PRIMORDIAL WORLD

Let us start with theoretical grounds to the answer to the question: is digital modification of human intelligence possible at all and how deeply pervasive could it be? To answer the question, we refer to Edmund Husserl's (1973) methodology of primordial reduction. Husserl's phenomenological reduction of "existing givenness" is followed by the "primordial reduction" with its goal to achieve the world of my own experience where my own presentations matter. The second stage of Husserl's reduction results in reduced consciousness and its correlated primordial world with no indication of the Other. The primary structure in the basis of the primordial world is the transcendental subjectivity as an immanent structure of consciousness with the transcendent being the immanent transcendental. Intentionality determines the primary layer of consciousness to be the transcendental ego with its inner intentional object – World as the Others. That is why we believe that it is impossible to directly modify the core of the consciousness' primordial world under the impact of digital technologies. However, since the world of primordial objects becomes open through physical experience, and an animate body becomes a flow of sensations which the world institutionalizes in (Husserl, 1973), this mediated modification turns out to be possible. What is more, it could be very deep, penetrating the entire experience of a subject's sensitivity, because digital technologies could change the entire physical experience of a human being.

TRANSFORMATION OF THE KINESTHETIC WORLD

Kinesthesia or flows of sensations associated with an animate body are the primary experience of orientation in the world (Merleau-Ponty et al, 2013). However, what happens when the children representing the generation of digital natives face the digital



technologies as their first objects constituted in their experience? In this case, the primary kinesthesia of sensitivity, materiality, space, and objectivity will be updated with one more kinesthesia when a human body is constituted not within its sensational experience, but rather perceives digital devices as its organs. These days, a child is typically provided with an access to the gadgets at an early age, and right from their early age their sensitivity horizon shifts *beyond* the borders of natural sensitivity to the sensitivity horizon of gadget sensors. Here the following analogy could be drawn: it is known that the point at which a vehicle driver (even with low driving proficiency) focuses his attention and which he subjectively associates with himself is located not inside his body but rather at some distance in front of the vehicle – and this distance increases with speed. Digital natives, on the other hand, dive into their sensitivity horizon with no sufficient experience of interacting with the objects of the material world via their natural sensitivity, that is why this subject dives into the horizon even deeper than a vehicle driver. Actually, a child with the parents tightly ‘merged’ with the gadgets at work, as well as at home, during their leisure time, has no other content choice.

As a result, one could observe a significant modification of space and time orientation of modern persons. The authors’ previous work (Vnutskikh and Komarov, 2023) has already mentioned that a cellphone’s space is manifested as a potential physical presence of a human body. Since geolocation is a basic feature of any cell device, space is synthesized a priori outside human consciousness. Space is given not as a type of external contemplation but rather as a type of representation of digital devices; it is made available as a function of this device rather than is synthesized with imagination. The actual geometry of space appears to be distorted by virtual topology, while the global space in reality acquires the structure of rigid cells of local existence. Social space is experienced as a set of ‘local’ places defining the possibility for social movements. This is manifested in greater disorientation of a person in space overall, a loss of connection with routine life, topological cretinism, and a loss of fundamental understanding of the global space beyond a mobile interface.

Like space, time also turns out to be a cumulative flow of all external rational processes and is grasped by sensations. That is why it appears to be a parameter of consciousness rather than self-consciousness of a modern person. This time is not constituted by our *I*, actions and activities of a person. On the contrary, it is defined by the flow of external events, time of the planners, clocks, and quantum generators of mobile devices. This is not human, not *my* time, but the time of devices.

Consciousness cannot catch up with reality, and thus becomes reproductive: the present turns out to be reconstructed from the archive of the moments of time. It means that memory is losing its function of storing things and identifying key moments of the past correlated with the present. A digital archive is the place to accumulate and store information, its timespan is not clear. This archive ontologically manifests hybrid reality on-demand once it is made: digitalized reality is on demand in this or that mode of time and depends on search samples. In other words, time and its modes are determined by the mechanism of filtration and data extraction.

This means human rationality is affectively inhibited so that time is experienced as the time started by external clocks, eternally reproduced and regularly multiplying the



same sensation. No matter how reactively time is represented, it is connected with understanding the difference between distribution and reproduction, representation and memory, the present and reconstruction, reality and archive. This is the contradiction between the time of *life* and the time of *clocks*, because human consciousness cannot be completely turned into a mechanical process. As a result, all sensations, evaluations and comments about the present as a whole lose their common basis and multiply. What is embodied in the digital devices is mirrored in social reality in the nature of discourse about time: the integral present uniting the past and the future is disintegrated into separate elements, and a modern person experiences these elements and speaks about them as isolated parts. For example, the genuine present can be experienced as the past or vice versa – the present can replace the past. Perhaps, it is a “digital anthropologist” who can understand these processes which change the humanway of being (Horst and Miller, 2013).

Here, Mark Coeckelbergh reasonably says that digital technologies could be used and understood provided there are a number of conditions for possibility or transcendental structures (Coeckelbergh, 2022a). He defines *language, social relations, a human body, material infrastructure* among these grammars or conditions which enable the meaning and use of digital technologies but which at the same time limit them. Indeed, for example, how we experience and think about something is embodied – cognition is an active attitude towards the world as it is manifested in a human body. Digital agents of communications or AI see that they are structurally perceived via our manifested mode of being and cognition. As a result, we can, for example, project a human body on an artificial agent of communication (Coeckelbergh, 2022a). However, in light of the above-mentioned transformation of the technogenic subject in relation to space and time, this projection turns out to be at least not quite a human one but rather mainly an artificial projection imposed on a subject by a machine.

Here, one could observe a clash of phenomenological and analytical traditions because the use of language defines new grammars and narratives – not only to describe the sensations per se (kinesthesia) and the respective things of experience (world) but also to describe the *I* (Descombes, 2011). In other words, the words not only describe things. In a sense they *do* things and make others do them; thus, words and things are a part of practices or, what Ludwig Wittgenstein called, the game.

Similarly, descriptions of emotions in words turn out to be a constructive practice of experiencing these emotions, just like in a cartoon ‘puzzle.’ It is worth reminding us here that kinesthesia is the flow of sensations associated with the body, as well as a well-known fact of a mimic expression of emotions. The pattern of bodily manifestation of emotions could definitely be expressed with machine languages, which could be proven by way of successful simulation of emotional mimics in a hyperrealistic robot which appears to be quite skillful in, for example, predicting a human smile or initiating a smile – or initiating some other things (Robo-C2, Promobot, 2019). So, emotions are associated with kinesthesia, while they, in turn, could be technologically mediated, have their images in gadgets and their names as their horizons. In this situation, a human body is a transcendental structure of meaning, and we are bodies-intelligences interacting with the environment (Coeckelbergh, 2022a). The concern is that our monopoly on emotionally



defined understanding of the world is getting lost – digital infrastructure “conceptualizes” us with digital technologies. All kinesthesia and emotions could be comprehended without face-to-face human communication but solely with the grammars defined by artificial machine languages or AI.

WORLD AS THE OTHERS. CONSTITUTIONALIZATION OF THE SOCIAL WORLD

What is a social world? In terms of phenomenology and phenomenological sociology, the social world is knowledge objectification in human practice. Alfred Schutz (1962) writes,

«I am a human being born into the social world, and living my daily life in it, perceive it as it is, constructed before my time, opened for my interpretation and actions correlated with me, a relevant biographically determined situation. A particular type of connections acquires its specific meaning in relation to me. I designate this meaning with a word “we.” The Others are “you” in relation to “us” with me at the center. A third party “they” is defined in relation to “you” that is correlated with me» (Schutz, 1962, p. 15).

At the same time, Schutz says that communication as a realized inner speech can rarely be seen in people: people automatically grasp situations and actions and rely on socially borrowed/socially approved typifications transferred via language as typifying agents.

However, the problem is that this typifying agent is referred to not only by people. Robots have long become the objects of people’s talk and *started talking not worse than many people in terms of their forms*: we can just refer to the latest versions of ChatGPT. It turns out that both people and non-people become the users of natural language, while the devices and machine can also do something with the words, typify the elements of the social world in a specific manner and make others (both machines and people) do something. This involves the integration of machines into social material practices (Coeckelbergh, 2020). We are so deeply immersed into interaction with the elements of digital infrastructure that only a “disconnection” can make its “logic” visible to us (Kaun, 2021).

It is clear that technologies do not think, human beings think. Technologies, on the other hand, speak in terms of a language game, while we learn this language, enter into a dialogue with it and then change it. Language is a transcendental structure of technology. In this case, features of this language start to define our (human) language practices, which results in specific features of both self-understanding and understanding the World as the Others. These Others act as addressants-communicators-counterparts. Does this experience have the criteria for clear distinction between human and non-human subjects of communication, for example, robots? In terms of intersubjectivity, these others are organized in our communication as *quasisocial subjects*, although Coeckelbergh (2022b) reasonably asks to be careful when referring to this notion.

What does AI change in the intersubjective world?



First of all, it changes the perception and understanding of the others: gadgets in their anthropologic dimensions are no different from human beings. That is why Coeckelbergh suggests something akin to „word-building” of technologies.

«I propose to call technology a formator. It is not just object and substance, not just a thing. It forms worlds... humans... also form. There is... co-formation» (Coeckelbergh, 2022a, p. 153).

Secondly, this leads to adopting the behavior practices which traditionally lack any reasons to distinguish robots from people, as well as to cheating. The way we speak about others and ourselves is particularly important. Language initially structures and defines the way we communicate, including with the machines. Wittgenstein showed that language is perceived by an individual as a natural language game learnt by practice. But what should be done if this “natural” language is actually artificial? Any language, including AI language, has its own metaphysics and is not neutral to the social reality perceived by an individual. That is why it is not only that robots are social subjects, but inversively we can realize ourselves as “natural” robots (thus, AI understanding becomes an explanatory model for natural intelligence).

The trend is *to understand the natural by analogy to the artificial*. The latest significant change in understanding the correlation between the natural and the artificial is mainly connected with the changes in digital technologies. It could be assumed that this could lead to a kind of turning over: the human natural is turned into the artificial in terms of human communication, consciousness, human subjectivity on the whole; and, vice versa, digital reality is built as the natural with its own objective laws.

Let us explain. Classical science from Plato and Aristotle defined the natural to be everything with the reason for its existence in itself, from nature (πρώτα από τί φύση), while everything with the reason in the other was seen to be artificial (το πρώτο στην τέχνη). The latter was typically understood as something created by a human being; even automata, at first sight, acted as the natural, which surprised Aristotle among others (Aristotle, 350 B.C.E./1965, 734b 7-19). However, at modern times the sciences assume that “the first by nature” could and should be artificially reproduced. This is what science and engineering of modern times deal with: they artificially create the conditions to reproduce the natural phenomenon which is now controlled, rationalized and is becoming more technological.

At the same time, the artificial colonizes what has been considered to be the natural up till now – the human body and consciousness (soul). The philosophy of modern times is dominated by the notions of correction and purification – to perceive the natural phenomena – of the human mind. This could be evidenced with the titles of the treatises written by the creators of the modern times’ philosophy and science – *On the Improvement of the Understanding* or *Rules for the Direction of the Mind* etc. Consciousness, sense, mind given to us by nature are not quite perfect, they are limited, blinded by various fictions, ambiguous and fake ones, which penetrate human consciousness (Spinoza, 1901). On the other hand, thinking itself for some time has been considered to be a *human* ability that is why it preserves the characteristics of the natural, an ability given to a person by nature.



However, the development of computing technologies in XX – XXI centuries brought about the so-called artificialization of human thinking as their natural ability. Now, the natural is artificially reproduced. For this it does not matter that analogue or digital operations just simulate the natural brain processes. Computation rate and amount determine and define understanding of thinking as the artificial process being more perfect than the natural human ability. This development of neural networks ultimately transforms the meaning of the natural and the artificial: human thinking is not just rationalized – what the philosophers of the modern age dreamt about; it is also not just technicalized and technologized – what the scientists of the 20th century assumed; *it is becoming the artificial in its strict sense*, since all its key elements can now function under the impact of external digital technological infrastructure (Vnutskikh and Komarov, 2023).

CONSTITUTIONALIZATION OF THE PHYSICAL WORLD

This question might be seen to be paradoxical, but it is still worth asking: do physical things exist in this world? Postphenomenology claims that the things are the agents between us and the world, although when the things turn into the users of language, they change their phenomenal *status*.

First of all, technologies constitute reality by creating material infrastructure and other material prerequisites. Phenomenologically, it means that the things per se are constituted in terms of the role they play in the technological sphere of our experience. It has been mentioned that digital technologies change our perception of time and space, imagination is deformed when reality is perceived, our ability to think (sense) is being transformed when it demands a colossal digital infrastructure for it to work properly (Vnutskikh & Komarov, 2023). For example, the internet is a special type of space which could set the meanings to a person, thus Homo virtualis or a personality in the internet is no longer one of the human identities. Ontologically it is defined within a self-regulating virtual reality, while the attempts to limit it are perceived by the users as artificial (Bylieva, 2016). From this perspective, virtual beings in the social networks replace their living, sensual corporeal being for a person. The virtual world becomes more real to a person than vague physical being. We rely on our virtual experience which, instead of personal experience, starts defining the non-virtual living being of a person.

Secondly, language defines the perception and understanding of the reality: it is known that grammar and syntax give language some metaphysics (Carnap, 1950). But all programming languages manifest their metaphysics in a different manner, and that is why they can perceive the outer world in a different manner (Bogost, 2012). The question whether artificial intelligence and all information reality affect the developing living intelligence and its perception of the physical world is quite ambiguous. If a child constitutes the physical world via the gadgets rather than their practical contact with physical things, then the things are not physical objects but rather animated ones with their role of talking assistants, virtual objects, etc. This means that things do not act on their own. It is next to impossible to see a thing as it is, per se, as its material entity, outside its instrumental, informational and other characteristics, in its pure materiality,



sensuality and form. Therefore “pure ontology,” as Aristotle or Heidegger understands it, is incompatible with “digital being” (Koulouris, 2020).

Thirdly, it has been mentioned that technologies talk. Daria Bylieva reasonably says that “language used to be seen as a purely human technology, but now language is acquired by non-people. Chat-bots, voice assistants, embodied dialogue agents and robots have acquired the ability to communicate via language and can often present themselves as humanoid personalities. People perceive them in an ambivalent manner; they recognize them to be the Others. Thus, artificial intelligence exploits language in a way which is not determined by the human method of using it” (Bylieva, 2022, p. 111). The phenomenological status of material things is changing: their physical givenness is replaced by their virtual image.

Fourthly, how does the real world constitute itself in this case, how is it described by all natural sciences? Husserl believed this layer of the lifeworld is based on the intersubjective experience of science (Husserl, 1973). The unified picture of the world described by the natural sciences is rooted in the intersubjective invariance of primary experience in perceiving the physical world and invariant language structures of its description. Modern conditions turn science into technological building of these objects rather than into research and acts as technoscience. Then, what is the real physical world in this context? The sciences become a complicated system of knowledge represented and mediated by intellectual networks. This knowledge makes things lose their material nature and thingness as they are mediated by language structures, knowledge arrays and artificial intelligence rather than human experience about them. Their scientific images are defined not by the primary experience of perceiving things, but, on the contrary, it is the knowledge images of things that substantiate the experience of their perception. In this context, things themselves appear to be *sets of data* as their intellectual images suggest, they are reduced to data and act as data: technologies act “as the author of language, as a user of language, and as a creator of the world” (Coeckelbergh, 2020, p 22). An objective world is presented as an information network image or an infinitely expanding database as a virtual copy, *a virtual construction of a digital language*.

THE SPIRITUAL WORLD

The world of values and ideals constitutes a special layer in a person’s *Lebenswelt*; this is the subject of humanities. The human spiritual world is being transformed like the physical world which becomes manifested in the databases of the natural knowledge. First of all, some values are being modified in that they are being replaced by the implicit principles of metaphysics of programming language (as Rudolf Carnap understands it) or by the metaphysics of natural language which has already been reformatted under the impact of digital technologies (as shown above). The function of reality reproduction and its never-ending reconstruction prevails in intellectual networks due to data digitalization and its reverse decoding in databases (digital archives). This is highly likely to modify fundamental values – since the perception of time modes and existence modalities are being modified. Here, for example, we can refer to the fact that the digital era is clearly focused on a structured future rather than the actual accomplished past, and that is why



history can be revised while objective reasons for ongoing events are ignored. Modality of possibility, in its turn, has its priority over modality of necessity (Vnutsikh and Komarov, 2023). We believe it means that values are being relativized as the „must“ or „has-to“ turns out to be just a status for the virtual and could be revised, if necessary. Therefore, we have to talk not just about the transition of traditional values into a digital format. We have to talk about a fundamentally new dimension of values and meanings (Bolshakova, 2024). Ideals of meaning can also act as the settings for the efficient exploitation of software.

Secondly, the way of thinking is changed under the impact of digital technologies so that the ability to put meanings in the strings of representations – sense – is also being deeply transformed. Today, we can talk about objectivation of sense into a colossal body of modern digital infrastructure. Technogenic factors develop a rigid but quite rational structure towards human subjectivity. Located outside the individual, transcendental structures of sense break through to the surface of the affective life of consciousness as a set of rational ties with reasons and motives in their technogenic nature. For persons this signifies the degradation of their logical reasoning; this degradation leads to a magical attitude towards machine reasoning (Bylieva and Zamorev, 2022). Poor logical reasoning, on the one hand, results in a loss of value of well-reasoned knowledge, while, on the other hand, weakens our ability to extract the meanings from fairy tales.

In this sense, ChatGPT can be looked at as a tool which simulates and even replaces dying-out human logical reasoning. The question-answer structures of Socratic dialogues, as well as the dialogues of the founders of modern science used to be valid tools for exploring objective reality but are no longer relevant under modern conditions. There are several reasons for this. First of all, the elements of this reality are turned into sets of data and constituted as regards to their role in technological patterns of our experience. Secondly, the reasoning skills of one party, namely of the person, in a modern digital dialogue leaves much to be desired as we go forward. For example, one commercial says that if you refer to the Alexa Voice Assistant for help to prepare your speech, this will be equal to having over 1,000 virtual assistants that could help you with your preparation. This raises the questions: where is the place for a person, and who is the owner of the prepared speech?

CONCLUSION

This study is preliminary by its nature, and it is worth noting here that the authors do not suffer from technophobia. The undeniable fact is that technologies accompany human beings from the moment of birth and define social progress. However, after Martin Heidegger's *The Question Concerning Technology* and subsequently the digital transformation of the *Lebenswelt*, technologies should not be perceived as a simple tools to satisfy human needs. One should be aware that the digital transformation, just like any other human-made thing (Mould, 2018), is ambivalent in its possible application, complicated and multi-faceted, and not all its aspects can be thoroughly examined in the context of an economically determined accelerated exploitation of digital technologies.



The results of these studies could provide a lot for *socially responsible* (in the full sense of the word) developments of modern technological systems.

Nowadays, the authors of humanities use digital methods and very often explore the interaction between human beings and digital technologies. These authors refer to these approaches as “digital humanities.” However, the transformation of the *Lebenswelt* in the contemporary context opens up prospects for deeper research, because the human being is no longer identical to the pre-digital human being. The use of digital tools is secondary, because it is “superimposed” on the already transformed human being, on the already transformed social.

The authors think the transformation of the *Lebenswelt* in modern conditions calls for important further study. First of all, digital technologies transform the corporeal experience of a person and one’s orientation in space. Here, *digital phenomenology merges with digital anthropology*.

Secondly, the transformation of *Lebenswelt* is manifested in the phenomenon when things lose their physical givenness under the impact of digital technologies: this givenness is replaced by phenomenological virtual being. Things speak the language of intellectual networks which are not the agents between a person and reality but rather reality itself. The objective world is only a virtual construction of a digital language (code). This aspect of digital transformation reveals the need for defining and developing *digital ontology*.

Thirdly, digital communication produces a person as an actant without *ego*, with emotions turned inside out, and subjectivity subordinated to digital infrastructure – represented and self-represented with language modified by the linguistic practices of digital devices. This constitutes the field of *digital psychology*.

Fourthly, digital technologies penetrate both the self-perception of a person and the perception of other people as being different from me. In modern conditions, the social world is being constituted through intellectual networks and the language of their representation. The language of intellectual networks appears to be an instrument for constituting social reality as a quasisocial being. It calls for the development of a specific *digital sociology*.

Fifthly, values are relativized due to their isolation from the actual physical world, the loss of value of objective knowledge and its reasoning becoming replaced by databases and quite challenging fact checking. As a result, the spiritual world as a layer of the lifeworld becomes a set of relative and replaceable values, while the ideals are constituted only as possible references for human behavior. This is dealt with in *digital axiology*.

It is worth emphasizing once again that our phenomenological analysis is preliminary and diagnostic. However, this examination opens up the new areas of study mentioned above. It also opens up new content for the constitution of *Lebenswelt* for a modern person – digital anthropology, digital psychology, digital sociology, ontology of digital being, and axiology of digital world. Each area is based on the physical experience which is associated with the introduction of digital technologies, digital code, and language, artificial intelligence in the respective area of a person’s lifeworld.



Thus, phenomenology reveals its new nature: it is not only a heuristic method for the analysis of a person's *Lebenswelt* in the digital era. First of all, it discovers the phenomena of human-AI interaction which cannot be identified through the analysis of AI or in just human-AI interaction. The digital codes and language structures become the phenomena of the lifeworld rather than the elements of software. Secondly, an analysis of the constitution of the digital lifeworld captures a genetic, although not always positive role of digital technologies, digital code, and language in accumulating the experience of a modern person. Digital codes appear to be the packages of meanings, while the gadgets and digital agents are only the machines of genesis of person's lifeworld. Databases, digital agents, electronic gadgets, neuronets, and AI are not pure physical material objects and derivatives only of engineering activity. We see digital phenomenology to be a promising area for the analysis of AI, digital agents, and digital codes as social phenomena.

REFERENCES

- Aristotle. (1965). *De generatione animalium* [Parts of Animals]. In *Aristotelis de generatione animalium*. Clarendon Press. (Original work published 350 B.C.E.)
- Bogost, I. (2012). *Alien Phenomenology, or, what it's like to be a Thing*. University of Minnesota Press.
- Bolshakova, A. S. (2024). Pareidolia of Meanings of Digital Axiology. In *Society: Philosophy. History. Culture* (pp. 99-106). Hors.
- Bylieva, D. (2016). The Internet as a New Type of Space. *Terra Linguistica*, 2(244), 124-130.
- Bylieva, D. & Zamorev, A. (2022). Father Christmas: Magic and Technology. *Technology and Language*, 3(4), 76-89. <https://doi.org/10.48417/technolang.2022.04.06>
- Bylieva, D. (2022). Language of AI. *Technology and Language*, 3(1), 111-126. <https://doi.org/10.48417/technolang.2022.01.11>
- Carnap, R. (1950). Empiricism, semantics, and ontology. *Revue internationale de philosophie*, 4, 20-40.
- Coeckelbergh, M. (2020). When Machines Talk: A Brief Analysis of Some Relations between Technology and Language. *Technology and Language*, 1(1), 22-27. <https://doi.org/10.48417/technolang.2020.01.05>
- Coeckelbergh, M. (2022a). The Grammars of AI: Towards a Structuralist and Transcendental Hermeneutics of Digital Technologies. *Technology and Language*, 3(2), 148-161. <https://doi.org/10.48417/technolang.2022.02.09>
- Coeckelbergh, M. (2022b). Response: Language and robots. *Technology and Language*, 3(1), 147-154. <https://doi.org/10.48417/technolang.2022.01.14>
- Descombes, V. (2018). *Le complément de sujet* [The Subject Complement]. Gallimard.
- Horst, H. A., & Miller, D. (2013). *Digital Anthropology*. Bloomsbury.
- Husserl, E. (1973). *Cartesianische Meditationen und Pariser Vorträge* [Cartesian Meditations and Paris Lectures]. Martinus Nijhoff.



- Ihde, D. (2009). *Postphenomenology and Technoscience: The Peking university lectures*. Suny Press.
- Kaun, A. (2021). Ways of Seeing Digital Disconnection: A Negative Sociology of Digital Culture. *Convergence*, 27(6), 1571-1583. <https://doi.org/10.1177/13548565211045535>
- Koulouris, T. (2020). Prolegomena to the Study of Digital Ontology. In *MeCCSA University of Brighton*. <https://research.brighton.ac.uk/en/publications/prolegomena-to-the-study-of-digital-ontology>
- Merleau-Ponty, M., Landes, D., Carman, T., & Lefort, C. (2013). *Phenomenology of Perception*. Routledge.
- Mould, R.F. (2018). Industrial Applications. *A Century of X-Rays and Radioactivity in Medicine* (pp. 103-107). CRC Press
- Robo-C2, Promobot. (2019). <https://promo-bot.ru/production/robo-c/>
- Spinoza, B. (1901). *Improvement of the Understanding: Ethics and Correspondence*. MW Dunne.
- Schutz, A. (1962). Common-Sense and Scientific Interpretation of Human Action. In M. Natanson (Ed.), *Collected Papers I. Phaenomenologica*, vol 11. (pp. 3-47). Springer. https://doi.org/10.1007/978-94-010-2851-6_1
- Vnutskikh, A., & Komarov, S. (2023). Technogenic Subjectivity and Its Social Projections. In E. Isaeva, & Á. Rocha (Eds.), *Science and Global Challenges of the 21st Century – Innovations and Technologies in Interdisciplinary Applications. Perm Forum 2022. Lecture Notes in Networks and Systems* (vol 622, pp 653-668). Springer. https://doi.org/10.1007/978-3-031-28086-3_61

СВЕДЕНИЯ ОБ АВТОРАХ / THE AUTHORS

Комаров Сергей Владимирович
philos.perm@gmail.com
ORCID 0000-0001-7358-6151

Sergey Komarov
philos.perm@gmail.com
ORCID 0000-0001-7358-6151

Внутских Александр Юрьевич
avnut@inbox.ru
ORCID 0000-0003-4162-1033

Alexander Vnutskikh
avnut@inbox.ru
ORCID 0000-0003-4162-1033

Статья поступила 1 марта 2024
одобрена после рецензирования 1 июня 2024
принята к публикации 15 июня 2024




Received: 1 March 2024
Revised: 1 June 2024
Accepted: 15 June 2024



<https://doi.org/10.48417/technolang.2024.02.07>

Research article

GPT Assistants and the Challenge of Personological Functionalism

Andrey Yurievich Alekseev¹   and Ekaterina Alekseeva Alekseeva² 

¹Peoples Friendship University of Russia named after Patrice Lumumba. st. Ordzhonikidze, 3, 115419, Moscow, Russian Federation

aa65@list.ru

²State Academic University of Humanities. Maronovsky Lane, 26, 119049, Moscow, Russian Federation,

eaalekseeva@gaugn.ru

Abstract

The paper discusses whether it is correct to speak of “generative artificial intelligence” – a concept that is not within the scope of AI research. The discussion suggests that it is premature to claim that humans are being replaced by GPT assistants such as ChatGPT in the field of sociocultural digital communication. Personological functionalism, which would justify the replacement of people by machines, is based on the psychofunctionalism of Ned Block, who proves the need to psychologize machine functionalism by introducing “meaning” as a criterion for passing the original Turing test. For personological functionalism, in addition to “meaning” the minimum necessary requirements of the Turing test include “creativity.” The paper shows that GPT Assistants do not pass this creativity test. To demonstrate the inability to pass a Turing test for meaningfulness, the Block machine was modified in a pair of 1978 and 1981 papers by combining the neurocomputer with symbolic versions. For the now further expanded Block test, the argumentation of previous versions is preserved and strengthened, leading to the conclusion that machines like GPT Assistants are not capable of fulfilling either the roles of psychological functionalism or personological functionalism.

Keywords: Generative AI; Complex Turing test; Block test; Lovelace test; Chinese nation; Block's machine; Psychofunctionalism; Personological functionalism

Aknowlegment The paper was prepared within the framework of the state task of the State Academic University of Humanities “Digitalization and the formation of a modern information society: cognitive, economic, political and legal aspects.” Registration number R&D 123022000042-0. Topic code FZNF-2023-0004. Registration number of topic 1022040800826-5-5.2.1;6.3.1;5.9.1

Citation: Alekseev, A. Y., & Alekseeva, E. A. (2024). GPT Assistants and the Challenge of Personological Functionalism. *Technology and Language*, 5(2), 80-99. <https://doi.org/10.48417/technolang.2024.02.07>



This work is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/)



УДК 1: 004.8

<https://doi.org/10.48417/technolang.2024.02.07>

Научная статья

Персонологический функционализм GPT-ассистентов

Андрей Юрьевич Алексеев¹  () и Екатерина Алексеевна Алексеева² 

¹Российский университет дружбы народов имени Патриса Лумумбы. Российская Федерация, ул. Орджоникидзе, д. 3, 115419, Москва, Российская Федерация

aa65@list.ru

²Государственный академический университет гуманитарных наук. Мароновский переулок, 26, 119049, Москва, Российская Федерация

eaalekseeva@gaugn.ru

Аннотация

В статье рассматривается вопрос о терминологической корректности понятия “генеративный искусственный интеллект”. Дискуссия позволяет сделать вывод о том, что преждевременно утверждать о замене человека GPT-ассистентами (Generative Pre-trained Transformer Assistant – GPTA) в среде социокультурной электронной коммуникации. Персонологический функционализм, обосновывающий замену людей машинами, опирается на психофункционализм Неда Блока, доказывающего необходимость психологизации машинного функционализма путем введения “смысла” в продукты прохождения оригинального теста Тьюринга. Для персонологического функционализма минимально необходимыми составляющими тьюрингового теста являются “творчество” и “смысл”. В работе показано, почему GPTA не проходят креативного теста. Для демонстрации неспособности пройти тьюринговский текст на осмысленность дорабатывается машина Блока путем совмещения нейрокомпьютерной (1978 г.) и символической версий (1981 г.). Для нового расширенного теста Блока сохраняется и усиливается аргументация предыдущих версий: машины типа GPTA не способны исполнять ни роли психологического функционализма – в частности, ни персонологического функционализма – в общем.

Ключевые слова: Генеративный ИИ; Комплексный тест Тьюринга; Тест Блока; Тест Лавлейс; Китайская нация; Машина Блока; Психофункционализм; Персонологический функционализм.

Благодарность: Статья подготовлена в рамках государственного задания ГАУГН «Цифровизация и формирование современного информационного общества: когнитивные, экономические, политические и правовые аспекты». Регистрационный номер НИОКТР 123022000042-0. Код темы FZNF-2023-0004. Регистрационный номер темы 1022040800826-5-5.2.1;6.3.1;5.9.1

Для цитирования: Алексеев А.Ю., Алексеева Е.А. Персонологический функционализм GPT-ассистентов // Technology and Language. 2024. № 5(2). С. 80-99. <https://doi.org/10.48417/technolang.2024.02.07>



This work is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/)



ВВЕДЕНИЕ

В настоящее время с появлением и повсеместным распространением генеративных предобученных трансформеров (Generative pre-trained transformers), компьютерные генеративные модели стали одним из самых заметных явлений современной электронной культуры. Им приписывается едва ли не революционный характер, их внедрение в различные сферы человеческой деятельности выглядит неизбежным, они спровоцировали новый виток научных и общественных дискуссий по поводу искусственного интеллекта. GPT-модели стали обозначаться словосочетанием “генеративный искусственный интеллект”, и это понятие получило признание в том числе и на правительственном уровне. И действительно, GPT-программы продуцируют тексты, изображения, музыку и многое другое, зачастую практически неотличимое от результатов человеческой деятельности, представленных в компьютерном формате. В силу этого возникает соблазн приписывать личностные характеристики и способности GPT-ассистентам – сервисным программам-помощникам, построенным на основе генеративных технологий различного рода. Подобное “персонифицирование” программных продуктов инициирует чрезвычайный интерес к ним. Из-за ажиотажа вокруг этих систем из поля зрения исчезают их очевидные недочеты: недостаточная глубина контекста, зависимость от качества обучающих данных, продуцирование “фейков”, “галлюцинаций”, прочей заведомо ложной информации и многое другое.

Разработка и массовое внедрение сразу несколькими технологическими компаниями таких систем стали возможны благодаря тому, что существенно возросла вычислительная мощность дата-центров, появились огромные массивы открытых данных, которые были сгенерированы пользователями, а затем по сути присвоены генеративными моделями в процессе обучения, что и дало такой заметный и неожиданный эффект. Однако при ближайшем рассмотрении GPT-системы не представляют собой никаких особых “прорывных” программно-аппаратных новаций. Они мало соответствуют приписываемым им свойствам “осмысленного” и “творческого” продуцирования текстовых или мультимедийных артефактов. По своим вычислительным принципам эти модели восходят к марковским цепям и байесовским сетям. В целом же такие модели являются “стохастическим попугаем”: “системой случайного соединения последовательностей лингвистических форм, которые присутствовали в обширных обучающих данных, в соответствии с вероятностной информацией о том, как они комбинируются, но без какой-либо ссылки на их значение” (Bender et al., 2021).

В данной работе мы ставим вопрос о правомерности использования понятия “генеративный искусственный интеллект”, о том, соответствуют ли в принципе GPT-ассистенты научным критериям исследований искусственного интеллекта. Проблема путей дальнейшего развития общества в контексте бума генеративных технологий не изучается. Об этом много других, как правило, восторженных работ.



GPT-АССИСТЕНТЫ И ПЕРСОНОЛОГИЧЕСКИЙ ФУНКЦИОНАЛИЗМ

Современные GPT-ассистенты (GPTA) часто демонстрируют персонологические параметры, которые принято считать характеристиками человеческой личности: разум, понимание, мотивацию, намерение, убеждение, смысл, творчество, мораль, красоту, целостность, субъектность и мн. др. Например, в текстах, продуцируемых с помощью данной технологии, помимо присутствия в них искомого пользователем определенного содержания, обнаруживаются лингвистические маркеры эмоциональных состояний, оценочных позиций, включая самооценку, параметры мотивации и т.п. Такие системы, в особенности при дополнении их мультимедийными аудиовизуализациями “личности”, способны сыграть роль актера коммуникации, вводящего в заблуждение взаимодействующих с ним людей. Обман при «подмене личности» средствами, подобными GPT, становится общим местом для обозначения рисков и угроз со стороны “генеративного ИИ”.

Выявление персонологических маркеров служит поводом для заявлений в СМИ, на научных и бизнес-конференциях о том, что GPT-системы встали на путь “сильного ИИ”. Декларируется, что дальнейшее развитие этих технологий может привести к формированию у них человеческих характеристик: креативности, способности к целеполаганию, принятию решений, моральному вменению и т.п. Здесь же возникает и вопрос о замене людей системами ИИ. Если для замены традиционными системами ИИ рассматриваются профессии с учетом работы в слабо формализуемых областях, но с опорой на автоматизацию рутинных процедур, то в случае с генеративным ИИ заменяются творческие, принципиально неформализуемые профессии. Вопросы приписывания персонологических феноменов результатам функционирования систем ИИ, оценка доверия к процедурам такого приписывания, начиная от дефиниций и выбора их компьютерной архитектуры, изучается в методологии, получившей в последние десятилетия название “функционализм искусственного интеллекта”. Эта методология лаконично представлена в (Алексеев, 2019) на уровне концептуальной организации современных проектов общего искусственного интеллекта: искусственной жизни, искусственного сознания (мозга), искусственной личности, искусственного общества, искусственного мира.

Функционализм ИИ – это направление научно-теоретических и инженерно-технологических исследований, ориентированное на построение и применение компьютерных репликаций (имитаций), репрезентаций (моделей) и репродукций когнитивных феноменов самого широкого спектра жизненных, психических, личностных и общественных проявлений. Исходной основой функционализма ИИ является психофизиологическая теория машинного функционализма Хилари Патнэма 60-х гг. прошлого века. В 1960 году он утверждал, что при изучении интерсубъективных вопросов ЭВМ может выполнять роль, инвариантную функциям мозга (Putnam, 1960).



Например, оценивать боль другого человека можно трояко: во-первых, бихейвиорально, по стонам; во-вторых, физикально, по изучению паттернов нейтральной динамики; в-третьих, функционально, по отказу “триггера № 36” в чипе, который репродуцирует нейроны, ощущающие боль, в составе работы мозга как системы. В 1997 году автор машинного функционализма отказался от собственных идей, указав на сложность выявления функциональных связей и в обществе, и в человеке: “Когнитивная физика невозможна, точно так, как ранее стала очевидной невозможность социальной физики О. Конта” (Putnam, 1997). Однако вряд ли это утверждение способно остановить сторонников современного функционализма ИИ. Сегодня в формировании функционализма ИИ выделяются собирательный, определительный и наблюдательный подходы (Алексеев, 2019). Собирательный подход отвечает за сбор, идентификацию, координацию, формализацию, систематизацию, унификацию, кодификацию всевозможных функционалистских теорий. Определительный подход обеспечивает анализ и выявление главных функционалистских характеристик, отношений, закономерностей, причинностей, инвариантных относительно содержания когнитивных феноменов. Наблюдательный подход позволяет оценивать различные отношения человека к компьютерному миру систем ИИ: онтологические, эпистемологические, логические, лингвистические, аксиологические, эстетические, этические, праксеологические и другие фундаментальные отношения.

Возникает вопрос: как вписывается идея “генеративного ИИ” в методологию современного функционализма ИИ, в обозначенные выше и фундаментальные отношения? И вообще вписывается ли она в эти подходы?

Машинный функционализм семидесятилетней давности апеллировал к идее машины Тьюринга. В те годы тест Тьюринга не достиг популярности и про статью (Turing, 1950) упоминали единицы. Поэтому Х. Патнэм в (Putnam, 1960) идею теста Тьюринга просто-напросто не заметил. Современный функционализм ИИ базируется на методологии комплексного теста Тьюринга, который систематически изучает принципиальные вопросы компьютерной реализации когнитивных феноменов самого широкого спектра. В (Алексеев, 2008, 2013, 2016; Лекторский, 2022) рассматривался “костяк” наиболее популярных тьюринговых тестов, отвечающих на вопросы: может ли машина, как человек, мыслить, сознавать, понимать, творить, любить, дружить, быть личностью, свободной, добродетельной и пр.

На первый взгляд кажется, что персонологический функционализм, изучающей вопросы “замены” естественных личностей функциональными “двойниками” задается тестами, которые близки проблематике компьютерного моделирования личности. Например, тест “искусственной личности” (достаточно подробно раскрыт в (Алексеев, 2016)) и тест “зомби” (Алексеев, 2020) воспроизводит метафизическую драму персонажей мысленных экспериментов философии сознания. Д. Деннетт полагает, что все мы, люди, на сегодняшнем этапе антропосоциогенеза, являемся всего лишь зомби, но можем стать сознательными в условиях глобального ИИ, который позволит четко осознать роль и место каждого



в системе человечества в целом. Дж. Серль, напротив, убежден, что мы в далеком прошлом обладали полнотой личностных способностей, а вот ИИ-системы ускоряют деградацию людей и скоро мы превратимся в зомби.

Эти метафизические картинки забавны, и при этом они полезны, так как выражают логически достаточные основания того, чтобы GPT-система проходила тест на “личностное”. Имеется много иных когнитивных феноменов, которые необходимо имитировать средствами GPT – понимание, подсознание, эмоция, самоопределение, воля... На наш взгляд, логически необходимыми для выявления функциональных персонологических инвариантов является “смысл” и “творчество”. Если нет этих феноменов, то ни о каких проявлениях “личности” утверждать нельзя, это просто-напросто логически невозможно.

ПРОХОДЯТ ЛИ GPT-АССИСТЕНТЫ ОРИГИНАЛЬНЫЙ ТЕСТ ТЬЮРИНГА?

GPTA проходит оригинальный тест Тьюринга, игру в имитацию интеллекта. Вопрос о прохождении тьюринговой игры в имитацию, как правило, предшествует другим вопросам любого частного теста из состава комплексного теста Тьюринга. Следует сразу отметить, что эта “игра” выходит за границы методологии бихейвиоризма: тьюринговые “диспозиции” вычислимы. Вычислимость трактуется не как арифметические преобразования, а как порядок следования программе. Этот порядок в общем случае кодируется целыми числами. В бихейвиоризме предметом когнитивного анализа выступает не некая ментальная структура сознания, а объективно фиксируемые параметры поведения (реакции), определяемые внешними воздействиями (стимулами). В лингвистическом бихейвиоризме стимулами и реакциями диспозиций являются вербальные выражения, и так как осуществляется манипуляция семиотическими комплексами, то не получается игнорировать ни семиотическую, ни прагматическую компоненту тьюринговой коммуникации.

Первая компонента предполагает отношение символьных выражений к объективной, субъективной или интерсубъективной формам реальности, т.е. выделение ментального параметра как *значения*. Вторая компонента определяет отношение носителя выражений к выражениям с учетом бесчисленного контекста познавательных, производственных, экономических, социальных, культурных, моральных и прочих отношений. Получается, что в диспозициях должен присутствовать *смысл* как специальный предмет программирования. Оценка успешности компьютерной имитации интеллекта осуществляется судьей и зависит от его наблюдательности, компетентности, ума, бескорыстности и пр. Судья вправе присудить игрокам не только правильный гендер, как изначально задумывал автор оригинального теста, но и наличие чувств, мотивов, воли и многого другого, в том числе и личностного. Конечно, необходим некоторый минимум компьютерных достижений имитации интеллекта, по крайней мере, способности вести диалог. Способен ли лингвистический бихейвиоризм, научившийся программировать диспозиции и, благодаря этому, преступивший порог функционализма,



имитировать креативность и осмысленность ответов на вопрос? Для оценки этого в составе комплексного теста Тьюринга изучаются тест Лавлейс и различные версии теста Блока.

СПОСОБНЫ ЛИ GPT-АССИСТЕНТЫ ПРОЙТИ ТЕСТ ЛАВЛЕЙС?

GPTA не проходит тест Лавлейс, т.е. тест на креативность. Впервые “тест” на компьютерное творчество предложила Ада Лавлейс в 1843 г.: машина способна на многое, например, вычислять, исполнять музыкальные симфонии. Но при этом способна делать только то, что ей указывает (инструктирует) сделать человек. Машина творить не может! (A.L.L., 1843).

А. Тьюринг в 1950 году сделал послабление в дефинициях: творческим является то, что удивляет наблюдателя и является неожиданным. Например, благодаря машинному обучению человек лишь предполагает общий ход продуцирования артефакта, конкретное же воплощение достигается в функциях (нейро)компьютера (Turing, 1960, p. 469). Но такая конкретика, по мнению А. Тьюринга, весьма необычна и её сложно предсказать.

В 2000 г. ряд авторов предложили улучшенный тест Лавлейс, в котором фигурировали “артефакты”, благодаря которым – подчеркнем это – GPT-системы стали популярны – картины, поэмы, музыка (Bringsjord, Bello, Ferrucci, 2001). Однако способна ли GPTA пройти тест Лавлейс? Компьютер пройдет этот тест тогда и только тогда, когда выполняется система аксиом: 1) GPTA способен продуцировать артефакт *A*; 2) артефакт *A* не является результатом обмана, случая или сбоя; 3) GPTA всегда может *A* репродуцировать (повторить); 4) судья *I* не способен объяснить, как GPTA продуцировал *A*, несмотря на полное понимание механизма GPTA. Предложенная аксиоматика приводит к следующим соображениям: 1) Креативный тест невозможно пройти путем обмана. 2) Тест Лавлейс всем современным GPTA присваивает статус некреативных. Модели запрограммированы разработчиками на основе больших данных, ничего “от себя” такие GPTA добавить не могут. 3) Необходимым условием прохождения теста Лавлейс является способность GPTA к программированию собственного программного кода, новых аксиом (для символического подхода), нового способа обучения нейронной сети (для коннекционистского подхода). Но это невозможно из-за аргумента Геделя-Лукаса: механическая система, которой является GPTA, не способна к аутоформализации. Для GPT даже отдаленно не стоит вопрос о том, способна ли программа порождать другие программы методами, скажем, “автогенеративного программирования”. Принцип работы GPTA принципиально исключает аксиоматику как вычислительную основу функционирования. Сегодня часто встречается утверждение о том, что GPT мгновенно формируют многообразие разных программ по требованию “промптов”. Однако на самом деле новые программы не создаются: GPT предлагает фрагменты кода, которые были ранее представлены в Интернете. Моделирование “свободной воли”, имитация мотивационно-волевых механизмов человека, стремящегося к созданию чего-то нового, дизрезультативно, так как продукт представляется в некотором



пространстве потенциально представимых решений. GPT-генерация не доходит до уровня аксиоматического представления этого пространства. Тест Лавлейс собственно и показывает невозможность компьютерной реализации феномена творчества. Однако возможна интерпретация продукта компьютерной генерации как квазикреативного, “как бы” творческого продукта, которая определяется компетентностью тьюрингового судьи, приписывающего программе креативность.

Чтобы повысить экспертные полномочия тьюрингового судьи в 2014 году М. Ридл в “тесте Лавлейс 2.0” (Riedl, 2014) предложил новую версию исходных предпосылок, которую применительно к нашему случаю представим таким образом: 1) GPTA продуцирует артефакт A типа T ; 2) A должен соответствовать набору ограничений C , где $c_i \in C$ – любой критерий, выражаемый на естественном языке; 3) человек-оценщик H , выбравший T и C , удовлетворен тем, что A является допустимым экземпляром T и соответствует C ; 4) человек-судья I определяет комбинацию T и C как нереальную для обычного человека. То есть помимо человека, выполняющего роль тьюрингового судьи, появляется человек (H), компетентно оценивающий гениальность, талантливость, креативность артефакта A или, напротив, бесталанность, обыденность этого произведения. Произошла четкая спецификация критериев оцениваемого продукта. Человек-оценщик и человек-судья должны быть экспертами в достаточно ограниченной области, т.е. нужен конкретный специалист, а не “человек вообще”, определяющий, может или не может творить машина. Самое главное, продукты генерации GPTA должны быть осмысленны.

В 2015 году Т. Пожарев предложил *тест Лавлейс 3.0*, когда для постижения смысла оценщик H должен выявлять факт осмысленного продуцирования артефакта A (Алексеев & Пожарев, 2020). Например, какая операция использовалась при создании A : абстрагирование или обобщение? Если оценщик H обнаружил абстрагирование, то возникает повод утверждать креативность. Если H обнаружил обобщение, то это повод отрицать креативность. Нейронные сети не формируют идеальные объекты, не оперируют ими, они способны только обобщать и ограничивать. Формально-логически это означает, что GPTA работает исключительно с объемами понятий, экстенционально. С содержанием понятий, со смыслами, интенционально, работает исключительно человек.

ПРОХОДЯТ ЛИ GPT-АССИСТЕНТЫ ТЕСТ БЛОКА?

GPTA не проходит тест Блока, т.е. тест на осмысленность. Тест Блока входит в состав комплексного теста Тьюринга и является совершенным частным тестом на “глобальный” искусственный интеллект. Тест Блока был сформулирован в двух работах Неда Блока (Block 1978, 1981) почти полувековой давности, но сегодня оказался чрезвычайно востребован в связи с GPT-экспансией. В ней представлены две версии проекта вычислительной машины для работы со “знаниями” всего человечества: коннекционистская и символьная. Одна выражает идею глобализации знаний в среде Интернета, вторая версия обеспечивает ответы по запросам, которые сегодня называются “промптами”. Первая версия,



коннекционистская (точнее, нейросетевая), предложена в статье Н. Блока 1978 г. “Проблемы с функционализмом”. Этот проект автор назвал “Китайской нацией”. Сегодня он иногда именуется “Китайский мозг” (“China brain”) (Bringsjord et al., 2001). В философии ИИ более известным является другой “китайский проект”, появившийся два года спустя – “Китайская комната” Джона Серля 1980 г. Комментаторы отмечают (например, (Cole, 2023)), что на серлевский мысленный эксперимент повлияла отмеченная нами работа Н. Блока (Block, 1978). Коннекционистская машина Блока масштабирует когнитивные функции психики отдельного индивида до уровня т.н. “общественного сознания”. На выбор названия повлиял размер населения Китая, сравнимый с количеством нейронов в коре головного мозга человека. Предлагается идея связать паттерны активности мозга китайца, которого избивают, с паттернами активности китайской нации в целом, когда каждый активный нейрон китайца-индивида активизирует нейрон китайца-члена общества. Если срабатывает нейрон, то через спутник Земли посылается сигнал на мобильное устройство члену китайской нации, скажем, прыгнуть. Если нейрон деактивируется, то поступает сигнал – замереть. В принципе, возможен функциональный инвариант между нейтральной динамикой мозга избиваемого китайца и китайской нацией. Ведь количество китайцев и количество нейронов сравнимо. Но будет ли нация испытывать боль, аналогичную боли китайца-индивида? На первый взгляд очевидно отсутствие каких-либо корреляций. Однако так ли это? Вопрос до конца не разрешим, как и вопрос об эпистемологическом реализме общественного сознания, ноосферы и глобалистики в целом. Конечно, тест Блока “Китайская нация” поднимает мировоззренческие проблемы, которые возникают в мысленных экспериментах в весьма далекой от реалистической презумпции лингвистической философии.

Вторая версия машины Блока, символическая, представлена в статье 1981 г. “Психологизм и бихейвиоризм” (Block, 1981). Она практичнее первой версии и позволяет изучать концептуальные модели построения глобальных интеллектуальных технологий. Эта версия называется по-разному: по-кибернетически звучит как “черный ящик” (“Black Block”). Иногда называют “голова-ящик”, иногда, ссылаясь на фамилию автора, “головой Блока” (Blockhead (Ned Block, 2024)). Н. Блок предлагает вообразить традиционную информационную, не нейрокомпьютерную, машину представления, хранения, обработки и выдачи ответов на все возможные вопросы. Глобальная машина собирает все эти лингвистические диспозиции, систематизирует их и по сути реализует поиск подстрок, так как в начало запроса включается дескриптор “диспозиции” в виде вопроса. Остается найти вопрос в глобальной базе данных и часть оставшейся подстроки подставить как ответ пользователю, который задал вопрос. Принципы работы машины поиска подстрок мало чем отличаются от общих способов функционирования GPT. Получается то, что сегодня реализуется поисковыми сервисами Интернета. Правомочен вопрос: является ли машина поиска подстрок системой искусственного интеллекта?

По отдельности эти версии машины Блока не подходят для концептуального анализа GPT-систем. Первая ориентирована на решение метафизических вопросов



философии сознания: можно ли представить в компьютерной системе феномены общественного сознания и каковы перспективы компьютерной реализации этого представления? Вторая версия не лишена практической ориентации, например, программирования интеррогативных (вопрос-ответных) сценариев использования семантических сетей на основе всех возможных открытых баз данных и знаний. Однако если две эти машины совместить в едином проекте, то получится вполне правдоподобная символно-коннекционистская машина реализации теста Блока. Назовем ее GPT-машиной Блока или *новой машиной Блока*. Эта машина учитывает традиционные для машинного обучения особенности построения GPT-систем: дискретизация “текстовых диспозиций” в “токенах” и хранение их цифровых аналогов в “мешках слов”, определение ассоциаций кодов и формирование пространства векторов, учет статистики частот встречаемости ассоциаций.

Попытаемся осуществить следующие аналогии в сфере методологии науки. Машинный функционализм концептуально реализуется индетерминированной машиной Тьюринга, психофункционализм – машиной Блока, а персонологический функционализм – новой машиной Блока. Психофункционализм сегодня сложился как самостоятельное направление в аналитической философии (Levin, 2023), в философии искусственного интеллекта в целом и в функционализме ИИ в частности. Психофункционализм дополняет идею функционализма Х. Патнэма: программа имитации психических феноменов, бихейвиорально реализующая реакции системы (выход) на ее стимулы (вход), необходимо включает психологические параметры. Как мы ранее отмечали со ссылкой на Х. Патнэма, боль человека представляется “*включением триггера № 36*” как коррелята соответствующих мозговых стимулов (Putham, 1960, p. 363). Помимо этого, согласно Н. Блоку и исходя из психофункционалистской интерпретации психофизического функционализма, в состав “теории боли” должна входить “*теорема № 17*”, которая на внутреннем языке мысли представляет “предикатор боли», определенный на “входе” из значений переменных “стол», “гвоздь», “сесть» и на “выходе” со значениями переменных “кровь”, “крик”, “стон” (Block, 1978, p. 236). Физикализм в этом построении опосредованно участвует как функциональное тождество физически отличных структур относительно компьютерного блока (*триггера № 36*), мозга человека, китайской нации, тины марсианина и пр., коррелированных с психологической *теоремой № 17* и с субъективно воспринимаемым феноменом “*боли*” от первого лица или с эмпатическим переживанием боли относительно оцениваемых субъектов (интерсубъективно, от второго лица).

Значимость психофункционализма обосновывается в следующем порядке. Вначале выдвигается как бы вполне очевидный тезис о том, что тест Тьюринга – разновидность бихейвиоризма. Тезис разрушается под натиском стандартных антибихейвиористских аргументов. Далее за счет нестандартных антибихейвиористских аргументов уничтожаются остатки тех нюансов лингвистического бихейвиоризма, которые привносит тьюринговая игра в имитацию. Программистские патнэмовские расширения лингвистического



функционализма разрушаются, то есть выявляется несостоятельность машинного функционализма и подчеркивается истинность психофункционализма.

К стандартным антибихейвиористским аргументам, применимым для обоснования психофункционализма GPT-ассистентов, относится, во-первых, очередь, *аргумент Чизхолма-Гича* (Block, 1981, p. 11) совмещающий идеи против бихейвиоризма при восприятии (выдвинуты Родериком Чизхолмом (Chisholm, 1957) и, в общем, ментальных актов (изучены Петером Гичем (Geach, 1957)). Эти идеи утверждают невозможность абстрагированного выделения в поведенческой диспозиции некоторого определенного психического состояния без учёта всей совокупности иных психических состояний. Допустим, бихейвиорист исследует желание человека съесть мороженое. “Желание” представляется в форме поведенческой диспозиции. В структуру этой диспозиции входит, например, намерение немедленно схватить мороженое, особенно если его дарят (“пока не передумали”). Однако возникает конфликт интересов. Возможно, что человек сомневается в том, что ему предлагают именно мороженое, а не муляж мороженого, предлагаемый шутки ради. Нужна так же уверенность в том, что за этим не последует противоречия с другими более важными желаниями. Не нужна ли ответная услуга, которая может оказаться непосильной? Таким образом, входы и выходы диспозиции изменяются весьма непредсказуемым образом. Относительно “боли” возможны следующие диспозиции, предложенные Х. Патнэмом. Возможны различные комбинации психических состояний, таких как: [боль + обычное переживание боли] или [отсутствие боли + желание обмануть, что боль переживается]. Во втором случае правдоподобны *суперспартанцы Х. Патнэма*, которым законы запрещают выказывать боль несмотря на ее испытывание. Правдоподобен *совершенный актер*. Он может успешно притворяться в том, что ему больно, хотя на самом деле это не так. Паралитики и Мозги в бочке не могут выказать боль, так как им нечем это осуществить.

На основе этих аргументов, но уже применительно не к боли, а к мышлению, Н. Блок утверждает: интеллектуальное поведение продуцируется комбинацией: [интеллект + склонность к мышлению]. Невозможны комбинации: [отсутствие интеллекта + желание обмануть в наличии интеллекта] или [отсутствие интеллекта + желание казаться разумным]. Умный может симитировать глупость, но глупый не способен симитировать разум – полагает Н. Блок.

Этих аргументов достаточно, чтобы разрушить лингво-бихейвиоральную концепцию игры в имитацию. После проведенной критической деконструкции Н. Блок формулирует новый тест Тьюринга: диалоговый интеллект, проявляющийся в разговоре людей – это способность продуцировать осмысленную последовательность вербальных реакций на некоторую последовательность вербальных стимулов, какими бы эти стимулы не были (Block, 1981, p. 18). Не совсем понятен логико-экзистенциальный статус последней части утверждения. Мы интерпретируем их в формате “смысловой индефферентности”: не важно, являются ли осмысленными либо бессмысленными входные стимулы. Главное, чтобы результат был осмысленным или был похожим на осмысленный.



Тест Блока как новый тест Тьюринга, *осмысленный*, а не случайно сгенерированный, снимает возражение Чизхолма-Гича: возможно сколь угодно много вариантов диспозиций, убеждений, намерений, желаний тому, что GPTA не будет расположен выдавать осмысленные ответы. Тем не менее, разнообразные варианты диспозиций никак не сказываются на *способностях* GPTA формировать такие ответы. Но такие способности не являются операциями машины. Это суть способности *людей* понимать стимулы и формировать реакции как осмысленные тексты. Снимаются аргументы “суперспартанец” и “совершенный актер”: надо обладать недюжинными интеллектуальными способностями для того, чтобы в совершенстве имитировать отсутствие разума. Невозможно притворяться разумным и не быть таковым. Для интеллектуалов-паралитиков и “умных мозгов в бочке” вообще нельзя найти опровергающие примеры – у них, возможно, имеется способность реагировать осмысленно, однако для реализации этой способности им не хватает средств, которые могут выразить способность постигать “смысл”. Теоретически “смысл” можно объективировать, например, вершиной треугольника Фреге. На практике же смысл, не являясь значением, значит. Это смысловое схватывание требует субъекта, владеющего и оперирующего смыслом, раскрытия внутреннего мира человека. Невозможно вообразить, что бессмысленная комбинация символов, демонстрирующая отсутствие интеллекта при ее производстве, способна, минуя человека, механистично, но осмысленно отвечать на произвольную последовательность стимулов.

Однако откуда в GPTA возможно продуцирование осмысленных реакций? Для “отражения” нестандартных антибихейвиористских аргументов Н. Блок кратко описывает, но детально раскрывает принцип работы машины, которая на уровне всего человечества способна производить осмысленную последовательность вербальных реакций на вербальные стимулы. Это так называемая машина поиска подстроки. Задача машины – по ответу на запрос сформировать поиск подстроки, представленной как фрагмент строки, которая была ранее сформулирована, отформатирована и сохранена в базе знаний диалоговых сценариев всего человечества. Машина Блока представлена почти полвека назад, в 1981 г., при этом она предвосхищает принцип работы GPTA.

Мы не будем выделять очевидные отличия, они прозрачны. Подчеркнем то, что машина Блока на поведенческом, бихейвиорально оцениваемом уровне выглядит как интеллектуальная. Однако знания её внутреннего устройства убеждают в полном отсутствии у неё “интеллекта”. Вся совокупность осмысленных вербальных выражений (которая к тому же является конечным множеством) на выходе такой машины задаётся исключительно человеческим коллективом (проектировщиков, инженеров, программистов и др.). Люди долго и упорно трудятся для реализации всевозможных последовательностей осмысленных ответов на предполагаемые последовательности вербальных стимулов. В ходе работы используются специальные инструментальные средства, например, средства автоматизации программирования, машинного обучения для предварительной разметки траекторий ответов, средства ранжирования статистики частоты встречаемости подстрок. Самое главное – люди творчески воображают и



решают, что считать осмысленной последовательностью языковых выражений, а что является бессмыслицей.

Такая машина, несомненно, является разумной. Однако интеллект, который машина “проявляет” как в форме актуального поведения, так и в форме диспозиции интеллектуального поведения – *суть интеллект людей, которые создают тексты, формулируют вопросы, осуществляют поиск ответов и при обнаружении неудовлетворительных ответов отрицают или исправляют*. То есть это система естественного интеллекта, но отнюдь не искусственного интеллекта. Машина Блока – это как двухстороннее радио, радиоприемник в одну и в другую стороны акторов диалога (Block, 1981, р. 21-22), которому абсурдно приписывать способность мышления и понимания. Столь же абсурдно приписывать “интеллект” компьютеру, если он исполняет роли, инвариантные функциям радиоприемника. Если теперь воспользоваться первой, нейрокомпьютерной версией машины Блока и соединить со второй, то получим функциональную эквивалентность GPTA с радиоприемником. Поэтому вся аргументация неинтеллектуальности машины Блока, тщательно разработанная для первой версии машины, переносится и для случая с GPTA.

Мы полностью согласимся с этой объемной аргументацией, указав на ряд интересных моментов, сохранив, по возможности, порядок последовательности аргументов, как было в (Block, 1981).

Первый абзац аргумента отражает аргумент Блока. Второй абзац – опровержение этого аргумента. Если нет концептуальных коллизий между GPTA и машиной Блока, в этом списке применяется аббревиатура анализируемой нами машины, т.е. “GPTA”.

1. *Чрезмерная антропологизация системы ИИ*. Интеллект, проявляемый машиной Блока, является интеллектом разработчиков – программистов, экспертов, операторов, инженеров по знаниям и др.

Машина не обнаруживает отличий между поведением, которое отражает собственный интеллект машины, и поведением, отражающим только интеллект разработчиков машины.

2. *Актуальность знаний*. В машине Блока содержатся только те знания, которые в нее введены. Принципы работы GPTA принципиально отличаются от классических принципов ИИ. В них нет гипотез, абдукции, индукции, традукции. В машине только факты, которые могут оказаться фейками.

Систему следует считать интеллектуальной, даже если в неё не заложены знания о текущих событиях.

3. *GPTA не является интеллектуальной системой*. GPTA как машина с внутренней механической структурой не интеллектуальна даже тогда, когда выглядит интеллектуальной в любом внешнем отношении. Система считается интеллектуальной в той степени, в какой её действия имитируют поведение действительно интеллектуальной системы. При этом игра внешне интеллектуальной системы ошибочно принимается за признак её интеллектуальности.



Так как операции машины Блока суть полностью имитация вопросно-ответных сценариев, то они не могут служить основанием для приписывания машине интеллекта.

4. *Интеллектуальный шовинизм.* Возможно, некоторый “шминтеллект” интеллектуальнее “интеллекта” в силу различия и даже несопоставимости информационных процессов для их реализации.

Так как внутреннее строение GPTA убеждает в том, что она успешно имитирует игру в имитацию интеллекта, не обладая интеллектом, то этого контрпримера, по мнению Н. Блока, достаточно для понимания того, “интеллект» – это нечто иное, нежели чем то, что имитирует интеллект.

5. *Легитимность машинного интеллекта.* Вообразим, что некто, обладающий позитивным правом, постановил считать, что человеческий интеллект – это машинная обработка информации точно такая же, как это осуществляет GPTA. Возможно так же, по утверждению позитивно-правового деятеля и то, что человек менее интеллектуален и, следовательно, менее интересен, нежели чем машина. Он же не имеет общечеловеческого масштаба хранимых данных, параллелизма, скорости вычислений. Можно ли утверждать, что люди менее разумны, нежели чем GPTA?

Подобного рода “законотворческие постановления” расходятся и со здравым смыслом, и с исследованиями когнитивной науки. GPTA не интеллектуальна, поэтому позитивное право приписывания интеллекта неинтеллектуальной машине, если и установится, то не надолго.

6. *Тождество мышления и GPTA.* Человек мыслит так же, как мыслит машина Блока (GPTA). Собирает знания на протяжении всей жизни и вероятно отвечает на вопросы.

Когнитивные процессы, без всякого сомнения, гораздо более механистичны, чем это принято считать. Однако огромна разница между утверждением, что люди более механистичны, и утверждением, что человек – это машина Блока.

7. *Комбинаторный взрыв машины Блока.* Является ли эмпирической гипотезой то, что интеллект – это способность выдавать осмысленные последовательности реакций, соответствующие входным последовательностям? Машина Блока невозможна из-за комбинаторного взрыва.

Существует порядка 10^{30} грамматических предложений длиной в 20 слов (Н. Блок ссылается на классическую книгу по когнитивистике Дж. Миллера, Е. Галантер и К. Прибрама (Miller, Galanter, & Pribram, 1960)). Предположим, что из них 10^{15} семантически корректны. Тесту Тьюринга длительностью в один час может понадобиться порядка 100 таких предложений. А это 10^{1500} строк – число, которое больше числа частиц во вселенной.

8. Однако проблемы комбинаторного взрыва преодолены GPTA. Пожалуй, одним из достоинств современного уровня вычислительной техники является суперкомпьютерная предобученная схема выдачи ответов на запросы.

В самом деле, для GPTA нет проблемы с этим взрывом. Машина обучается на последовательностях текстов, которые ранее уже были осмыслены и представлены человеком. Самостоятельно GPTA ничего осмысленного не формирует.



9. *Редуцируемость интеллекта к диалоговому интеллекту.* Насколько правомерна эмпирическая гипотеза о том, что интеллект – это способность выдавать осмысленные последовательности реакций, соответствующие входным последовательностям?

Отождествление интеллектуальности со способностью выдавать осмысленные ответы – это второстепенный принцип эмпирической психологии. Подобного рода отождествление может быть предложено как рациональная реконструкция неопределенно общей концепции интеллекта, которая, возможно, найдёт полезное применение в будущей теории эмпирической психологии. Люди со здравым рассудком могут неправильно понимать предложения, путаться. Любому же нормальному человеку, который участвует в длительном тесте Тьюринга, вскоре должно наскучить это мероприятие, он станет невнимателен. Говоруну с самого начала будут нести бессмыслицу, иногда извиняясь, что не слушали собеседника. Многие люди на замечания собеседника будут отвечать свободными ассоциациями, нежели напрягать свой разум, давая вразумительные ответы. Некоторые будут постоянно жаловаться на непривлекательность этих бесконечных тестов Тьюринга.

10. Самым лучшим проектом интеллектуальной машины станет *замена памяти разумом*. “Для теста Тьюринга неопределённой длины машина становится настолько большой, что попытка ее построить станет причиной провала средств финансирования проекта в черную дыру” (Block, 1981, p. 34).

Техническая изобретательность, сколь бы изощрённой она не мыслилась, выходит за рамки методологии моделирования речевых или диалоговых способностей человека. Эвристика недостаточно для этого.

11. *GPTA и аниматроника.* Лишь немногие люди смогут пройти тест Тьюринга, который длится девятьсот лет, и ни один человек не сможет пройти тест Тьюринга длиной в пятьсот лет (Block, 1981, p. 34).

Длина машины Тьюринга не является необходимым условием интеллектуальности. Лучше сравнить машину Блока с печально известными “машинами” прошлых столетий, в которых незаметно прятался человек, слушал вопросы и отвечал на них посредством механизмов. Казалось, что отвечает машина, но на самом деле отвечал человек. GPTA вполне способна играть роль такой “машины”. Она хранит “знания” людей, ничего не добавляя от себя.

12. *GPTA и робототехника.* GPTA – робот может вести себя так, как ведет себя человек, набивая текст на клавиатуре, когда философствует.

GPTA не опровергает иные точки зрения на искусственный интеллект, в частности, на робототехнику. Просто утверждается, что GPTA не является интеллектуальной машиной. Подобная робототехническая ситуация номологически возможна, т.е. не противоречит науке, когда машина Блока, помимо диалога, вычисляет траектории элементарных частиц при нажатии на клавиатуру манипуляторами робота. Однако на самом деле этот GPTA не философствует. То, что он делает, – это вычисляет траектории элементарных частиц, чтобы подражать человеку, занимающемуся философией. Манипуляция роботом дескрипциями человеческих размышлений само по себе не является размышлением. Ни в



классическом, ни в общем, ни в сильном, ни в глобальном, ни еще каком-то ином смысле трактовки понятия искусственного интеллекта.

Таким образом, лингво-бихейвиористская способность искусственной системы продуцировать “осмысленные” ответы не является логически необходимой для приписывания интеллекта GPTA-системе. В GPTA нет ни ума, ни глупости. Она включает совокупность подстрок, каждая из которых бессмысленна. За исключением одного смысла – когда-то кем-то этот кусок был осмыслен как часть некоторой целостной осмысленной конструкции. Порядок этой совокупности не имеет смысла. За исключением одного смысла: когда-то кем-то этот кусок текста был осмыслен порядком следования фрагментов текста. В ответе как в статистически обусловленном результате генерации нет никакого смысла. Осмысленность GPTA приписывается исключительно воображением наблюдателя.

“Мышление” машины” – суть мышление экспертов (программистов, в первую очередь). GPTA ни актуально, ни диспозиционально мыслить не может. Диспозиционально мыслит разработчик информационных средств GPTA. Актуально мыслит пользователь GPTA, приписывающий смысл продуктам ИТ.

С нашей стороны добавим еще один аргумент, который назовем *аргумент предикат предиката*. Он обусловлен психофункционалистским нюансом строения машины Блока. В аргументе Чизхолма-Гича (Block, 1981, pp. 11-12) демонстрируется идея системной обусловленности психологических феноменов и невозможности целочисленного кодирования диспозиций “вход – выход” из-за того, что представление некоторого феномена динамически расширяет состав связанных с ним феноменов. Поэтому требование построения динамической системы классификации и кодирования психических феноменов вряд ли сегодня можно прямолинейно реализовать, нужен язык второго порядка, определенный как “предикат предиката”. Однако второпорядковое понимание логического языка идет врозь с теорией и практикой современной вычислительной математики и информатики.

Персонофункционализм на порядок сложнее психофункционализма, если продолжить блоковское развитие формулировок “функционализма”. Дело в том, что “боль” – это весьма отчетливый феномен сознания. Его в самом деле можно обозначить, формализовать, указать диспозиции, выделить типовые и индивидуальные корреляты мозговой активности. А как быть с онтологическими параметрами личности – “смыслом”, “самостью”, “ценностью”, “творчеством” и пр.? По всей видимости, претензии GPTA на персонофункционализм, т.е. на замену цифровыми помощниками естественных личностей, станут правдоподобными тогда, когда GPTA воплотит проект включения ментального в машинно-функциональное, чтобы личности, и естественные, и искусственные, обживали “ментальную третью империю” Г. Фреге (Фреге, 2000, с. 471) или, скажем, “третий мир” К. Поппера, вместе со осмысленными высказываниями, истинными теоремами, оптимальными алгоритмами.



ЗАКЛЮЧЕНИЕ

Методологические исследования GPT-ассистентов, т.е. принципы построения и использования систем т.н. “генеративного ИИ”, следует изучать на основе теста Блока как тьюрингового теста, реализуемого в условиях глобальных информационных ресурсов Интернета посредством нейрокомпьютерной генерации осмысленных подстрок как фрагментов осмысленного текста. Обобщенные функции машины Блока инварианты общим принципам работы генеративного ИИ. Архитектура и принципы работы машины Блока убеждают в том, что она никоим образом не объясняет принципы функционирования естественного интеллекта. Поэтому системы генеративного ИИ не способны претендовать на роль системы “сильного ИИ”.

Более того, понятие “генеративного ИИ” невозможно подвести под понятие “система искусственного интеллекта”, потому что принципиально элиминирует из системы управления способами представления знаний эвристики как правила формирования способов функционирования компьютерной системы ИИ. Понятие “генеративный ИИ”, сформулированный на основе GPT-систем представляется поспешным и нелегитимным. Введение этого термина на уровне государственных стандартов не прошло предварительных серьезных исследований и научных дискуссий.

Произведения GPT-ассистентов не критически интерпретируются как высоко осмысленные, разумные, креативные. При этом неясным остается юридическая правомерность коммерческого присваивания фирмами-разработчиками генеративных систем знаний различных форм человеческой культуры, переработка их в обезличенном формате с потерей и трансформацией смыслов.

Интересно было бы изучить причины, почему нелегитимные системы ИИ неожиданно получили статус легальных систем ИИ (Перечень поручений Президента РФ, 2023). От тех, кто критически оценивает стратегии развития ИИ и перспективы электронной культуры многое зависит, чтобы “генеративный искусственный интеллект” не послужил причиной глобального дегенеративного естественного интеллекта.

СПИСОК ЛИТЕРАТУРЫ

- Алексеев, А. Ю. (2008). Трудности и проблемы проекта искусственной личности. *Искусственные общества*, 1(3). <https://artsoc.jes.su/s207751800000077-3-1/>
- Алексеев, А. Ю. (2013). *Комплексный тест Тьюринга: философско-методологические и социокультурные аспекты*. ИИнтелЛЛ.
- Алексеев, А. Ю. (2016). *Философия искусственного интеллекта: концептуальный статус комплексного теста Тьюринга*. [Дис. ... д-ра филос. наук]. Моск. гос. ун-т им. М. В. Ломоносова.
- Алексеев, А. Ю. (2019). Общефункционалистский концепт искусственной потребности как основа общего искусственного интеллекта. *Философские науки*, 62(11), 111-124. <https://doi.org/10.30727/0235-1188-2019-62-11-111-124>



- Алексеев, А. Ю. (2020). Мультиагентные категории робофилософии: роболичность и робозомби. *Искусственные общества*, 15(2). <https://doi.org/10.18254/S207751800009761-6>
- Алексеев, А. Ю., & Пожарев Т. (2020). Принципы креативной робототехники. *Технологос*, 4, 5-17. <https://doi.org/10.15593/perm.kipf/2020.4.01>
- Лекторский, В. А., Алексеев, А. Ю., Винник, Д. В., Гарбук, С. В., Черногор, Н. Н. (2022). *Методика оценки доверия к “искусственному интеллекту”*. Президиум РАН.
- Перечень поручений Президента РФ по итогам конференции “Путешествие в мир искусственного интеллекта”, состоявшейся 23–24 ноября 2022 года. (2023). <http://www.kremlin.ru/acts/assignments/orders/70418>
- Фреге, Г. (2000). *Логика и логическая семантика*. Аспект Пресс.
- A.L.L. (1843). Sketch of the Analytical Engine invented by Charles Babbage, Esq. By L.F. Menabrea of Turin, Officer of the Military Engineers. In R. Taylor (Ed.), *Scientific Memoirs, selected from the Transactions of Foreign Academies of Science and Learned Societies, and from Foreign Journals* (pp. 666 – 731). Richard and John E. Taylor.
- Bender, E., Gebru, T., McMillan-Major, A. & Shmitchell, S. (2021). *On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?* 🦜. In *FAccT '21: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (pp. 610-623). ACM. <https://doi.org/10.1145/3442188.3445922>
- Block, N. (1981). Psychologism and Behaviorism. *The Philosophical Review*, 90(1), 5-43.
- Block, N. (1978). Troubles with Functionalism. *Minnesota Studies in the Philosophy of Science*, 9, 261-325.
- Bringsjord, S., Bello, P. & Ferrucci, D. (2001). Creativity, the Turing Test, and the (Better) Lovelace Test. *Minds and Machines*, 11, 3–27. <https://doi.org/10.1023/A:1011206622741>
- Chisholm, R. M. (1957). *Perseving: A Philosophical Study*. Cornell University Press, Ithaca.
- Cole, D. (2023). The Chinese Room Argument. In E. N. Zalta & U.Nodelman (Eds.), *The Stanford Encyclopedia of Philosophy* <https://plato.stanford.edu/archives/sum2023/entries/chinese-room/>
- Frege, G. (2000). *Logic and logical semantics*. Aspect Press
- Geach, P. (1957). *Mental Acts. Their Content and their Objects*. Routledge & Kegan Paul.
- Levin, J. (2023). Functionalism. In E. N. Zalta & U. Nodelman (Eds.), *The Stanford Encyclopedia of Philosophy* <https://plato.stanford.edu/archives/sum2023/entries/functionalism>
- Miller, G.A., Galanter, E., & Pribram, K. H. (1960). *Plans and the Structure of Behavior*. Henry Holt & Co.
- Ned Block (2024, 22 February). In *Wikipedia*. https://en.wikipedia.org/w/index.php?title=Ned_Block&oldid=1209557312



- Putnam, H. (1960). Minds and Machines. In S. Hook (Ed.), *Dimensions of Mind*. New York University Press. Reprinted in *Mind, Language, and Reality* (pp. 362 – 385). Cambridge University Press.
- Putnam, H. (1997). Functionalism: Cognitive Science or Science Fiction?. In D. M. Johnson and C. E. Erneling (Eds.), *The Future of the Cognitive Revolution* (pp. 32–44). Oxford University Press.
- Riedl, M. O. (2014). The Lovelace 2.0 Test of Artificial Intelligence and Creativity. *arXiv*. <https://doi.org/10.48550/arXiv.1410.6142>
- Turing, A. M. (1950). Computing Machinery and Intelligence. *Mind*, 59(236), 433-460. <https://doi.org/10.1093/mind/lix.236.433>

REFERENCES

- A.L.L. (1843). Sketch of the Analytical Engine invented by Charles Babbage, Esq. By L.F. Menabrea of Turin, Officer of the Military Engineers. In R. Taylor (Ed.), *Scientific Memoirs, selected from the Transactions of Foreign Academies of Science and Learned Societies, and from Foreign Journals* (pp. 666 – 731). Richard and John E. Taylor.
- Alekseev, A. Yu. (2008). Difficulties and Problems of the Project of Artificial Personality. *Artificial Societies*, 1(3). <https://artsoc.jes.su/s2077518000077-3-1/>
- Alekseev, A. Yu. (2013). *Comprehensive Turing test: Philosophical and Methodological and Sociocultural Aspects*. AIntell.
- Alekseev, A. Yu. (2016). *Philosophy of artificial intelligence: the conceptual status of a Comprehensive Turing test*. [Dr. Phil. Dis.]. Moscow State University named after M. V. Lomonosov.
- Alekseev, A. Yu. (2019). The General Functionalist Concept of Artificial Need as the Basis of General Artificial Intelligence. *Philosophical sciences*, 62(11), 111-124. <https://doi.org/10.30727/0235-1188-2019-62-11-111-124>
- Alekseev, A. Yu. (2020). Multigenate Categories of Robofilosophy: Robolic and Robozobi. *Artificial societies*, 15(2). <https://doi.org/10.18254/s207751800009761-6>
- Alekseev, A. Yu., & Pozharev, T. (2020). The Principles of Creative Robotics. *Technologos*, 4, 5-17. <https://doi.org/10.15593/perm.kipf/2020.4.01>
- Bender, E., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). *On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?* 🦜. In *FACCT '21: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (pp. 610-623). ACM. <https://doi.org/10.1145/3442188.3445922>
- Block, N. (1978). Troubles with Functionalism. *Minnesota Studies in the Philosophy of Science*, 9, 261-325.
- Block, N. (1981). Psychologism and Behaviorism. *The Philosophical Review*, 90(1), 5-43.
- Bringsjord, S., Bello, P. & Ferrucci, D. (2001). Creativity, the Turing Test, and the (Better) Lovelace Test. *Minds and Machines* 11, 3–27. <https://doi.org/10.1023/A:1011206622741>



- Chisholm R. M. (1957). *Perseving: A Philosophical Study*. Cornell University Press.
- Cole, D. (2023). The Chinese Room Argument. In E. N. Zalta & U. Nodelman (Eds.), *The Stanford Encyclopedia of Philosophy*
<https://plato.stanford.edu/archives/sum2023/entries/chinese-room/>
- Frege, G. (2000). *Logic and Logical Semantics*. Aspect Press.
- Geach, P. (1957). *Mental Acts. Their Content and their Objects*. Routledge & Kegan Paul.
- Lectorsky, V.A., Alekseev, A. Yu., Vinnik, D.V., Garbuk, S.V., & Chernogor, N. N. (2022). *Methodology for Assessing Trust in "Artificial Intelligence"*. Presidium of the Russian Academy of Sciences.
- Levin, J. (2023). Functionalism. In E. N. Zalta & U. Nodelman (Eds.), *The Stanford Encyclopedia of Philosophy*
<https://plato.stanford.edu/archives/sum2023/entries/functionalism>
- Miller, G.A., Galanter, E., & Pribram, K. H. (1960). *Plans and the Structure of Behavior*. Henry Holt & Co.
- Ned Block (2024, 22 February). In Wikipedia.
https://en.wikipedia.org/w/index.php?title=Ned_Block&oldid=1209557312
- Putnam, H. (1960). Minds and Machines. In S. Hook (Ed.), *Dimensions of Mind*. New York University Press. Reprinted in *Mind, Language, and Reality* (pp. 362 – 385). Cambridge University Press.
- Putnam, H. (1997). Functionalism: Cognitive Science or Science Fiction?. In D. M. Johnson and C. E. Erneling (Eds.), *The Future of the Cognitive Revolution* (pp. 32–44). Oxford University Press.
- Riedl, M. O. (2014). The Lovelace 2.0 Test of Artificial Intelligence and Creativity. *arXiv*. <https://doi.org/10.48550/arXiv.1410.6142>
- The List of Instructions of the President of the Russian Federation Following the Results of the Conference “Journey to the World of Artificial Intelligence”, held on November 23–24, 2022. (2023)
<http://www.kremlin.ru/acts/assignments/orders/70418>
- Turing, A. M. (1950). Computing Machinery and Intelligence. *Mind*, 59(236), 433-460.
<https://doi.org/10.1093/mind/lix.236.433>

СВЕДЕНИЯ ОБ АВТОРАХ / THE AUTHORS

Андрей Юрьевич Алексеев, aa65@list.ru,
 ORCID 0000-0003-2230-311X

Andrey Yurievich Alekseev, aa65@list.ru,
 ORCID 0000-0003-2230-311X

Екатерина Алексеевна Алексеева,
 eaalekseeva@gaugn.ru
 ORCID 0000-0002-0006-5942

Ekaterina Alekseeva Alekseeva, eaalekseeva@gaugn.ru
 ORCID 0000-0002-0006-5942

Статья поступила 20 февраля 2024
 одобрена после рецензирования 11 апреля 2024
 принята к публикации 3 июня 2024

Received: 20 February 2024
 Revised: 11 April 2024
 Accepted: 3 June 2024



Discussion: André Leroi-Gourhan – Technology, Language, Humanity




<https://doi.org/10.48417/technolang.2024.02.08>

Research article

The Origin and Dissemination of Scientific Knowledge

André Leroi-Gourhan

Translation by Saskia Brown¹

Introduction by Nathan Schlanger²  (✉)

¹French Translators' Society, 22, rue de la Pépinière, 75008, Paris, France

²École nationale des chartes, 65, rue de Richelieu, 75002 Paris, France

saskia.brown@free.fr

schlanger1@gmail.com

Abstract

André Leroi-Gourhan (1911-1986) was a French ethnologist, prehistorian and paleo-anthropologist who is today appreciated for his influence on the philosophy of technology. His first publications on *L'Homme et la matière* and *Milieu et techniques* (1943, 1945) secured his reputation as a specialist in the study of material civilizations and in comparative technology. This perspective was enriched by evolutionary and anthropological considerations in his best known work, *Le geste et la parole* (1964, 1965). This book has appeared in English as *Gesture and Speech* in 1993, but not all of his relevant publications have been translated, and several aspects of his technological approach remain little known. The translation here of his March 1952 lecture at the Maison des Sciences in Paris, as part of a lecture series on “The structures of the universe and their scientific perception,” is an opportunity to highlight the interest and relevance of Leroi-Gourhan for contemporary reflections about technology. For example, a jointly haptic and cognitive “material engagement” is for Leroi-Gourhan characteristic of specifically human manufacture, of “materially creative activities” as undertaken by artisans of all times. We can recognize here Leroi-Gourhan's adhesion to Henri Bergson's philosophical tenet regarding the epistemological primacy of action over contemplation, and consequently the active, dynamic, vital origins of knowledge.

Keywords: André Leroi-Gourhan; Technology; Rationality; Physical and social evolution; Prehistoric flintknapping; *Chaine opératoire*

Acknowledgment We thank the estate of André Leroi-Gourhan for permission to translate and publish this paper.

Citation: Leroi-Gourhan, A. (2024). The Origin and Dissemination of Scientific Knowledge (S. Brown, Transl., N. Schlanger, Introd.). *Technology and Language*, 5(2), 101-115. <https://doi.org/10.48417/technolang.2024.02.08>



This work is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/)



УДК 001.9



<https://doi.org/10.48417/technolang.2024.02.08>

Научная статья

Происхождение и распространение научных знаний

Андре Леруа-Гуран

Перевод Саскии Браун¹

Введение Натана Шлангера²  

¹Общество французских переводчиков, Пепиньер, 22, 75008, Париж, Франция

²Национальная школа хартий, 65, де Ришелье, 75002 Париж, Франция

saskia.brown@free.fr

schlanger1@gmail.com

Аннотация

Андре Леруа-Гуран (1911-1986) - французский этнолог, специалист по истории доисторических времен и палеоантрополог, которого сегодня ценят за его влияние на философию технологии. Его первые публикации о “L’Homme et la matière” и “Milieu et techniques” (1943, 1945) закрепили за ним репутацию специалиста по изучению материальных цивилизаций и сравнительному анализу технологий. Эта точка зрения была обогащена эволюционными и антропологическими соображениями в его наиболее известной работе “Le geste et la parole” (1964, 1965). Эта книга вышла на английском языке под названием “Gesture and Speech” в 1993 году, но не все его соответствующие публикации были переведены, и некоторые аспекты его технологического подхода остаются малоизвестными. Публикуемый здесь перевод его лекции, прочитанной в марте 1952 года в Доме наук в Париже в рамках цикла лекций на тему “Структуры Вселенной и их научное восприятие”, является возможностью подчеркнуть интерес и актуальность Леруа-Гурана для современных размышлений о технологиях. Например, совместная тактильная и когнитивная “материальная вовлеченность”, по мнению Леруа-Гурана, характерна для специфического человеческого производства, для “материально-творческой деятельности”, которой занимались ремесленники всех времен. Здесь мы можем отметить приверженность Леруа-Гурана философскому принципу Анри Бергсона, касающемуся эпистемологического примата действия над созерцанием и, следовательно, активного, динамичного, жизненно важного происхождения знания.

Ключевые слова: Андре Леруа-Гуран; Технология; Рациональность; Физическая и социальная эволюция; Доисторический кремль; *Chaîne opératoire*

Благодарность: Мы благодарим наследников Андре Леруа-Гурана за разрешение перевести и опубликовать эту статью.

Для цитирования: Leroi-Gourhan, A. The Origin and Dissemination of Scientific Knowledge (S. Brown, Transl., N. Schlanger, Introd.). // Technology and Language. 2024. № 5(2). P. 101-115.
<https://doi.org/10.48417/technolang.2024.02.08>



This work is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/)



INTRODUCTION

Nathan Schlanger

The early 1950s represented something of a watershed for André Leroi-Gourhan (1911-1986). A lecturer in “Colonial ethnology” at the University of Lyon since 1944, he was beginning to reorient his research interests towards prehistoric archaeology, including prospections and excavations in nearby caves. At the same time, he was clearly keen to explore new research domains and bring together hitherto unrelated perspectives and fields of knowledge. To his previous concerns with the characterization of “material civilizations,” as inherited from the 1930's *musée de l'Homme* of Paul Rivet and Marcel Mauss, he now added an interest in the all-encompassing notion of technical behavior [*comportement technique*]. Setting aside ethnology – travel in space –, he was reaching towards prehistoric archaeology – travel in time – as well as technology,¹ comparative psychology and, admittedly in a more allusive way, cybernetics too. Thus, alongside modern and ancient humans, the horizons of his empirical research and philosophical reflections expanded to include also animals and machines (see on those issues Schlanger 2023, chapter 7, and Leroi-Gourhan, 2024).

On the strength of his manifest pluridisciplinarity, Leroi-Gourhan was invited in March 1952 to contribute to a lecture series at the Sorbonne on the broad theme of “The structures of the universe and their scientific perception.” Like other conferences in this cycle, his lecture was subsequently published in *Structure et Évolution des Techniques* (SET), a journal which notably served as the semi-official organ of the *Cercle d'études cybernétiques* – itself launched soon after the publication in 1948, simultaneously in France and the USA, of Norbert Wiener's *Cybernetics or Control and Communication in the Animal and the Machine* (Wiener, 1948) As intended by its editor, the philosopher of technology Pierre Ducassé, the journal aimed to provide an interdisciplinary venue for debating techniques, their forms and developments. Thus, issue 31-32 of the SET included contributions by historical epistemologist Gaston Bachelard and by Jacques Lafitte, author of the 1932 *Réflexions sur la science des machines* (Lafitte, 1932); issue 35-36, dedicated to “*La cybernetique*”, featured papers by mathematician Louis Couffignal, neurophysiologist Alfred Fessard, and Nobel prize winner physicist Louis de Broglie. Issue 33-34, the first of a “documentary publications” series, was dedicated to Leroi-Gourhan's article here translated.

Addressing the eminent scientists in the audience, Leroi-Gourhan provided a long-term perspective (what would be called much later a “deep history”) on the origins and development of scientific thought. The trajectory he outlined in these 10 pages extended from primate-like proto-hominids all the way to the rise of agriculture, sedentism and

¹ In the French research tradition, from Alfred Victor Espinas and Marcel Mauss to André Leroi-Gourhan and Pierre Lemonnier, the term “technology” essentially designates the logos or study of techniques, just like musicology is the study of music. Among other advantages, this specificity serves to distinguish between technicians, practitioners and engineers, and students of technology in the social and human sciences. As well, this designation helps us avoid commonplace or a priori demarcation between small-scale or ethnographic “techniques” and modern, science-based and industrial “technology” (see more in Sigaut, 1994, Schlanger, 2023).



literate civilizations – a succinct overview that prefigured the sweeping “anthropogenesis” panorama he went on to elaborate a decade later in his 1964-1965 masterpiece, *Le Geste et la parole* (Leroi-Gourhan, 1964; 1965). In this 1953 publication, it must be noted, his linear conception of cumulative progress was occasionally simplistic and partly outdated, regarding for example the supposed cleanliness of *Homo sapiens*, the role of military hierarchy in later prehistory, or the recourse to unwarranted ethnographic analogies. Where Leroi-Gourhan proved to be highly innovative, however, was in his understanding of both the processes of prehistoric stone tools manufacture and their wider implications. To the latest experimental results secured by prehistorians Léon Coutier and François Bordes he added insights from comparative psychology, partly inspired by the physiologist and behavioral psychologist Henri Piéron. This led him to conceive of flintknapping as a structured sequence of fairly complex operations involving various degrees of foresight, memory and consciousness – a process he was beginning to call the “*chaîne opératoire*.” This 1953 paper represents in any case a pioneering contribution to the study of “the prehistoric mentality” – nowadays known as “cognitive archaeology” – whereby, as Leroi-Gourhan put it in a contemporary paper for the *Revue de synthèse historique*, “to follow the gestures, flake by flake, is to reconstruct with certainty an important part of the mental structure of the maker” (Leroi-Gourhan, 1952, pp. 83-84).

This approach informed his contribution to the question at stakes, namely the origins of scientific thought: can we indeed see in the gestures and tools of the most ancient humankind (as evidenced in lithic technology) “a rational perception of causes and effects, which represents a fundamental characteristic of scientific research”? A key notion here, on which Leroi-Gourhan repeatedly insisted, was the close and continuous “contact” between the technical operator and the material in their hands. Specifically to stone tool manufacture, this tactile proximity reflected the inherent indetermination entailed in flintknapping, since both the variable properties of the raw materials (the flint nodules) and the uncertain delivery of the knapping blows make it necessary for the technician to constantly monitor and adjust their course of action. For our author, this confirmed that from the onset technical operations necessarily entailed technical speculations, which in turn heralded speculations *tout court*. Interestingly, in view of his own ideological (Catholic) commitments, Leroi-Gourhan also tied the emergence of speculation with that of religion. Far from setting religion and science in dichotomous or potentially conflictual relations, he saw in the former the fount of non-utilitarian practices and beliefs. Since the Reindeer age of the Upper Palaeolithic, at least, these beliefs found their expression in the symbolism of painted and engraved cave art and were thereby superadded to the “strictly technical” pursuit of subsistence needs. As Leroi-Gourhan further indicated in the text translated here, the “priests” who practiced such religious techniques were the first to be so “liberated,” thereby effectively leading the way towards the full-time specialized votaries of modern science.

Returning in conclusion to the key notion of material contact, two further implications can be suggested. From a phenomenological perspective, this jointly haptic and cognitive “material engagement” is for Leroi-Gourhan characteristic of specifically human manufacture, of “materially creative activities” as undertaken by artisans of all times. In some of his more fundamental anthropological conjectures, notably in *Le Geste*



et la parole, he insisted on the quintessentially human imperative of making and creating with our ten fingers – a constitutive engagement radically distinct from the ultimately impoverishing use of a single finger to flick machines and mechanisms into preordained motions.² At another level, we can recognize here Leroi-Gourhan's adhesion to the philosophical tenet, advanced notably in Henri Bergson's *L'Évolution créatrice*, regarding the epistemological primacy of action over contemplation, and consequently the active, dynamic, vital origins of knowledge (Bergson, 1907). Indeed, to use another Bergsonian notion which considerably preoccupied Leroi-Gourhan in these years (though not explicitly in the paper here translated), the challenge was to show just how important was the continuous and cumulative contribution of *homo faber*, the fabricator, in the emergence of the rational and scientific *homo sapiens* that we are.

THE ORIGIN AND DISSEMINATION OF SCIENTIFIC KNOWLEDGE³

André Leroi-Gourhan

This evening, I would like to explore with you the paths taken by humankind from its origins to the point at which it entered the period of major scientific speculation, to see when the first concerns with rational research emerged in the history of human societies.

Prehistory and ethnology may be able to answer the questions we must ask at the outset, at least partially. We can ask whether technical progress is really linked to [scientific] research as we understand it today. We can also ask what can be grasped of the first forms of scientific awareness and to what extent humans' first gestures and first tools indicate a rational perception of causes and effects, which represents a fundamental characteristic of scientific research. We can therefore ask from what exact moment researchers can be distinguished from manufacturers [*fabricants*], and indeed whether this distinction should be made at all. This prompts us to untangle the links existing for many animals between psychical reactions and technical behavior, between the latter and the manufacturing techniques specific to human beings, between manufacturing techniques and invention, and between invention and scientific speculation.

Let us first define the means available to us for carrying out this research.

Before us we have the whole field of history, but it is actually extremely short, covering a bare three or four thousand years, if one includes in it much of proto-history. Beyond it lies the field of prehistory, whose extent defies the imagination when compared to historical times. Indeed, if we take the timespan between ourselves and the beginnings of Egyptian civilisation to be a single unit, it would make up perhaps no more than 1/100,000th of the time between us and the first manifestations of properly human activity.

² As Leroi-Gourhan put it, for humans “not having to think with their ten fingers amounts to missing part of their normally, phylogenetically human thought” (Leroi-Gourhan, 1965, p. 61-62 [translated Leroi-Gourhan, 1993, p. 255]). Such questions of material engagement and skills are nowadays associated with anthropologist Tim Ingold, who is himself much indebted to the thoughts of Leroi-Gourhan and of Bergson (see Ingold 2004; 2013; 2014).

³ “Origine et diffusion de la connaissance scientifique”. Conference given at the Maison des Sciences, Paris, on Wednesday 5th March 1952, see Leroi-Gourhan (1953) [note added by editors].



What documents can we use?

In truth, the prehistorian possesses very few elements for these 999/1000th of the time of humankind. For the most recent phases of prehistory, from about 40,000 to 10 or 15,000 years before the present era, we have stone objects and also a large number of works of art whose content and situation give us a vague idea of the intellectual concerns humans had in the age of the reindeer.⁴ Further back, that is, between 40,000 and perhaps 500,000 or 600,000 years, we have only the sparsest vestiges of human beings, most often limited to a small fragment of the jaw or the skull, along with the thousands of knapped flint objects that have survived all the geological periods without deterioration, and bring us proof of the presence of beings from hundreds and thousands of years ago (we shall not call them humans for the time being, but we return to this shortly). It ensues that these beings could make objects of a certain form and for a certain end. On the basis of this very sparse material evidence, prehistorians have managed to divide prehistoric times into three periods: the Lower Palaeolithic, the Middle Palaeolithic and the Upper Palaeolithic. We will see later that the Upper Palaeolithic corresponds to the period when humans quite similar to us lived, from about 40,000 to about 10,000 years ago; further back, in the Middle Palaeolithic and the Lower Palaeolithic, they were physically very different from us.

However, are these testimonies – some very rare human physical remains, and the innumerable worked stones that have survived the disappearance of all that surrounded these humans at the time – sufficient to provide us with some trace of the evolution of human thought, are they reliable enough for us to be able to identify an evolutionary pathway? We can answer Yes, at least in the broadest terms. There is no question here of producing a complete history of the evolution of human thought, but we seem to be able to follow up two or three particularly clear paths. These humans of the Lower Palaeolithic were very different from us, and on the whole they corresponded to forms that can be considered to be intermediaries between primates (of which current days anthropoid primate give us only an approximate idea, not being bipeds) and the *Homo sapiens* we ourselves represent. What can we imagine the mind of these beings to be? What do we know about their technical thought?

THE ELEMENTARY STAGE OF TECHNIQUES

For these intermediate beings our evidence comes almost exclusively from large shaped stones which are called bifaces, and are found over an extremely long period, not only in France and Western Europe, but in most of Africa and large areas of Asia. We shall consider the biface to be the most characteristic object from this period (there are others, but we know this one best). What were they used for?

Everything suggests that these were the knives of the period, and not, as has sometimes been proposed, handaxes [*coup-de-poing*] or axe blades. They were the tools

⁴ The chronological values mentioned here are symbolic only. They are based on geologists' speculations and express probable orders of magnitude, confirmed in part by the methods of astronomy and physics. However, we still have no absolutely tried and tested method for measuring geological time over the whole of the Quaternary period.



with which humans cut up rhinoceros or the deer they had killed, and with which they probably chopped down the branches they used to make clubs and spears. They were an all-purpose tool, like today's knives or machetes.

I will use this example to follow up the development of technical thought over time with you, and to see to what extent we can go further than purely material forms of stone. Take this object, which weighs 500-600 g. It was knapped out of a flint block weighing approximately 1 kg at least. At the beginning of the Quaternary Era, humans obtained from this 1 kg a useable blade of about 20 cm. So we know that in the Lower Palaeolithic period, 1 kg of flint was required for 20 cm of cutting blade. I should add that experiments have shown that a flint blade wore down quite quickly, so knives had to be changed frequently. When we get to the second Palaeolithic period, the Middle Palaeolithic, we find human beings who are already entirely different physically; humans of this period belonged to the Neanderthal type, that is, with a reduced cranial capacity for the most ancient, and equal to ours for the most recent, with very low foreheads, enormous eyebrows, and much more prominent and longer faces than ours. Their limbs still retained certain primitive features, and what we know of how they lived seems to indicate that their material customs were at the very least quite similar to the Tasmanians or Fuegians encountered by nineteenth-century travellers. They hunted, gathered roots and tubers, and lived either in the open or in the shelters they could find, and we know from excavations of prehistoric sites what these "interiors" must have looked like. Large quantities of the Neanderthal's food remains were found in these caves. There is often a sort of bulge around the edge of the rooms, some 50 to 75cm deep, composed of bisons' ribs, the remains of horses' hooves, of reindeer antlers, rhinoceros fragments, etc., all smashed into bits in order to extract the marrow before being thrown to the edge of the room just as they were after eating. One can well imagine the comfort of these dwellings, not to mention the prevailing odours.

What do we know about the intelligence of Neanderthals?

We have a large number of documents in our possession that can give us a more human idea of their techniques than we might have supposed from their living spaces and the form of their skulls.

THE DEVELOPMENT OF RATIONAL TECHNICAL OPERATIONS

A considerable distance has been covered since the rudimentary bifaces of the Lower Palaeolithic. If we keep to the example of the flint core weighing 1 kg that was used to make the biface, we can reconstruct the operations used by Neanderthals to manufacture their knives. Using a pebble to trim the mass of the core, the Neanderthal shaped it so that he could quickly obtain, for example, some thirty triangular flakes, which were just as efficient as cutting tools as was the biface. So we now have thirty flakes, the equivalent of three metres of usable cutting edge. This is very important, because it is the first precise picture we have in the history of humankind of a technical process of such complexity: the systematic knapping of triangular flakes characteristic of the Mousterian represents a series of a dozen gestures following each other in an absolutely rigorous order. This is the first evidence we have of the sustained unfolding of technical



intelligence. It is also the first time in all of human history that we can observe the setting up of such a complex arrangement of operations [*un dispositif opératoire complexe*]. The “serial” production of triangular flakes supposes, in strict order: the choice of the block of raw material, its roughing out according to the shape to be given to the core, the preparation of the striking platform and the ridges, which determine the precise shape of the future flakes, the preparation of the point of impact, the prior shaping of the butt of the flake, the thinning-out of the prehensile part of the flake, and finally the precisely placed delivery of the final blow, which detaches the finished product. It is important to note that the product that emerges from the block in its final form is the result of a series of technical actions, whose sequencing [*enchaînement*] requires at every stage a veritable technical reasoning and a best possible fit between the operatory knowledge of the flintknapper and the [unpredictable] variations of the material used.

Alongside this first body of evidence, we have another one. I was talking earlier about the countless numbers of bone fragments found in shelters and caves from the Middle Palaeolithic. When these debris are examined very closely we find traces of cutting with flint knives on many pieces. After making a precise record of all these traces, we managed to reconstruct, in a rigorous manner, the gestures or series of gestures which were performed by the Neanderthals to cut up reindeer or bison⁵. We know, for example, thanks to the traces that have sometimes survived on the cervical vertebrae, that the flint blows used to detach the head were always aimed at a specific place, the one where a butcher today places his knife to detach the head of cattle. We know for certain that, when skinning a reindeer, a flint blow was made to the carpals in exactly the same place as the skinning knife still cuts today. And when one wanted to detach the foot entirely, the blows were designed to land a little lower down. I will not go further with describing the Neanderthals' butchery procedures, but once again we have before us a sequence of relatively complex technical operations, which enable us to extract from the heaps of detritus cluttering up the inside of a cave some moving evidence of a very ancient state of technical thought. What preliminary conclusions can we draw from this?

Our findings show that even before homo sapiens appeared, that is, before humans resembling us physically became cave dwellers, the technical domain already evidences standardised operational thinking [*une pensée opératoire très rationalisée*] involving series of gestures set in precise sequence. Furthermore, at the end of the Neanderthal period we find the beginnings of concerns of a very different order, with the appearance of the first graves, and the first aesthetic preoccupations. That is, we discovered on several occasions that in the most recent layers of Neanderthal occupation skeletons that had been deliberately buried, as well as objects of curiosity such as large fossil shells or blocks of iron pyrite or pieces of red ochre, which had been collected from afar and brought back to the caves.⁶ This represents the very first evidence of a psychical orientation that departs

⁵ Doctor Henri Martin's research. [Léon Henri-Martin (1864-1936), a physician and archaeologist, was a pioneer in the field of archaeozoology; added by editors]

⁶ Leroi-Gourhan is referring here to his recent find at his excavations at Arcy-sur-Cure (Grotte de l'hyène) of several fossils and unusual mineral items. As he would go on to argue, this “Palaeolithic collection” represented not only the first evidence of a deeply-rooted human fascination with the aesthetics of nature, but also “the adolescence of the natural sciences, for the dawning of science in all civilizations took place



from strictly alimentary concerns. Such evidence is extremely sparse, but very valuable, since we can see in these humans, whose physical appearance was not yet like ours, the emergence of certain preoccupations that would later become of capital importance.

As regards fabrication, we can see that from the outset techniques have followed a very steady line of progress. I have spared you the interminable stages between the first biface and the perfected triangular flake; these advances occurred imperceptibly, through the contact between the flintknapper and the material he had in his hands, the aim he set to achieve and the block in his possession. This manual contact shows that, at the origin, there was no possible divorce between operational activity and technical speculation [*entre l'opération et la spéculation technique*]. We will see later that this attitude of progressive contact between manufacturers and the material on which they operate is one of the foundations of technical development.

The dimension of progressive contact is a specifically human feature. We might to a certain extent consider that the action of the human or the anthropoid who uses the first stone is comparable, at the origins, to what today's anthropoids do when they use a stick to hit an object, or to the classic scene of the monkey aiming a coconut on the explorer's head. In fact, what is at stake is something quite different; in all times, monkeys have never done anything other than throw their coconuts, whereas, as soon as the first stone was knapped by a being from the human lineage, we can see that stone evolve, in exactly the same way as an ammonite or a series of mammals has evolved through geological strata.

THE DEVELOPMENT OF AESTHETIC AND RELIGIOUS TECHNIQUES

We have now reached the Late Palaeolithic, that is, the time when, at least in European regions, human beings emerged who were very similar to ourselves. Concerning these humans, considerable evidence exists of activities that are not strictly focused on food. When excavating a prehistoric site, it is possible to tell, even from the nature of the ground alone, that something important happened after the disappearance of the Neanderthals. We no longer find the heaps of detritus left around the hearths, all the mess of a dwelling place where the inhabitants had no concern for good housekeeping, except for occasionally making a space for themselves so that they could sit down among the carcasses of reindeer and bison. We find, on the contrary, that the interior floors have been regularly swept, and that the living surfaces have been levelled, and that there is evidence of aesthetic and religious activity from the very earliest times of occupation by “homo sapiens.”

The advent of the Upper Palaeolithic marked a major increase in aesthetic and religious production. This was the period when humans covered the depths of their caves with engravings and paintings, when spears and spear throwers were carved, when shaft straighteners were used to straighten wooden lances and arrow shafts, when pebbles were

amid the bric-a-brac of so-called curios” (Leroi-Gourhan, 1965, p. 213-214 [translated Leroi-Gourhan, 1993, p. 367-368]) [note added by editors].



engraved, and statuettes sculpted. We can observe that some of the individual's time of activity is now devoted to tasks that are no longer immediately related to food.

In fact, this first impression should be rectified by pointing out that all Palaeolithic art is linked to food by magic, since apart from the female statuettes, which are part of a fertility cult, all the representations we know of are of animals that are hunted. Yet it is still important to note that an aesthetic-religious activity has now been grafted alongside this strictly technical activity. A new development has also taken place in the techniques themselves.

We know that clothing became more elaborate. While we have no evidence of what humans wore in the age of the reindeer (since all furs and skins have disappeared), we do at least have their sewing needles and a certain number of tools used for working leather and hide. And we can tell from the fineness of the needles that sewing techniques were well developed.

A much greater variety of weapons appeared. If we return to the example of knives, we can observe a significant economic change in flintworking techniques. Starting with our initial block weighing 1000g, what is produced is no longer a single biface, which needed to be replaced after a relatively short period of use, nor are these thirty triangular blades which represent (in comparison with the biface) three meters of cutting edge, but rather a hundred or so much finer and much lighter blades, which represent about ten meters of usable cutting edge. This is important from an economic point of view, if we bear in mind that flint-rich regions are relatively few in number, and that the distance generally travelled in order to procure the raw material was around ten and sometimes more than a hundred kilometers. With the progress initiated by the appearance of the biface, we have the first quantifiable evidence of economic evolution in the history of humankind. The evidence is precise, because we can weigh and measure the tools, make statistical calculations, and represent this evolution through graphs. The humans we are dealing with occupy a field identical to that of present-day populations, who are still attached to an economy based essentially on hunting and fishing.

These societies still offer only limited possibilities for the development of collective means of action, because these groups number few, if any, specialists who are freed from directly food-related activities. Only in the case of religious techniques can the specialist be said to benefit from material freedom [*libération technique*]. Given our experience of present-day primitives and from what we can conjecture (I would almost say observe) through the countless religious documents left to us by prehistoric humans, we realise that in this domain there must already have been individuals who held priestly roles. A large body of evidence also convinces us that the priest's function involved new techniques and in particular the use of graphic symbols. The symbolic transposition attested by the reindeer-age art attests opens up a new path for thought, one that leads to writing.

We have come a very long way. We can now see in human societies the emergence, in addition to technical progress that is related to a normal biological evolution that would be comparable to that of animal series, also the onset of concerns which will directly give rise to scientific speculation.



THE ECONOMIC LIBERATION OF THE TECHNICIAN

The next period begins at the end of the last glacial extension, around 8,000 BCE. In the eastern Mediterranean and central Asian regions, where conditions were favourable, agriculture took over from hunting. We do not yet exactly understand the origins of agriculture. But we do know enough to imagine that, in a favourable botanical environment, groups devoted largely to gathering gradually shifted to domesticating a certain number of plants. The transition to agriculture (including animal husbandry) determined all future development, such that it is the most important event to have occurred as regards the evolution of human societies, from the beginning of the Quaternary Period to the present day.

Agricultural techniques, which spread with extraordinary speed over most of the inhabited world, brought about a complete transformation of the political and economic systems. First, populations became almost completely sedentarised, due to the obligations imposed by agriculture. People were required to stay on the land, at least as long as it took for the plants to grow. Resources were amassed in large quantities at the point where they were cultivated, which led to a much greater concentration of human beings than in the case of hunting or fishing resources. This brought about a need for defense systems against both animals and surrounding human groups. Concentration and defense systems: this is almost urbanisation. From the earliest developments in agriculture, one can find enclosures and other mechanisms for isolating the tiny agricultural communities that had managed to settle in the region at points of strategic importance. Sedentarisation, concentration and defense inevitably led to the need for a single, individual command. The hunting-fishing communities we know of rarely concentrate authority in the hands of a single individual. More often than not, there was a sort of diffuse authority that existed between the elders and between the middle-aged men, but as soon as it became necessary to build a camp or to defend it, a social hierarchy, largely military in nature, was established, which gradually led to kingships and empires. From this point onwards, the social organism took over partially from the individual, and this fact is of great importance because in the societies we were observing earlier, individuals were versatile: within the group, everyone was capable of carrying out all the tasks that enabled the group to survive, and this made it possible for individuals or isolated families to meet their needs.

A completely different social organisation emerged; its possibilities were considerably enhanced by the fact that certain individuals could be asked to do work that was no longer focused on the immediate search for food. We have proof of this: as soon as agriculture appeared, techniques appeared extremely rapidly that implied knowledge going far beyond the elementary exploitation of the environment - pottery, weaving and metallurgy, as well as surgical and medical techniques.

We have considerable evidence on this later point, since the graves, dolmens and funerary stone chests frequently found show us the marks of surgeon's flint knives on the skeletons of the exhumed individuals. We can thus observe an extraordinary technological blossoming, and simultaneously more complex social structures of command and hierarchy. This is also the period during which we see the priest's functions



reflected in large monuments and megaliths, and documentary evidence of the first calendars and astronomical activities. As a consequence, we find with the apparition of agriculture a totally different intellectual horizon from the one I outlined at the beginning of this lecture. A total transformation took place in the relations between technical activities and the technician. We observe that groups are truly in a position to act as super-individuals, as an individual-collective [*une sorte de super-individu, d'individu-collectif*], if I may venture such an expression, whose means were significantly increased and diversified than were those of hunter-fishermen. Research at this stage is still (and will largely remain in the future) connected to the processes involving material contact; the manufacturer is one who works on matter, who follows it as it were, and never loses manual contact with it – but it was the liberation of a certain number of specialists that allowed invention to flourish. A whole series of innovations appeared in the technical field. Metallurgy continued its rapid development, and in the space of a few centuries it reached the heights that led to the historical period. This is due to the fact that the blacksmith no longer devoted his time to hunting, but to making metals, the best bronze and the best iron that the community might demand of him.

From this point on, the specialists split into several currents. The religious concerns that had appeared progressively in earlier stages now led to growth in the techniques which long remained related to magic: these were the healing techniques, which gave rise to medicine and pharmacy, and the techniques linked to cosmogony, which gradually led to astrology and geomancy. Thus religious techniques began to break out of their religious confines around the figure of the priest. Another current was related to the functions of command and of military action. Military action determined the development of metallurgical research and defensive architecture: following the progress in metallurgical techniques and the geographical movements of bronze and iron over the last three millennia, we can see that everything is linked to research (still very empirical by the blacksmith or the founder), which was carried out under the pressure of the group. In another field, a trend appeared that was to be of prime importance. Command implies economic concentration which implies in turn the constitution of an aristocracy and ruling classes. It was noted that the origins of writing were largely linked, on the one hand, to economic concerns, that is, to put it coldly, to accounting and, on the other hand, to the techniques of prestige of the aristocracy of the time. Whether we look at the origins of Chinese or Mediterranean writing, we find in both initial concerns for the material recording of facts of an economic or religious nature, such as temple accounts or lists of dynasties, as documents justifying power. The other techniques, those not directly related to political and economic concentration, remained purely manual. Their progress is not related to the extraordinary impetus given by the concentrated nature of groups and their need both to defend themselves and to conquer new territories. Such techniques benefit therefore only very indirectly from this evolution, and they will constantly lag behind. This is true of pottery and of weaving, and it is especially true of agriculture and animal husbandry which, in all societies up to and including our own, have remained at a stage relatively close to that of their origins. Farming, weaving and pottery remained Neolithic until an extremely late stage, even in the most advanced societies of the Mediterranean and Chinese civilisations.



SYMBOLS, WRITING, AND SPECULATION

This process leads us to the expansion of the Greek and Chinese cultures. In the organisation that emerged from agriculture and sedentarisation, we see the possibility of individual research. But it only appeared at a very late stage, and partially; it was never completely separated from manual techniques. The Sumerians, the Mesopotamians, the Egyptians until the Middle Kingdom, the Chinese until their Confucian proto-history, and the great civilisations of pre-Columbian America had not yet detached speculation from utilitarian and religious practices. These civilisations still belonged to an intermediate stage between those we have already considered and the one we are about to enter.

We find the conditions of the Upper Palaeolithic in hierarchical societies, which constitute higher social forms. An important event occurred during this period: using methods borrowed from divination and techniques of prestige, writing came into being. What it introduced was not so much, perhaps, a means for preserving and transmitting knowledge as a means for expressing this knowledge in graphic symbols, abstracted from manual practices. Just as agriculture was pivotal for human societies, so the emergence of writing was pivotal for scientific thought. As soon as graphic symbolism left its imprint on thought, current forms of research appeared. In some limited fields, especially religion, particular practices became scientific speculation. As a result, we can follow the evolution of human thought towards speculation, not through reasoning (by projecting into the past the forms of our current thinking, a process which can only lead to extremely vague notions), but through the materiality of the documents provided by prehistory, archaeology and the earliest texts.

The first stage [of human development] is known to us through precise facts: with the earliest tools, contact between humans and the material world involved relationships that were initially very similar to the psychomotor reactions of animals. But since that stage, when the human body was still far from having reached its present degree of development, humankind was distinguished by the progressive character of its industry. This fact clearly sets apart even the pithecanthropoids and the sinanthropoids from anything that may have preceded them in the order of primates.

We observed a second stage when, at the margins of the technical domain, seemingly useless activities appeared. These were the aesthetic-religious activities of the reindeer age, which added something to the simple facts of material contacts. Finally, there gradually appeared a third stage, with the emergence of agriculture, beginning with intensive gathering in a favorable botanical environment. The advent of agriculture radically altered the relationships between the individual and the group. We have observed how an ever-greater impetus to progress became concentrated in the vital techniques of collective defense, after which writing soon appeared. From writing onwards, we see increasingly fluid symbolisation and a growing dissociation between manual contact and speculation. It should be noted that the relationships between the human being, the material world and ideas, which characterise all the stages we have considered, do not disappear at each new stage. Rather than substitutions, we observe an enrichment, the gradual constitution of a physical and mental system [*dispositif*] of extraordinary diversity, well able to extend into new fields. But until today, manual



practice retains its primordial value. Thus, such is the all too rapid overview I have presented to you of what we can know about the earliest forms of the evolution of human thought. We have seen that our materials of study are taken from the complementary sciences that are human palaeontology and prehistory. I have only alluded here to the field of human palaeontology, which will be the subject of another lecture in this series. I have confined myself in my presentation to what the prehistorian can perceive, but these two sciences complement each other; one shows the paths followed by hominids to arrive at present-day forms, the other shows the progressive enrichment of the techniques of material life.

There remains a very important area to examine, namely the progressive detachment of abstract thought, as we have tried to follow it up over the course of time. This progressive development represents humankind's most distinctive and original feature. I have tried to show in this lecture that in this respect [scientific] research is not totally devoid of means of investigations. While considerable gaps in the information available still remain, it is perhaps not illusory to seek in the most material and concretely verifiable facts provided by prehistory and archaeology a very first outline of the evolution of scientific knowledge.

REFERENCES

- Bergson, H. (1907). *L'Évolution créatrice* [Creative Evolution]. Presses universitaires de France.
- Ingold, T. (2004). André Leroi-Gourhan and the Evolution of Writing. In F. Audouze, & N. Schlanger (Eds.), *Autour de l'homme: contexte et actualité d'André Leroi-Gourhan* (pp. 109-123). Éditions APDCA.
- Ingold, T. (2013). *Making: Anthropology, Archaeology, Art and Architecture*. Routledge.
- Ingold, T. (2014). A Life in Books. *Journal of the Royal Anthropological Institute*, 20, 157-159. <https://doi.org/10.1111/1467-9655.12088>
- Lafitte, J. (1932). *Réflexions sur la science des machines* [Thoughts on machine science]. Bloud & Gay.
- Leroi-Gourhan, A. (1952). Homo faber... Homo sapiens. *Revue de Synthèse*, 30, 79-102.
- Leroi-Gourhan, A. (1953). Origine et diffusion de la connaissance scientifique [The Origin and Dissemination of Scientific Knowledge]. *Structure et evolution des techniques (S.E.T.)*, 4(33-34), 1-15.
- Leroi-Gourhan, A. (1964). *Le geste et la parole. Vol. 1: Technique et langage* [Gesture and Speech. Vol. 1: Technics and Language]. Albin Michel.
- Leroi-Gourhan, A. (1965). *Le geste et la parole. Vol. 2: La Mémoire et les Rythmes* [Gesture and Speech. Vol. 2: Memory and Rhythms]. Albin Michel.
- Leroi-Gourhan, A. (1993). *Gesture and Speech* [A. Berger, Trans.]. MIT Press.
- Leroi-Gourhan, A. (2024). *André Leroi-Gourhan on Technology. A Selection of Writings from the 1930s to the 1960s*, (N. Schlanger, Ed. & Intr.; translated by N. F. Schott, Trans.). Bard Graduate Center.
- Sigaut, F. (1994). Technology. In T. Ingold (Ed.), *Companion Encyclopaedia to Anthropology* (pp. 420-459). Routledge. <https://doi.org/10.4324/9781003209263>



Schlanger, N. (2023). *L'invention de la technologie. Une histoire intellectuelle avec André Leroi-Gourhan* [The Invention of Technology. An Intellectual History with André Leroi-Gourhan]. Presses universitaires de France.

Wiener, N. (1948). *Cybernetics: Or Control and Communication in the Animal and the Machine*. J. Wiley & Sons; Hermann & Cie.



<https://doi.org/10.48417/technolang.2024.02.09>

Research article

Leroi-Gourhan and the Object of Technology

Xavier Guchet (✉)

University of Technology of Compiègne, Rue du docteur Schweitzer CS 60319, Compiègne
60203 Hauts-de-France, France

xavier.guchet@utc.fr

Abstract

The emphasis on technical artifacts is a hallmark of contemporary philosophy of technology. How can Leroi-Gourhan's conceptualization of the technical object enrich current discussions among philosophers of technology? This article aims not to exhaustively address this question but to briefly outline how Leroi-Gourhan, as an ethnologist, reconfigures the concept of the technical object inherited from ethnology.

The article begins by presenting Leroi-Gourhan's ambition to revisit the central question of ethnology: what is the origin of the division of the human mass into distinct ethnic units called "peoples", distributed across the globe? According to Leroi-Gourhan, ethnology did not divide humanity at its natural junctures, leading to inaccurate historical conclusions. For him, "peoples" are not fixed and uniform entities defined by constant, specific characteristics. Instead, they arise from the temporary convergence of traits, such as language and technology, which have their own independent existence. These traits may come together at a certain point, but beyond that, they diverge. Ethnology should focus on these traits, not on the "peoples." In particular, technology serves as a reliable indicator of how the human mass has been divided and dispersed across space and time. However, to draw solid conclusions on this matter, it is essential to approach the extensive technical documentation with a rigorous method of classification and analysis. The article examines this method, leading Leroi-Gourhan to redefine the very concept of the technical object. The article highlights Leroi-Gourhan's focus on the concepts of "fact" and "tendency" in his analysis of technical objects. These objects are viewed both as solutions to general human challenges in transforming matter (representing "tendencies") and as culturally significant items with varying "degrees of fact." Thus, Leroi-Gourhan assigned a dual nature to technical objects, but in an interestingly different way than those analytical philosophers who have been discussing the dual nature of artefacts.

Keywords: Ethnology; Fact; Technology; Tendency

Citation: Guchet, X. (2024). Leroi-Gourhan and the object of technology. *Technology and Language*, 5(2), 116-124. <https://doi.org/10.48417/technolang.2024.02.09>



This work is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/)



УДК 39:62

<https://doi.org/10.48417/technolang.2024.02.09>

Научная статья

Леруа-Гуран и объект технологии

Ксавье Гуше (✉)

Технологический университет Компьена, ул. доктора Швейцера CS 60319, Компьень 60203 О-де-Франс, Франция

xavier.guchet@utc.fr

Аннотация

Акцент на технических артефактах является отличительной чертой современной философии технологий. Как концептуализация технического объекта, предложенная Леруа-Гураном, может обогатить текущие дискуссии среди философов техники? Целью данной статьи является не исчерпывающее рассмотрение этого вопроса, а краткое изложение того, как Леруа-Гуран как этнолог переконфигурирует концепцию технического объекта, унаследованную от этнологии. Статья начинается с представления амбиций Леруа-Гурана вернуться к центральному вопросу этнологии: каково происхождение разделения человеческой массы на отдельные этнические единицы, называемые “народами”, распространенные по всему миру? По мнению Леруа-Гурана, этнология не разделила человечество на его естественных этапах, что привело к неточным историческим выводам. Для него “народы” не являются фиксированными и однородными образованиями, определяемыми постоянными специфическими характеристиками. Вместо этого они возникают в результате временного сближения черт, таких как язык и технология, которые существуют самостоятельно. В какой-то момент эти черты могут сойтись вместе, но в дальнейшем они расходятся. Этнология должна сосредоточиться на этих чертах, а не на “народах”. В частности, технологии служат надежным индикатором того, как человеческая масса была разделена и рассеяна в пространстве и времени. Однако, чтобы сделать убедительные выводы по этому вопросу, необходимо подойти к обширной технической документации со строгим методом классификации и анализа. В статье рассматривается этот метод, что привело Леруа-Гурана к переопределению самого понятия технического объекта. В статье подчеркивается внимание Леруа-Гурана к понятиям “факт” и “тенденция” в его анализе технических объектов. Эти объекты рассматриваются как решения общих человеческих задач по преобразованию материи (“тенденции”), так и как культурно значимые предметы с различной “степенью фактичности”. Таким образом, Леруа-Гуран приписывал техническим объектам двойственную природу, но совершенно иначе, чем те философы-аналитики, которые обсуждали двойственную природу артефактов.

Ключевые слова: Этнология; Факт; Технология; Тенденция

Для цитирования: Guchet, X. Leroi-Gourhan and the object of technology // Technology and Language. 2024. № 5(2). P. 116-124. <https://doi.org/10.48417/technolang.2024.02.09>



This work is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/)



PUTTING AN END TO THE CONCEPT OF “PEOPLE”

André Leroi-Gourhan (1911-1986) was primarily an ethnologist and an archaeologist, and he approached many fields of study, including technology, from this perspective. Ethnology focuses on human diversity and the origins of this diversity. Although humanity is one, it is also a very heterogeneous mass spread across the globe. Where does this heterogeneity come from? How has humanity fragmented into distinct ethnic groups across space and time? According to Leroi-Gourhan (1945), ethnology has traditionally answered these questions by uncritically accepting the division of humanity into “peoples” that are defined by unequivocal characteristics: a physical type, a language, a social and political organization, and specific cultural traits (p. 306). Social and political organization is most often considered primary. For example, we start by defining the “Chinese people” within a given territory as a political unit. From this initial delineation, the anthropologist will establish the physical type of the “Chinese,” the linguist will study the language or languages spoken by these people, the sociologist will examine their social and political organizations, and the historian of religions will explore their myths and beliefs, and so on. Thus, the existence of well-defined peoples in both space and time is an unquestioned reference point for the ethnologist.

This produces certain methodological biases, such as explaining all observable cultural phenomena (including technology) by referencing the “people” to whom these traits are attributed. Thus, when ethnologists study technical objects, they immediately draw conclusions about the people who created them, much like tracing an effect back to its cause. For instance, if two similar objects are found in distant locations, the ethnologist concludes that the respective peoples were historically connected, or that something from one culture (an object or an idea) reached the other through contact, borrowing, or diffusion. The “people,” assumed to be a primary, consistent, and unequivocal reality, are regarded as the actual cause of all ethnological phenomena.

Leroi-Gourhan challenges this assumption that humanity is naturally divided into well-defined ethnic units or “peoples” with fixed characteristics. He defends two major theses: first, the concept of a perfectly defined “people” cannot be a valid starting point for ethnology; second, we do not know the laws by which humanity has been distributed and divided into distinct groups across the globe: the general laws governing ethnic phenomena are still largely unknown. Therefore, the ethnologist's task is to use rigorously controlled archaeological documentation to uncover the natural divisions in this process, akin to Plato's ideal butcher finding the natural joints. The general history of humanity is not the history of “peoples” and their movements: ethnology must free itself from the very notion of “peoples” and adopt a completely different way of dividing up the whole of humanity.

Instead of viewing ethnic groups as having stable properties within well-defined territories, Leroi-Gourhan considers physical types, techniques, and social-religious structures as elements that each have a life of its own; some of these elements predate the group, others outlast it. Societies, he argues, are “temporary colonies” or constellations of elements, rather than cohesive units with consistent and stable characteristics over time (Leroi-Gourhan, 1945, pp. 403-404). What ethnologists refer to as a “people” is actually the result of the local and temporary convergence of some such elements. Beyond a



certain zone where these elements actually converge, they diverge and the very notion of a “people” is blurred (Leroi-Gourhan, 1945, p. 306). The ethnologist's role should not be to artificially construct the “average personality” of a supposedly well-defined ethnic group by combining specific characteristics (such as a physical type or a language). Instead, the ethnologist should trace these characters, considering them as distinct entities, each with its own dynamic. This approach may show that, indeed, these themes can locally and temporarily converge and align at specific points in space and time. However, beyond these points, the convergence dissolves.

As a consequence, when studying technology from an ethnological perspective, it is essential to separate humans from their products and to stop assuming that the spread of these products is dependent on the movement of people. In *L’homme et la matière* (Leroi-Gourhan, 1943) and his ethnology thesis on the archaeology of the North Pacific (Leroi-Gourhan, 1946), Leroi-Gourhan observes that humans have travelled much less frequently and over shorter distances than their products (Leroi-Gourhan, 1946, p. 7). There is abundant evidence of objects being transported independently of human migration. Humans and their products should not be seen as an inseparable unit.

TECHNOLOGY AND THE CLASSIFICATION OF TECHNICAL OBJECTS

To break away from common-sense divisions (“people”) and uncover the natural divisions of humanity, Leroi-Gourhan focused on technology, particularly in the extensive North Pacific region. He believed that it was crucial to reverse the usual ethnological assumption that objects are explained by the movements and activities of distinct ethnic groups. Instead, he argued that the formation of ethnic groups at specific points in space and time is not so much exposed but, much rather, caused by the circulation of objects.

Thus, the mistake of ethnology was to draw historical conclusions about the “peoples” who created and circulated objects, based solely on the dispersion of these objects in a given area. Leroi-Gourhan points out that conclusions about migrations have often been drawn from overall similarities in objects, even though these similarities rarely show indisputable identical traits. In other words, ethnologists have been too quick to make assumptions, relying on purely morphological aspects (i.e., the shape of objects) to derive historical insights without thoroughly verifying these conclusions through the establishment of proper technological connections (i.e., relating to the objects’ technical functions). Leroi-Gourhan argues that similar forms found in different locations do not necessarily indicate contact between peoples. It could simply be that people in different places faced similar challenges with only a limited number of solutions available. While two similar technical features might result from a historical point of contact, they could also be explained in purely technological terms—namely, the limited ways humans could interact with materials at a given time and place.

Technology is the science of how humans interact with materials. Like ethnologists, “technologists” group and classify technical objects, but they do so in a different way, namely based on the technical operations applied to materials. These are characterized



not by form but by the type of treatment they can undergo (solid, semi-solid, flexible, supple, fluid, etc.). Without the rigor of such technological classification, the historical conclusions of ethnologists are incredibly fragile. Therefore, every technical fact must be viewed from two distinct perspectives: history and technology. The ethnologist's error has been to conflate these perspectives, or rather, to neglect the technological perspective in the rush to draw historical conclusions.

Leroi-Gourhan acknowledges that ethnologists have also applied a logical framework to technical facts. Typically, they have reasoned similarly to philologists by analyzing each technical characteristic comparatively, viewing them as variations from an original motif—just as philologists analyze textual elements. Leroi-Gourhan argues that this approach assumes that there already exists a phonetics and syntax of objects, allowing us to analyze an axe or a hammer, as if they were a part of a sentence. However we do not yet understand the laws that govern the grouping of technical objects, the rules by which they form coherent sequences and inseparable sets, and the regularities by which these sequences and sets transform into others. In other words, we lack a syntax for the material world of humans (Leroi-Gourhan, 1946, p. 9). The goal of technology is to construct this syntax, but it cannot simply adopt the philologist's comparative method. It needs its own unique analytical approach.

These considerations prevented Leroi-Gourhan from relying on the form and similarity of objects to draw general conclusions. Instead, he suggests viewing objects as evidence of an interaction between humans and materials. On one side is the biomechanics of human action, and on the other are the material properties. These two factors define the limited range of ways humans can work with materials, regardless of where in the world the work is done.

All tools, despite their great diversity, fall into this typology that reflects the set of constraints on human action. The key to the classification tools lies not in the tool alone, but in the combination of the tool and the gesture that employs it (Leroi-Gourhan, 1964; 1965; 1993). Leroi-Gourhan is explicit on this point: the technologist must move beyond focusing solely on the object and instead see it as the result of a dialogue between human action and material properties, offering a richer understanding than a purely morphological classification. This means applying the same approach to “objects” that was applied to “peoples.” It means moving beyond immediate perception and reclassifying technical facts in a way that transcends common sense. The aim of technologists is to uncover the natural divisions by which these facts are distributed across space and time.

FACT AND TENDENCY

Technological documentation can be categorized in various ways and at different levels. At the most general level, an object represents a specific solution to a particular problem, chosen from a very limited set of possibilities. However, a technical artifact is more than just a functional solution; it also reflects the specific characteristics of its environment (such as the availability or scarcity of certain materials) and the unique cultural traits of the human group that uses it. On the most specific level, an object may



appear unique and specific to a particular group. On the most general level, it appears as a functional response to a problem, similar to how other objects from other groups address the same problem. As Leroi-Gourhan puts it, technical objects are both unique, context-specific “facts” and the materialization of “tendencies,” which are the general functional principles governing the interactions between humans and materials. Technical objects can be classified according to an evolutionary logic, but they can also be categorized from a historical perspective. How can we ensure the alignment between a technological series and a chronological one? How can the same fact be both a historical singularity and part of a technological classification independent of history? Is it possible to unify these two perspectives?

Leroi-Gourhan employs very Kantian terms to discuss this hypothetical concordance between these two viewpoints, which he describes as incommensurable: the unity of the two perspectives on the technical fact is not itself a fact and should not be sought in a theory of a higher level; “This unity is transcendental.” (Leroi-Gourhan, 1945, 58) This means, to continue using Kantian words, that the duality of the “technical fact” and the “technical tendency” is a condition of possibility for knowledge in ethnology. Technical objects fall under two kinds of research, one historical and the other reflexive: this is not a paradox, but the condition of possibility for ethnology as a science. Based solely on technical objects considered in their inexhaustible and bewildering cultural diversity, ethnologists can do nothing but indulge in speculative constructions based on apparent formal similarities among these objects. Very quickly, they realize that these constructions are artificial and that they can develop opposing theories from the same documents. They then unavoidably fall into skepticism.

Ethnologists made two mistakes: firstly, they believed they could classify technical objects on a purely morphological criterion which is very imprecise; secondly, they drew erroneous general historical conclusions from this morphological classification, without any caution. When Leroi-Gourhan writes that *Évolution et techniques* (Evolution and Techniques) should be read as the critique of a great book whose author has yet to be born (Leroi-Gourhan, 1943, p. 37), the word “critique” should be understood in its Kantian meaning: it is indeed about defining the transcendental conditions for valid knowledge in ethnology. What is this “great book”? The systematic collection of all knowledge about humans and the products of their activity, based on complete and perfectly controlled archaeological documentation, on a global scale. In his thesis on the archaeology of the North Pacific, Leroi-Gourhan imposed severe constraints in order to achieve maximum security in the treatment of documentation and in the conclusions he allows himself to draw. In light of these constraints, one can’t help but imagine that the completion of this “great book” is not for tomorrow...

The Kantian reference leads to a remark regarding the “tendency.” It is known that this term owes to an inspiration by Henri Bergson. Should we understand this concept of “tendency” as a sort of vital impulse, a force acting within the human milieu? In short, should we see it as an eruption of metaphysics in ethnology? The answer is no: the tendency is, as Leroi-Gourhan argues, “a convenience cut that our logic introduces into the activity of men” (Leroi-Gourhan, 1943, pp. 34-35). It is not a metaphysical concept,



but a transcendental one. Admittedly, transcendental concepts, in Kant, are certainly not classificatory concepts, as tendency is. The connection with Kant stops there.

How does this classification of technological documentation, using the concepts of “fact” and “tendency”, work in concrete terms? The example of harpoons illustrates this particularly well. In this regard, Leroi-Gourhan (1946) states his intention to establish a framework that is broad enough to encompass all possible variants of harpoons, and precise enough to capture the finest local nuances (p. 326). This is facilitated by organizing the documentation (which is vast) within the two poles of “tendency” and “fact” (for a detailed analysis, see Leroi-Gourhan, 1943, pp. 30-35).

The tendency embodied by the harpoon is “to kill a marine animal.” Countless human groups have hunted marine animals across the globe. Since the word “harpoon” conveys nothing more than its function (“to kill a marine animal”), with no added ethnic specification, Leroi-Gourhan identifies this first degree of the fact with the tendency itself. It is found widely throughout the world. From there on, Leroi-Gourhan classifies the documents according to the “degrees of the fact,” and in doing so, he individualizes them more and more. Comparing technical objects from different place around the world is only beneficial “if a list is created for each object. This list should begin with the most prominent feature, move on to the more apparent characteristics, and finally include the most specific details, such as the symbolic meaning of the tool.” (Leroi-Gourhan, 1943, p. 30) In the case of harpoons, a second degree of the fact specifies them as “harpoons with bone tips,” found in the Pacific Ocean – still very broad, but no longer the entire world since in other regions of the world, harpoons do not have bone tips. This is already an ethnically determined specification, a first level of individualization. At the third degree of the fact, there are “harpoons with bone tips and a float,” among the Eskimos – a new level of individualization, a further narrowing of the geographical area concerned. At a fourth level, there are “harpoons with bone tips and a bladder float,” among the Eskimos located in Alaska (unlike the Eskimos who live elsewhere). The degrees of the fact can be multiplied as necessary to achieve the maximum level of individualization of the technical fact.

Classifying technological documentation in this way makes it possible to carefully control historical conclusions. Indeed, often the general constructions of ethnologists consist of linking together two facts at the second or third degree (or even higher) by a fact considered at the first degree, in order to conclude that the two corresponding “peoples” have had historical contact. Let us reiterate that the first degree of the fact (for example, the harpoon without further specification) is the materialization of a technical tendency, that is, the solution to a problem to which many human groups have independently responded, drawing from a limited stock of possibilities. Therefore, referring to the presence of harpoons considered at the first level of the fact (that is, without ethnic peculiarities) to prove historical connections between distant human groups is misleading. The similarity in the shapes of countless harpoons found worldwide does not necessarily indicate historical connections between distant ethnic groups. Instead, these similarities highlight the universal nature of technical functions and the constraints these functions impose on tool design.



CONCLUSION: LEROI-GOURHAN AND THE PHILOSOPHY OF TECHNICAL ARTIFACTS

In conclusion, a technical object only holds scientific value for ethnology if it is precisely described and characterized. A technical object is meaningless without its accompanying documentation, making it both material and textual. The criteria for this documentation were established by Mauss (1947) and later adopted by Griaule and Leiris (1931/2015) during the Dakar-Djibouti mission of 1931-1933, with Leroi-Gourhan further building on this foundation.

This framework is very different from that proposed by contemporary philosophers analyzing technical artifacts. Analytical philosophers view a technical artifact as the interaction between material and intention. Leroi-Gourhan would have considered these concepts of material and intention overly general. According to him, material in itself does not exist; instead, there are specific materials with distinct properties that either assist or resist the technician. These materials can be stable solids, fibrous, plastic, semi-plastic, flexible, or fluids. Depending on how they are treated, their properties can change. For instance, bronze is a stable solid when chiselled, plastic under the hammer, and fluid when cast in a mold. Therefore, a sculptor might work with the same material but encounter different properties, requiring different techniques. Intention must also be broken down according to its tendencies and degrees, encompassing both the general technical function and the desire to impart a unique ethnic quality to the tool.

Analytical philosophers tend to analyze a technical artifact in isolation, whereas Leroi-Gourhan believed that an isolated technical object cannot be conclusively analyzed scientifically; it must be part of a functional series. An isolated technical object, even with detailed documentation, does not qualify as a scientific document; only a series can establish solid knowledge. Between the isolated object and the series lies the collection. Properly cataloged and documented objects form collections, which can then be classified into functional series based on the duality of “tendency” and “fact.” Consequently, Leroi-Gourhan's technical object is less like the artifact analyzed by today's philosophers and more akin to the biological samples that biobankers compile into collections of scientific relevance for biomedical research.

REFERENCES

- Griaule, M., & Leiris, M. (2015). Instructions sommaires pour les collecteurs d'objets ethnographiques. In É. Jolly, & M. Lemaire (Eds.), *Cahier Dakar-Djibouti* (pp.169-207). Éditions Les Cahiers. (Original work published 1931)
- Leroi-Gourhan, A. (1943). *L'Homme et la matière. Vol. 1 of Évolution et techniques*. Albin Michel.
- Leroi-Gourhan, A. (1945). *Milieu et techniques. Vol. 2 of Évolution et techniques*. Albin Michel.
- Leroi-Gourhan, A. (1946). *Archéologie du Pacifique nord. Matériaux pour l'étude des relations entre les peuples riverains d'Asie et d'Amérique*. Travaux et mémoires de l'Institut d'Ethnologie de l'Université de Paris, XLVII.
- Leroi-Gourhan, A. (1964). *Le Geste et la Parole. 1 Technique et langage*. Albin Michel.



- Leroi-Gourhan, A. (1965). *Le Geste et la Parole. 2 La mémoire et les rythmes*. Albin Michel.
- Leroi-Gourhan, A. (1993). *Gesture and Speech*, [A. Berger, Trans.] MIT Press.
- Mauss M. (1947). *Manuel d'ethnographie*. Payot

СВЕДЕНИЯ ОБ АВТОРЕ / THE AUTHOR

Ксавье Гуше, xavier.guchet@utc.fr
ORCID 0000-0003-4912-4912

Xavier Guchet, xavier.guchet@utc.fr,
ORCID 0000-0003-4912-4912

Статья поступила 3 марта 2024
одобрена после рецензирования 10 июня 2024
принята к публикации 17 июня 2024

Received: 3 March 2024
Revised: 10 June 2024
Accepted: 17 June 2024



<https://doi.org/10.48417/technolang.2024.02.10>

Research article

“Une véritable syntaxe”. Some Notes on Leroi-Gourhan and Structural Linguistics

Simone Aurora (✉)

Università degli studi di Padova, Padova, Piazza Capitaniato 3, 35139, Italy

simone.aurora@unipd.it

Abstract

In what follows, I will try to offer a brief overview of the relationships between some of Leroi-Gourhan's anthropological insights on technology and some of the fundamental theoretical claims that form the general framework of structural linguistics. This hermeneutical movement runs an evident risk, which needs to be addressed and overcome: the risk of including Leroi-Gourhan's works in the wide range of the structuralist *corpus*. For this reason, in the introduction I clarify what I mean by “structuralism” so that, in the subsequent sections, I can try to show the epistemological relationship between Leroi-Gourhan's ethnology and the “structuralist turn,” as described in the introduction. To this end, I will point out the possible theoretical influence exerted by structural linguistics, and especially by the structural phonology developed within the Prague linguistic circle, on Leroi-Gourhan's conceptual toolbox. More specifically, in the paper, I will focus on some passages of *La geste et la parole* (1964), which I will consider in connection with two more minor and older texts, namely *Origine et diffusion de la connaissance scientifique* (1953) and *L'homme et la nature*, an article published in 1936 in the *Encyclopédie française*.

Keywords: Structuralism, Linguistics, Phonology, Syntax, Technological system

Citation: Aurora, S., (2024). “Une véritable syntaxe”. Some Notes on Leroi-Gourhan and Structural Linguistics. *Technology and Language*, 5(2), 125-135. <https://doi.org/10.48417/technolang.2024.02.10>



This work is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/)



УДК 81-116: 39

<https://doi.org/10.48417/technolang.2024.02.10>

Научная статья

“Une Véritable Syntaxe”. Некоторые примечания о Леруа-Гуране и структурной лингвистике

Симон Аврора (✉)

Университет Падуи, Падуя, Площадь Капитаниато 3, 35139, Италия

simone.aurora@unipd.it

Аннотация

Здесь я попытаюсь предложить краткий обзор взаимосвязей между некоторыми антропологическими взглядами Леруа-Гурана на технологию и некоторыми фундаментальными теоретическими утверждениями, которые формируют общую основу структурной лингвистики. Это герменевтическое движение сталкивается с очевидным риском, который необходимо учитывать и преодолевать: риск включения работ Леруа-Гурана в широкий спектр структуралистских изданий. По этой причине во введении я разъясняю, что я подразумеваю под “структурализмом”, чтобы в последующих разделах попытаться показать эпистемологическую связь между этнологией Леруа-Гурана и “структуралистским поворотом”, как описано во введении. С этой целью я укажу на возможное теоретическое влияние структурной лингвистики, и особенно структурной фонологии, разработанной в рамках пражского лингвистического кружка, на концептуальный инструментарий Леруа-Гурана. Более конкретно, в этой статье я сосредоточусь на некоторых отрывках из “La geste et la parole” (1964), которые я рассмотрю в связи с двумя меньшими и более старыми текстами, а именно “Origine et diffusion de la connaissance scientifique” (1953) и “L’homme et la nature”, статьей, опубликованной в 1936 году во Французской энциклопедии.

Ключевые слова: Структурализм, Лингвистика, Фонология, Синтаксис, Технологическая система

Для цитирования: Aurora, S. “Une véritable syntaxe”. Some Notes on Leroi-Gourhan and Structural Linguistics // Technology and Language. 2024. № 5(2). P. 125-135.
<https://doi.org/10.48417/technolang.2024.02.10>



This work is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/)



INTRODUCTION: LEROI-GOURHAN AND STRUCTURALISM

In what follows, I will try to offer a brief overview of the relationships between some of André Leroi-Gourhan's anthropological insights on technology and some of the fundamental theoretical claims that form the general framework of structural linguistics. To this end, I will focus on some passages of *La geste et la parole* (1964), which I will consider in connection with two "minor" and older texts, namely *Origine et diffusion de la connaissance scientifique* (1953) and *L'homme et la nature*, an article published in 1936 in the *Encyclopédie française*.

This hermeneutical movement runs an evident risk, which needs to be addressed and overcome: the risk of including Leroi-Gourhan's works in the wide range of the structuralist *corpus*.

The term "structuralism" usually refers to a well-defined research trend in the area of the social sciences, especially popular within the French culture during the 1950s and 1960s, whose "pilot science" – to borrow an expression from François Dosse's (1991/2012) *History of Structuralism* – was represented by Saussurian linguistics and, above all, by the structural linguistics developed by the schools of Prague, Moscow and Copenhagen. In fact, however, the history of structuralism begins much earlier and is not at all limited to the field of linguistics and social sciences and still less to a particular period of the French culture. On the contrary, it constitutes a proper and very broad epistemological paradigm, as recent works have tried to show. As Ernst Cassirer wrote in an essay published in 1945 with the title *Structuralism in Modern Linguistics*, it is thus possible to affirm that "[...] structuralism is no isolated phenomenon; it is, rather, the expression of a *general tendency of thought* that, in these last decades, has become more and more prominent in almost all fields of scientific research" (Cassirer, 1945, p. 120).

The term "structuralism" was originally coined in the field of psychology, in order to define the psychological approach introduced by Edward Bradford Titchener, the most important and influent student of the German psychologist Wilhelm Wundt – who is usually regarded as the founder of modern scientific psychology – in a famous article published in 1898 and entitled *The Postulates of a Structural Psychology* (Titchener, 1898). However, although it is, without any doubt, Titchener who explicitly introduced the term "structural" within the field of scientific research, his scientific approach cannot be defined as a form of structuralism in the strict sense of the term. Indeed, the proper structuralist currents that emerge in the years between the late nineteenth and the early twentieth centuries are instead, most notably, the following:

1) In the field of psychology, we can consider the descriptive psychology of Franz Brentano and the mereological analyses developed in his "school" and, especially, the inquiries on perceptual experience developed by Gestalt psychology as initiated by Carl Stumpf's students Max Wertheimer, Kurt Koffka, and Wolfgang Köhler. "[T]he most spectacular form of psychological structuralism," Piaget notes, "was undoubtedly the theory of Gestalten" (Piaget, 1968, p. 53). Gestalt psychology – whose foundational text



is generally considered to be von Ehrenfels's (1890) essay *Über Gestaltqualitäten* – developed, like Brentano's psychology, in direct opposition to Titchener's structuralism, and especially in opposition to its atomism and positivistic attitude. According to Mitchell Ash, “[t]he Gestalt theorists opposed the assumption that sensory ‘elements’ are the basic constituents of mental life then characteristic of psychological theory and research in Germany and elsewhere” (Ash, 1995, p. IX). Indeed, the fundamental idea of Gestalt psychology is represented by the mereological thesis, according to which a whole differs from the mere sum of its parts and it is therefore impossible to investigate the structure of a complex psychological fact on the basis of its ground elements. This is due to the fact that it is not possible to consider these elements separately, since they can exist only within the system of relations which connects them in a law-governed whole, that is, a structure. As Daniel Lagache puts it, “*Gestalt* theory rejects the idea of simple elements, the composition of which would explain the whole. Whole and parts are given at the same time. The knowledge of the whole cannot be inferred from the knowledge of the parts. The latter cannot be complete without reference to the whole” (Lagache, 1962, p. 81).

2) In the mathematical domain, we can count Evariste Galois's “group theory” as a forerunner of mathematical structuralism (Piaget, 1968), although the term “structure” was actually introduced only with the development of the calculus of variations and of topology. Structuralist approaches can also be found in the works of David Hilbert and, above all, in the program of the so-called “Bourbaki group.” Indeed, in a 1948 essay, *L'architecture des mathématiques*, the Bourbakists propose the following definition of the concept of structure: “The common character of the different concepts designated by this generic name, is that they can be applied to sets of elements whose nature has not been specified; to define a structure.” They write further: “one takes as given one or several relations, into which these elements enter [...] then one postulates that the given relation, or relations, satisfy certain conditions. To set up the axiomatic theory of a given structure [thus] amounts to the deduction of the logical consequences of the axioms of the structure, excluding every other hypothesis on the elements under considerations” (Bourbaki, 1950, pp. 225-226). As Robert Hannah (2010) summarizes,

[m]athematical Structuralism [...] says that mathematical entities (e.g., numbers or sets) are not ontologically autonomous or substantially independent objects, but instead are, essentially, positions or roles in a mathematical structure, where a mathematical structure is a complete set of formal relations and operations that defines a mathematical system. What counts as an individual object of the system is thereby uniquely determined by the system as a whole. (p. 158)

3) Finally, with reference to linguistics, structuralism emerged with the general theory of language established by Saussure and then with the structural linguistics elaborated by the schools of Prague and Copenhagen. Indeed, in the field of linguistics, the term first appears with reference to phonology, more precisely, in the very well-known theses of the Prague Linguistic Circle, presented at the first Congress of Slavists held in



Prague in 1929. As the prominent linguist Émile Benveniste sums up, the Prague-school phonologists “advocated ‘a method suitable for permitting the discovery of the laws of structure of linguistic systems and their evolution.’” Moreover:

the notion of ‘structure’ was closely linked with that of ‘relationship’ within the system: ‘The sensory content of phonological elements is less essential than their reciprocal relationships within the system (*structural principle of the phonological system*).’ Hence this rule of method: ‘The phonological system must be characterized [...] by an obligatory specification of the relationships existing among the said phonemes; that is, by tracing the structural scheme of the language being considered’. These principles are applicable to all parts of the language. (Benveniste, 1971, p. 81)

General reconstructions of structuralism usually neglect many of these currents – for instance in the Brentanian school or in early phenomenology – or underestimate them – as in the case of structuralist tendencies in the mathematical domain. On the contrary, when considered in its complexity, the “structuralist turn” allows to better understand and clarify several tendencies in the social and human sciences of the last century, at least until the 80s. This contribution aims to show that some of Leroi-Gourhan’s theoretical insights on the relationship between language and technology show a deep solidarity with the “structuralist turn” and, especially, with some fundamental features of structural linguistics. In this sense, it is not by chance that Leroi-Gourhan’s anthropology and ethnology show an often-underestimated epistemic relationship with Lévi-Strauss’s structural inquiries (see Collins, 2021):

In rereading his work, I am struck by the fact that, working in different domains, he and I were trying to do basically the same thing [...] The guiding idea of his work was always to study the relations between things rather than the things themselves, to try to reduce the chaotic diversity of empirical facts to invariant relationships. (Lévi-Strauss, 1988, pp. 203-204)

In the following sections, I will try to show the relationship between Leroi-Gourhan’s ethnology and structuralism by highlighting the theoretical influence exerted by (structural) linguistics on Leroi-Gourhan’s conceptual toolbox.

***L’HOMME ET LA NATURE* (1936): A STRUCTURAL PRINCIPLE OF TECHNOLOGICAL SYSTEMS?**

Leroi-Gourhan was more than acquainted with linguistic studies since, as he himself acknowledges his “*première formation ayant été de linguistique et d’anthropologie anatomique*” (quoted in Schlanger, 2023, p. 26) and, especially in the



30s, his scientific training took the form of a “*quadruple formation en langues, linguistique, ethnologie, anthropologie biologique*” (de Beaune, 2011, p. 200).

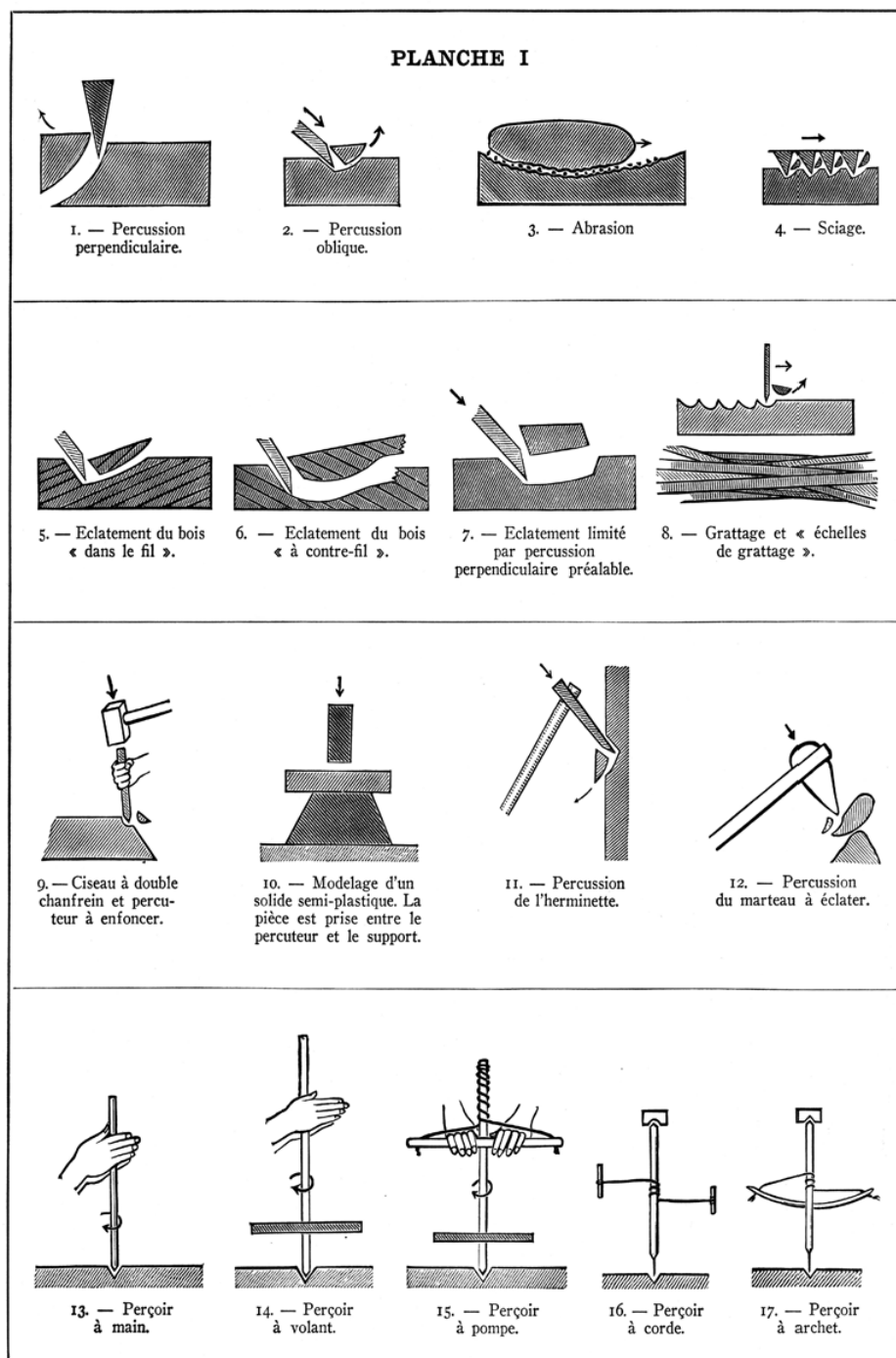
Moreover, the linguistic apprenticeship of Leroi-Gourhan took place in years of great importance for the development and establishment of a new paradigm in the sciences of language, namely structural-functional linguistics: In 1916, Charles Bally and Albert Sechehaye edited the *Cours de linguistique générale* from notes on lectures given by Ferdinand de Saussure at the University of Geneva between 1906 and 1911; in 1929, the newly born *Cercle linguistique de Prague* presented its theoretical manifesto, in French, at the First Congress of Slavic Philologist, the famous “Thésés,” which were bound to chart the general direction of structuralism for the subsequent years; finally, in 1933, an important issue of the *Journal de psychologie normale et pathologique* devoted to the psychology of language was published, featuring contributions by some of the most important linguists and philosophers of the time, with a strong presence of structural linguists and phonologists, such as Sechehaye, Vendryes, Brøndal, Trubetzkoy, Sapir, Bally. Among these potential influences, I believe the scientific work of the Prague linguistic circle to be certainly the most important, even in consideration of its reception in French anthropology and ethnology in the 30’s and the 40’s (see for instance Lévi-Strauss, 1945).

In his 1936 article for the *Encyclopédie française*, Leroi-Gourhan (1936) proposes a “mechanically logical classification” of general technics (p. 206), namely a critical inventory showing that technics must be described as the result of a combinatorial system of specific material options and potential actions (fig. 1). In a similar way, in the famous “theses” the Prague linguists claim that “acoustic-motor representations,” namely objective-material sounds (the *phones*), “are elements of a language system only insofar as they serve to differentiate meanings”, that is only insofar as they are *phonemes*; on this basis, they establish “*the structural principle of the phonological system*,” according to which “the sensory contents of [...] phonological elements are less essential than their interrelations in the system” (Steiner, 2014, p. 8). On these grounds, the Prague phonologists identify three general fundamental tasks of phonology, namely:

1. [...] to describe the phonological system, that is to establish the set of simplest acoustic-motor representations which create meanings in a given language (*phonemes*) [...]
2. [to] determine *the combinations of phonemes realized* in a given language compared to all the theoretically possible combinations of these phonemes, the variations in the sequence of their grouping, and the scope of these combinations [...]
3. [to] determine the degree of utilization and the frequency of realization of the given phonemes and the combinations of phonemes of different scope.” (Steiner, 2014, pp. 8-9)



SECTION A - FORMES ÉLÉMENTAIRES DE L'ACTIVITÉ HUMAINE



7'10 - 8

Figure 1. A mechanically logical classification of elementary forms of human activity (Leroi-Gourhan, 1936, p. 8).



In my opinion, the descriptive framework that one can find in *L'homme et la nature* is epistemically equivalent to the methodological principles outlined in the Prague manifesto and can represent a theoretically coherent derivation thereof. This means that what is claimed in the *Thèses* with reference to language has been “transferred” by Leroi-Gourhan into the domain of technology. Understood in these terms, one can easily rephrase the Prague phonological principles in technological terms. Accordingly, the *structural principle of the technological system* would claim that *the sensory contents of technological elements are less essential than their interrelations in the system* and the general tasks of a philosophy of technology would thus be the following: 1. To describe the technological system, that is to establish the set of simplest relations between a given material (stable solids, semi-plastic solids, plastic solids, flexible solids, fluids) and a given action (perpendicular, oblique or circular percussion) which create tools in a given society (we could call them *technemes* or, as Leroi-Gourhan writes, “elementary forms of human activities”) 2. To determine *the combinations of technemes realized* in a given society compared to all the theoretically possible combinations of these technemes, the variations in the sequence of their grouping, and the scope of these combinations 3. To determine the degree of utilization and the frequency of realization of the given technemes and the combinations of technemes of different scope.

ORIGINE ET DIFFUSIONE DE LA CONAISSANCE SCIENTIFIQUE **(1953): THE TECHNICAL PROCESS**

In the March 1952 lecture at the Maison des Sciences in Paris (*Origine et diffusion de la connaissance scientifique*), translated for the first time in English in this collection, Leroi-Gourhan aims “to explore [...] the paths taken by humankind from its origins to the point at which it entered the period of major scientific speculation, to see when the first concerns with rational research emerged in the history of human societies” (Leroi-Gourhan, 2024, p. 105). According to Leroi-Gourhan, a fundamental premise for such an investigation lies in the claim according to which one has to admit that “technical progress is really linked to [scientific] research” insofar as there is a “link between psychical reactions and technical behavior, between the latter and the manufacturing techniques specific to human beings, between manufacturing techniques and invention, and between invention and scientific speculation” (p. 105). The considerations offered by Leroi-Gourhan in this lecture are thus a coherent development, on a diachronic level, of the theoretical framework established in 1936 on a synchronic level, stating that technology is nothing but a systematic interrelation between materials and actions or *gestures*, as Leroi-Gourhan calls them in this lecture and, notably, in *La geste et la parole*. In this sense, the analysis of flintknapping by the Neanderthals mentioned in the lecture reveals the first precise picture we have in the history of humankind of a complex technical process, since “the systematic knapping of triangular flakes characteristic of the Mousterian represents a series of a dozen gestures following each other in an absolutely



rigorous order” (p. 107). According to Leroi-Gourhan, “this is the first evidence of technical intelligence”, which seems thus to coincide with the “setting up of [...] a complex arrangement of operations [*un dispositif opératoire complexe*]” and with a “standardised operational thinking [*une pensée opératoire très rationalisée*] involving series of gestures set in precise sequence” (p. 107 and p. 108). As Charles Lenay writes, in this sense “we may even speak of a sort of technical *syntax*, insofar as the fabrication of the tools proceeds by ordered sequences of operations, and a different arrangement would produce different products” (Lenay, 2018, p. 219). As it is well known, in *La geste et la parole* one can find explicit references to syntax and linguistic structures. However, this reference to syntax is, in my opinion, neither new nor metaphorical; on the contrary, as I’ve tried to suggest in this paper, this reference can be found in Leroi-Gourhan’s earlier texts, although in an implicit way, and shows a deep epistemological solidarity with some of the main theoretical principles of structural linguistics and phonology,¹ as the following concluding section will try to show.

LA GESTE ET LA PAROLE (1964): UNE VERITABLE SYNTAXE

In Leroi-Gourhan’s 1964 masterpiece, *La geste et la parole*, the term “syntax” occurs six times, while language plays of course a pivotal role throughout the whole text. Although, in my opinion, the references to syntax are always coherent with the theoretical framework previously outlined, there are two occurrences that seem to be particularly relevant for my argument. In the first of these references, Leroi-Gourhan clearly connects the idea of “syntax” to that of an “operating sequence,” which cannot but recall the notion of *dispositif opératoire* and of *pensée opératoire*, already encountered in the 1952 lecture:

We shall revert to the concept of operating sequences [...] but mention of it must be made here if we are to understand the link between technics and language. Techniques involve both gestures and tools, sequentially organized by means of a ,syntax‘ that imparts both fixity and flexibility to the series of operations involved. This operating syntax is suggested by the memory and comes into being as a product of the brain and the physical environment. If we pursue the parallel with language, we find a similar process taking place. On the basis of what we know of techniques from pebble culture to Acheulean industry, we could adopt the hypothesis of a language whose complexity and wealth of concepts corresponded

¹ It must be said that an eminent scholar like Nathan Schlanger totally disagrees with the hermeneutical hypothesis advocated in this contribution: “la notion de ,syntaxe‘ [...] donne une inflexion linguistique ,grammaticale‘ ou ,générative‘ qui, tout en étant de mode dans les sciences humaines des années 1960 et 1970, est en fait assez éphémère dans le bagage théorique de son auteur. Outre la section du Geste et la Parole intitulée ,Le langage des “préhominiens,” et mis à part ses interprétations de l’art pariétal paléolithique, cette conception est peu présente et non théorisée dans ses écrits antérieurs” (Schlanger, 2023, p. 542).



approximately to the level of those techniques. (Leroi-Gourhan, 1993, pp. 114-115)

In the second instance, Leroi-Gourhan explicitly poses an equivalence that appears to be of particular interest for the purposes of this reflection, namely the equivalence between word-tool and syntax-gesture. Moreover, he seems to understand tools and linguistic elements (together with rhythmic components) as parts of the same process:

[...] the purpose of verbal figures – words and syntax – is, like the purpose of tools and manual gestures, *their equivalents*, to provide an effective hold on the world of relationships and of matter [...] we see that tools, language, and rhythmic creation are *three contiguous aspects of one and the same process*. (Leroi-Gourhan, 1993, pp. 365-366, italics added)

As I have tried to suggest in this contribution by means of a brief exploration of some important texts, I believe that some of Leroi-Gourhan's theoretical insights on technology show a deep epistemological solidarity with the “structuralist turn,” as it has been described in the introduction, and especially with some of the fundamental tenets of structural linguistics and phonology. Against this background, I think that the 1952 lecture on “The origin and dissemination of scientific knowledge” plays a quite important and significant role.

REFERENCES

- Ash, M. G. (1995). *Gestalt Psychology in German Culture, 1890-1967. Holism and the Quest for Objectivity*. Cambridge University Press.
- Benveniste, E. (1971). *Problems in General Linguistics*. Miami University Press.
- Bourbaki, N. (1950). The Architecture of Mathematics. *The American Mathematical Monthly*, 57(4), 221-232. <https://doi.org/10.1080/00029890.1950.11999523>
- Cassirer, E. (1945). Structuralism in Modern Linguistics. *Word*, 1, 99-120. <https://doi.org/10.1080/00437956.1945.11659249>
- Collins, J. (2021). Parallel structures: André Leroi-Gourhan, Claude Lévi-Strauss, and the making of French structural anthropology. *History of the Human Sciences*, 34, 307-335. <https://doi.org/10.1177/095269512091153>
- De Beaune, S.A. (2011). La genèse de la technologie comparée chez André Leroi-Gourhan. Introduction à son article “L’homme et la nature” paru dans L’Encyclopédie française en 1936 [The Genesis of Comparative Technology in André Leroi-Gourhan. Introduction to his Article “Man and Nature” Published in L’Encyclopédie Française in 1936]. *Documents pour l’histoire des techniques*, 20, 197-223. <https://doi.org/10.4000/dht.1826>
- Dosse, F. (2012). *Histoire du structuralisme* [History of structuralism]. La Découverte. (Original work published 1991)
- Ehrenfels, von C. (1890). Über Gestaltqualitäten [About Gestalt Qualities]. *Vierteljahrsschrift für wissenschaftliche Philosophie*, 13, 249-292.



- Hannah, R. (2010). *Mathematical Truth Regained*. In M. Hartimo (Ed.), *Phenomenology and Mathematics* (pp. 147-181). Springer. https://doi.org/10.1007/978-90-481-3729-9_8
- Lagache, D. (1962). Structure en psychologie [Structure in Psychology]. In R. Bastide (Ed.), *Sens et usages du terme structure dans les sciences humaines et sociales* (pp. 81-82). Mouton.
- Lenay, C. (2018). Leroi-Gourhan: Technical Trends and Human Cognition. In S. Loeve, X. Guchet, & B. Bensaude Vincent (Eds.), *French Philosophy of Technology, Classical Readings and Contemporary Approaches* (pp. 209-226). Springer. https://doi.org/10.1007/978-3-319-89518-5_13
- Leroi-Gourhan, A. (1936). L'homme et la nature [Human and nature]. In L. Febvre (Ed.), *L'encyclopédie française: Vol. 7. L'espèce humaine* (pp. 3-13). Société des gestion de l'Encyclopédie française
- Leroi-Gourhan, A. (1993). *Gesture and Speech*. Massachusetts Institute of Technology.
- Leroi-Gourhan, A. (2024). The Origin and Dissemination of Scientific Knowledge. *Technology and Language*, 5(2), 101-115. <https://doi.org/10.48417/technolang.2024.02.08>
- Lévi-Strauss C. (1988). '...Nous avons lui et moi essayé...' [He and I Tried...]. In A. Leroi-Gourhan, & Bernot, L. (Eds.), *André Leroi-Gourhan, ou Les voies de l'homme: actes du Colloque du CNRS, mars 1987* (pp. 201-206). Albin Michel.
- Piaget, J. (1968). *Le structuralisme* [Structuralism]. Puf.
- Schlanger, N. (2023). *L'invention de la technologie. Une histoire intellectuelle avec André Leroi-Gourhan* [The Invention of Technology. An Intellectual History with André Leroi-Gourhan]. Puf.
- Steiner, P. (Ed.) (2014). *The Prague School. Selected Writings. 1929-1946*. University of Texas Press.
- Titchener, E. B. (1898). The Postulates of a Structural Psychology. *Philosophical Review*, 7, 449-465. <https://doi.org/10.2307/2177110>

СВЕДЕНИЯ ОБ АВТОРЕ / THE AUTHOR

Симон Аврора, simone.aurora@unipd.it

Simone Aurora, simone.aurora@unipd.it

Статья поступила 14 января 2024
одобрена после рецензирования 18 мая 2024
принята к публикации 7 июня 2024

Received: 14 January 2024
Revised: 18 May 2024
Accepted: 7 June 2024



<https://doi.org/10.48417/technolang.2024.02.11>

Research article

Between Endorsement and Disavowal: André Leroi-Gourhan's Russian Interactions

Nathan Schlanger  

École nationale des chartes, 65, rue de Richelieu, 75002 Paris, France

schlanger1@gmail.com

Abstract

Material-based interpretations of everyday undertakings have long been of interest to the French social sciences, including anthropology and history. André Leroi-Gourhan (1911-1986) follows to some extent this trend, insofar as his pioneering contributions to ethnographic and prehistoric technology – from the “elementary forms of human activity,” to studies of stone tool manufacture, to the formulation of the “chaîne opératoire” – shed much light on the more tangible and infrastructural dimensions of human existence. At the same time, his predominantly idealist recourse to evolutionary “tendencies,” “vital thrusts” (élan vital), and suchlike metaphysical notions rather held him at bay from would-be historical and dialectical understandings of primitive socio-economic formations – and this, despite his ready access to and close acquaintance with the professional literature from the other side of the Iron Curtain. Hence the paradox, as outlined here, of Leroi-Gourhan's distant attitude towards the conceptual (historical-materialist) substrate of Russian-cum-Soviet archaeology, on whose practical achievements he nonetheless remained well-informed and appreciative. In turn, this ambivalence may partly explain the rather superficial and incomplete perception of Leroi-Gourhan's works within Soviet archaeology and anthropology, limited to his publications on Prehistoric art and religion while ignoring his broad-ranging contributions to “anthropogenesis.”

Keywords: André Leroi-Gourhan; S. A. Semenov; Palaeolithic archaeology; Prehistoric technology; Stone tool manufacture; “Proletarian archaeology”; Planimetric excavations; Pincevent

Acknowledgment I thank our Paris discussion group (Xavier Guchet, Bernadette Bensaude-Vincent, Simone Aurora, Alfred Nordmann, Oliver Schlaudt, Saskia Brown, Alberto Romele) for the encouragement to publish this paper, which is a considerably reworked version of a broader study to appear in a collective publication on Archaeology and the Cold War.

Citation: Schlanger, N. (2024). Between Endorsement and Disavowal: André Leroi-Gourhan's Russian Interactions. *Technology and Language*, 5(2), 136-152. <https://doi.org/10.48417/technolang.2024.02.11>



This work is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/)





УДК 130.2(091)

<https://doi.org/10.48417/technolang.2024.02.11>

Научная статья

Между одобрением и отрицанием: взаимодействие Андре Леруа-Гурана с Россией

Натан Шлангер  

Национальная школа хартий, 65, де Ришелье, 75002 Париж, Франция

schlanger1@gmail.com

Аннотация

Материально-ориентированные интерпретации повседневности уже давно представляют интерес для французских социальных наук, включая антропологию и историю. Андре Леруа-Гуран (1911-1986) в некоторой степени следует этой тенденции, поскольку его новаторский вклад в этнографические и доисторические технологии – от “элементарных форм человеческой деятельности” до исследований изготовления каменных орудий и формулирования “*chaîne opératoire*” – пролили много света на более материальные и инфраструктурные аспекты человеческого существования. В то же время его преимущественно идеалистическое обращение к эволюционным “тенденциям”, “жизненным толчкам” (*élan Vital*) и тому подобным метафизическим представлениям скорее удерживало его от потенциального исторического и диалектического понимания примитивных социально-экономических формаций – и это, несмотря на его свободный доступ и близкое знакомство с профессиональной литературой по другую сторону железного занавеса. Отсюда описанный здесь парадокс отстраненного отношения Леруа-Гурана к концептуальному (историко-материалистическому) субстрату российско-советской археологии, о практических достижениях которой он, тем не менее, был хорошо информирован и высоко ее оценивал. В свою очередь, эта двойственность может отчасти объяснить довольно поверхностное и неполное восприятие работ Леруа-Гурана в советской археологии и антропологии, ограничивающееся его публикациями по доисторическому искусству и религии и игнорирующее его широкий вклад в “антропогенез”.

Ключевые слова: Андре Леруа-Гуран; С. А. Семенов; Археология палеолита; Доисторическая технология; Изготовление каменных орудий; “Пролетарская археология”; контурные раскопки; Пенсеван

Благодарность: Я благодарю нашу парижскую дискуссионную группу (Ксавье Гуше, Бернадett Бенсоуд-Винсент, Симона Аврору, Альфреда Нордмана, Оливера Шлаудта, Саскию Браун, Альберто Ромеле) за поддержку публикации этой статьи, которая представляет собой значительно переработанную версию более широкого исследования, которое появится позже. в коллективном издании по археологии и холодной войне.

Для цитирования: Schlanger, N. Between Endorsement and Disavowal: André Leroi-Gourhan's Russian Interactions // Technology and Language. 2024. № 5(2). P. 136-152.
<https://doi.org/10.48417/technolang.2024.02.11>



This work is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/)



INTRODUCTION

Need prehistoric archaeology, a scientific discipline specialising in the meticulous study of ancient material remains, be materialist at heart, or by default? The study of techniques, from early prehistory to modern times, is clearly predicated on rigorously empirical and experimental research, but does it necessarily follow that technology is or should be exclusively concerned with the material dimensions, the substrate of social life? The Russian interactions of André Leroi-Gourhan (1911-1986), the central figure around which this paper revolves, highlights some of the complications at stakes. While his pioneering studies of ethnographic and prehistoric technologies clearly shed much light on the infrastructural dimensions of human existence, his predominantly idealist recourse to evolutionary “tendencies” and suchlike “vital thrusts” rather held him at bay from a would-be material-historical understanding of primitive socio-economic formations. To these ambiguities can be added Leroi-Gourhan's evident linguistic propensities and conceptual dispositions towards Russian archaeology – including, in some respects, its Soviet versions –, affinities that, paradoxically enough, were barely reciprocated, even until today, on the other side of the (former) Iron curtain.

ANTECEDENTS

The French social sciences and humanities have long displayed considerable interest in the material forms of historically and anthropologically known societies. Without reaching back to Saint-Simon or Auguste Comte, we can begin with Émile Durkheim's (1858-1917) identification, early in his career, of possible causal or explanatory links between what he called the moral and the material “densities” of modern societies. This apparent proximity to materialist tenets led Durkheim himself to an abrupt volte-face; from then on, he deliberately emphasized the importance of religion in social life as a means to counter the hypothesis of its economic or infrastructural determination. By the interwar years, nevertheless, the French scientific milieu was far better disposed to welcome Marxism into the social and human sciences (Gouarné, 2013; Gouarné & Kirtchik, 2022). It was partly in this spirit that Durkheim's nephew Marcel Mauss (1872-1950) effectively introduced technology – the study of techniques, defined as “traditional efficient acts” (pursued with mechanical or physico-chemical ends) – in his anthropological teachings. From 1928 onwards, Mauss also contributed to the conceptual foundations of the new *Musée de l'Homme*, together with physical anthropologist and anti-fascist militant Paul Rivet (1876-1958). In this spirit, both scholars placed distinctive explanatory and epistemological emphasis on ordinary life and mundane material practices.

What is important to know – as Rivet indicated in 1936 – are all the aspects, or at least the average aspects of a civilization, and not the exceptional forms it takes among the privileged classes. For the ethnologist, the house of the poor is as precious and informative as is the palace of the rich. The humblest and most imperfect tool, the coarsest pottery, have as much if not more value as the most finely decorated vases. It is only on the average cultural state [*état culturel*]



moyen] of a given region that the ethnologist's comparisons should bear (Rivet, 1936, p. 13).

Two of Mauss' well-known students took these lessons to heart. One, André-Georges Haudricourt (1911-1996) was something of a maverick rationalist, left-leaning and Russophile to boot. Given his interests in agronomy and in agricultural practices, Mauss sent him to Leningrad and Turkmenistan in 1934 – not an easy time-period to dwell in the USSR – to study plant domestication and dispersal under the famous botanist Nikolai Ivanovich Vavilov (Haudricourt, 1987; Haudricourt & Dibie, 1987; Barbe & Bert, 2011; as well as Haudricourt's 1994 account of his times in Soviet Russia in Haudricourt & Bertrand, 2002).

Mauss's other technology student was André Leroi-Gourhan.¹ Already in one of his earliest publications, in the 1936 *Encyclopedie française* edited by Rivet and historian Lucien Febvre, Leroi-Gourhan proposed to replace the standard classification of techniques (according to their finality: hunting, fishing, basketry, weaving) with a new approach which highlighted the universal mechanical and functional features implicated in the “elementary forms of human activities” (Leroi-Gourhan, 1936). Further enriched by meticulous observations gathered during his ethnographic and archaeological fieldwork in Japan (from 1937 to 1939), Leroi-Gourhan's “elementary” technomorphological insights were considerably expanded in the two volumes of *Évolution et techniques* (*L'Homme et la matière* 1943, *Milieu et technique* 1945). These publications rapidly secured their author's reputation among ethnologists, museum professionals, archaeologists, historians and philosophers too, as a leading specialist in the study of “materially creative activities” in the social and human sciences (Leroi-Gourhan, 1943, 1945; partly translated into English in Leroi-Gourhan 2024).

Yet, alongside the comprehensive scope and descriptive qualities of these volumes, their more theoretical pages left somehow undecided the ultimately materialist or idealist tenor of his claims, whereby, to give an example, “the inevitable, foreseeable, rectilinear character [of the tendency] drives the hand-held flint [tool] to acquire a handle, or the bundle dragged on two poles to equip itself with wheels” (Leroi-Gourhan, 1943, p. 27). While readers such as Haudricourt, Febvre or linguist Marcel Cohen rather deplored the intrusion here of idealist speculations, others – including philosophers Emmanuel Mounier, Raymond Ruyer, and to an extent Georges Canguilhem – were clearly attuned to these metaphysical overtures. This simultaneously empirical and abstract conception of a continuously incremental “universal technicity” – inspired in part by the *élan vital* philosophy of Collège de France professor Henri Bergson (1907) and by the evolutionism of Jesuit-palaeontologist Pierre Teilhard de Chardin (1955) – came to fruition in Leroi-Gourhan's two volume masterpiece *Le Geste et la parole* in 1964-1965. Alongside their archaeological reception, these volumes also reached and influenced a particularly widespread anthropological, philosophical and indeed post-modern and post-humanist readership – in French and, in function of their availability, in translation as well.²

¹ On Leroi-Gourhan's life and work, see mainly Soulier 2018, as well as Audouze 2002, Audouze & Schlanger, 2004, and Schlanger, 2023; 2024, on which several of the following discussions are based.

² Leroi-Gourhan, 1964, 1965; translated into English in Leroi-Gourhan 1993. For some comments on Leroi-



The broad-ranging achievements detailed in *Le Geste et la parole* reflected two distinct expansions in Leroi-Gourhan's research horizons from the mid-1940s onwards. First, the emphasis he placed there on human cultural and biological evolution corresponded to his gradual shift from the fields of ethnology and museology towards prehistoric archaeology and, to a lesser extent, human palaeontology. Although this shift was partly motivated by disciplinary challenges and manoeuvres (notably due to the emergence of Lévi-Straussian structural anthropology and the rise of Marxist thought in the social sciences), it also confirmed Leroi-Gourhan's growing research interests in the origins and developments of humankind – as embodied in the figure of *Homo faber*, this “tool-making animal” successively invoked by Benjamin Franklin, Karl Marx and Henri Bergson.

A quite distinctive (though not necessarily antithetical) expansion in Leroi-Gourhan's worldviews ensued from the reaffirmation of his Catholic faith. As already intimated, these religious sensibilities occasionally transpired in his scientific work, usually implicitly or inconspicuously – and in any case without proselytising intents –, as a source of intellectual and spiritual guidance. This affinity was manifest by the mid-1950s, when Leroi-Gourhan became a “lay sympathiser (*sympathisant laïque*)” at the *Centre catholique des intellectuels français* (CCIF), a resolutely open-minded gathering of intellectuals, scientists and clergy concerned with the place and the relevance of religion in contemporary life. Leroi-Gourhan gave several presentations at the CCIF – on such topics as “Human origins” (1955), “Techniques and society among humans and animals” (1957) and “The technological illusion” (1960) – and these effectively foreshadowed the sweeping perspectives he went on to elaborate in *Le Geste et la parole* (see translations in Leroi-Gourhan, 2024).

A technologist, then, a prehistorian and a believer as well, Leroi-Gourhan was also – as yet another variable in our shifting kaleidoscope – a keen and well-informed Russophile. His adolescent encounters with post-revolutionary émigrés spurred his lifelong fascination with the culture and language of “old Russia” – that is, of “white” coloration, as distinct from the “reds” favoured by Haudricourt. These interests, coupled with his manifest linguistic abilities, proved their intellectual and strategic worth throughout his life. Mastery of Russian undoubtedly contributed to his secondment in 1935 to the Arctic department of the *musée de l'Homme*, whose director Anatole Lewitsky (himself a Russian émigré) was then developing the field of “comparative technology”. In his doctoral research on *Archéologie du Pacifique nord* (published in 1946) Leroi-Gourhan was able to include a range of relevant Russian sources. Finally, from the 1950s onward, his linguistic competencies served him to present to a French audience recent prehistoric excavation and studies published in leading Russian-language journals such as *Sovetskaya arkheologiya*, *Sovetskaja Etnografiya*, and *Materialy i issledovaniya po arheologii SSSR* (Leroi-Gourhan and Mazon, 1950; Leroi-Gourhan, 1954, 1957, 1958a and 1961).

Gourhan's reception through “French theory” and beyond, see Audouze, 2002; Schlanger, 2024.



CONTACT ZONES

In addition to this general interest in Russian archaeology, Leroi-Gourhan also forged specific affinities on matters technological with Sergey Aristarkhovich Semenov (1898-1978), the founder of the laboratory for archaeological technology at the Leningrad-based USSR-wide State Academy for the History of Material Culture (GAIMK). The strength of Semenov's approach derived from his use of distinctive experimental and forensic methodologies (allegedly mastered during his former career as a security agent) for investigating the functions of prehistoric tools (Semenov, 1957, 1964, 1965/2005, 1970/2005; and Klejn, 2012, p. 307-308). In addition to its intrinsic interest, this “traceological” contribution is noteworthy for its exceptionally rapid and wholehearted endorsement by Western scholarship – a laudable recognition of a Soviet scientific breakthrough that was nevertheless wholly contingent on the 1964 translation into English (by M. W. Thompson, 1955/1961) of his 1957 *Primitive technology* (Semenov, 1957; 1964)³. As part of his technological investigations, Semenov also undertook a range of dedicated experiments on the manufacture of stone tools by percussion. Some of his preliminary results were presented (by his colleague V.V. Bounak) at a symposium on *Les processus de l'homínisation* held in Paris in 1958, where they caught Leroi-Gourhan's attentions.

Stone tool production and use were actually a fairly new research topic for Leroi-Gourhan. Since the late 1940s, he had been intrigued by the flintknapping experiments of L. Coutier and François Bordes, but – unlike these prehistorians who mainly valued the potential of these techniques as distinctive chrono-cultural markers – Leroi-Gourhan rather intended to take this “technical behaviour” to more fundamental psychological and cognitive directions. These emerging interests made him all the more attentive to Semenov's experimentally based propositions at the 1958 symposium, whereby:

the manufacture of australopithecine tools required some 3 to 5 [striking] blows, Chellean tools needed 20 to 35 well directed blows in a single operation, Acheulean tools already required two operations and 60 to 70 blows, Mousterian points implied four distinct operations and nearly 100 different blows – and, in the case of Cro-Magnon hafted blades, as many as eleven operations and up to 200 or 250 blows. (Bounak, 1958, p. 104-105)

In the discussion that followed, Leroi-Gourhan specifically welcomed these results “on the enrichment of operational series in the course of the Palaeolithic [, which] correspond perfectly with those of my own research” (Leroi-Gourhan, 1958b, p. 110)⁴ – namely a conception of incremental accumulation in prehistoric flintknapping procedures and products which Leroi-Gourhan championed well into the 1960s. Similarly, while he did not fully endorse all of Bounak's phonetic theory of primitive language (presented at the 1958 symposium), Leroi-Gourhan “totally approved the [latter's] idea of a primordial

³ On traceological research in the USSR see Levitt, 1979; Philipps, 1988; as well as Longo & Skakun, 2005.

⁴ “*Les résultats des travaux cités de Semenov sur l'enrichissement des séries opératoires au cours du Paléolithique correspondent rigoureusement avec ceux de mes propres recherches.*” On this topic, see also Semenov, 1964, p. 42-45. and Semenov, 1970/2005.



link between technical gesture and language (...)” (p. 110) – this being a precocious formulation of the insight that was to feature so prominently in *Le Geste et la parole*, regarding the mutually enriching developments of techniques and of language.⁵

In any case, Leroi-Gourhan put to good use his unmediated linguistic access to the approaches and achievements of the Leningrad Institute, and of Soviet archaeology more generally. Besides listing selected titles and journals dealing with prehistoric matters, he also translated in 1954 a brief extract of a 1949 report by Pëtr Petrovich Efimenko (1884-1969) on his excavations at the Upper Palaeolithic open-air site of Kostienki (Leroi-Gourhan, 1954; also Efimenko, 1949). Of even greater significance was Leroi-Gourhan's 1961 review of recent archaeological publications from the Lower Palaeolithic to the Middle-Ages for the *Cahiers du monde russe et soviétique* – a review in which he was particularly attentive to the conceptual foundations of Soviet research. He readily endorsed there the notion of “the history of material culture,” and attributed the originality of Soviet archaeology to its “search for the evidence of material culture as a means for investigating the technical and economic history of human groups who are not accessible through written archives” (Leroi-Gourhan, 1961, p. 262).⁶ Soviet archaeology was no longer restricted to noble works of art, he concurred, but rather sought to recover tools, domestic waste, dwellings and sites of cult, thus aiming for “a complete history, at once economic, social, technical and intellectual, of those people which have not been touched, or have been neglected, by writing” (p. 262).⁷ Foreshadowing a “truly ethnological archaeology,” this proximity to the “man of every day” made it all the more understandable, so concluded Leroi-Gourhan with evident approval, why “the popular republics [of the Soviet Union] have given such a considerable development to this 'proletarian' archaeology [*cette archéologie "prolétarienne"*]” (p. 262).⁸

AMBIGUITIES

This exceptional endorsement of “proletarian' archaeology” – expressed in a generalist “area studies” journal dedicated to the broader social, cultural and political dimensions of Russian and Soviet history – was probably the closest Leroi-Gourhan ever came to an intellectual engagement with the tenets of Marxist historical materialism. This unicum may well have been inspired from contemporary Soviet publications (if such an

⁵ See Leroi-Gourhan, 1964, p. 162, p. 306 note 10 [Translated Leroi-Gourhan, 1993, p. 114, p. 412 note 10.

⁶ “(...) *ce qu'il y a de plus propre et original dans les travaux russes récents : la recherche des témoins de la culture matérielle comme moyen d'investigation de l'histoire technique et économique des groupes humains non accessibles à travers les archives écrites.*”

⁷ “*On fouillera [en archéologie soviétique] beaucoup moins pour découvrir des œuvres d'art ou des inscriptions (documents qui ne sont certes pas à négliger, mais qui viennent en surcroît du reste) que pour découvrir sous forme d'outils, de débris ménagers, de traces d'habitats ou de lieux de culte, une histoire complète, c'est-à-dire à la fois économique, sociale, technique et intellectuelle, des hommes que l'écriture n'aura pas atteints de leur vivant ou qu'elle aura négligés.*”

⁸ “*A côté donc de la noble archéologie classique est née une archéologie véritablement ethnologique, beaucoup plus proche de l'homme de tous les jours que ne l'était celle des grands monuments[,] et les raisons du développement considérable que les républiques populaires ont donné à cette archéologie 'prolétarienne' sont particulièrement compréhensibles.*”



expression was in use), unless it somehow harked back to the confrontation between “proletarian” and “bourgeois” science spurred by the infamous Lysenko affair in the late 1940s.⁹ As for these concerns with the ordinary remains of everyday life, taken as a key to “total history”, they undoubtedly reprised and extended some of Mauss's and Rivet's intuitions of the 1930s. The French research tradition as a whole clearly favoured the notions of “techniques” and “technology”, and Leroi-Gourhan reference in 1961 to the “history of material culture” served him to acknowledge both the conceptual and the institutional centrality of this notion in Soviet research.

Leroi-Gourhan's long-established interests were in documenting and explaining the evolution of technical tendencies (“the pre-conceivable lines [of development] from undifferentiated flint to finely worked blades, to copper knives and steel sabers”, Leroi-Gourhan, 1943, p. 13; translated in Leroi-Gourhan, 2024). These interests did not really encompass, to use Marxist parlance, the productive activities of historical social units in prehistoric times. Since the early 1950s, Leroi-Gourhan's favoured the notions of “behaviour (*comportement technique*)”, as applicable to humans, animals and machines too¹⁰, and that of embedded and immemorial “craft” undertakings (*artisanat*). In between, there remained little room for an appreciation of “labour” as a socio-economic, productive and even creative process. In his publications and his teachings (including to students who would later create the “*Technologie culturelle*” and “*Anthropologie des techniques*” research movements), Leroi-Gourhan mostly remained uncommitted, if not indifferent, on these theoretical positions. This attitude surfaces in his 1982 book of conversations-reminiscences, *Les Racines du monde*:

I see no contradiction [between attaching importance to 'infrastructures' and not being a 'materialist']. It is a point of view that creates a contradiction [only] if it is transposed to the political level. I do not feel it. Since the publication of *L'Homme et la matière*, I have had numerous contacts with Marxists. They have recognised me, whereas I did not recognise them... I have practiced Marxism like Monsieur Jourdain [of Molière's play], without realising it, and I continue. (...) [Marx] did not play much of a role [in my intellectual formation]. I have read fragments of his work, and never took it as a whole. Generally speaking, I am not inclined towards exegesis. I will not claim that I have rediscovered the America of Marxism, but there is a bit of that in my attitude. After the fact, I found in certain texts by Marx that I have been made to read things that I have been thinking about on my own. (Leroi-Gourhan, 1982, p. 229)¹¹

⁹ On the controversies raised in France and in Western Europe by Trofim Denisovich Lysenko's attempted refutation of genetic biology thorough a Marxian-compatible form of Lamarckism, see mainly Lecourt, 1977.

¹⁰ See Leroi-Gourhan 1952, and 1953, translated in this issue of *Technology and Language* (Leroi-Gourhan, 2024).

¹¹ “*Je n’y vois pas de contradiction [entre le fait d’accorder de l’importance aux “infrastructures” et ne pas être un “matérialiste”]. C’est un point de vue qui, si on le transpose sur le plan politique, crée une contradiction. Moi, je ne la ressens pas. Depuis la publication de L’Homme et la Matière, j’ai eu beaucoup de contacts avec les marxistes. Ils m’ont reconnu alors que je ne les reconnaissais pas... J’ai fait du marxisme comme Monsieur Jourdain [de Molière], sans m’en rendre compte, et je continue. [Marx] n’a*



In fact, when Leroi-Gourhan mentioned this “proletarian archaeology” in 1961, he was already identified (and self-identified) as a faithful “*compagnon de route*” – not however of free-thinking, left-leaning, or even Communist party related intellectuals (as were quite a few of his contemporary ethnologists and historians), but rather of Catholic ones. Seeking to maintain for humanity an irreducible core of spiritual solace against the assaults of atheist propaganda – a *supplement d'âme*, as it were – the CCIF organised several debates to counter or deflect the secularist threats emanating from the French Rationalist Union, and indeed from Marxist Russia (Schlanger, 2023, p. 333-341.). These intermittent conflicts and their latent entrenchments may well explain, at least in part, Leroi-Gourhan's paradoxical position: holding a resolutely agnostic attitude towards the conceptual (historical-materialist) substrate of Russian-cum-Soviet archaeology – while nonetheless remaining manifestly well-informed, and appreciative, of its practical achievements.

MUTUAL IGNORANCE

A comparable disjunction between knowledge and ignorance (both *sensu* lack of awareness and refusal to acknowledge) seems to have prevailed on the other side of the Iron curtain. Better documented research on this issue is certainly needed,¹² but it is clear that some explanation is called for regarding Leroi-Gourhan's intermittent and underwhelming presence across Soviet archaeology and anthropology. An already encountered generic factor has to do with linguistic limitations: few Russian scholars were able at that time to engage with his works in the original French, and translations were clearly wanting.¹³ Likewise, our protagonist's own cross-disciplinary or indeed eclectic propensities can explain this relative lack of visibility. Like many of his readers elsewhere (including in France itself) Russian scholars could be forgiven for not fully appraising Leroi-Gourhan's remarkably diversified contributions in the fields of ethnography, prehistory and technology – as if each domain were somehow held apart, possibly deliberately so, through some opaque disciplinary curtains of their own.

This diversity, occasionally verging on dispersion, is perceptible also within the broad field of prehistoric or Palaeolithic archaeology. In Russia, Leroi-Gourhan's

pas joué un bien grand rôle [dans ma formation intellectuelle]. J'ai lu de lui des fragments et je n'ai jamais pris l'œuvre dans son ensemble. D'une façon générale, je ne suis pas porté vers l'exégèse. Je ne vais pas dire que j'ai redécouvert l'Amérique du marxisme, mais il y a un peu de cela dans mon comportement. Après coup, j'ai trouvé, dans certains textes de Marx qu'on m'a fait lire, des choses que j'avais pensées de mon côté.”

¹² Beginning with a deeper appreciation of the tenets and specificities of Russian and Soviet archaeology, from Miller, 1956; Mongaït, 1951, 1955/1961; as well as Thompson, 1961; to Bulkin et al., 1982; Davis, 1983; Vasiliev, 2004, 2011; Lozny, 2017; and Klejn 2012, 2017; see also Gellner, 1980; Plotkin & Howe, 1985; Bertrand, 2002, and Alymov, 2022 on relevant aspects of Soviet and Russian anthropology.

¹³ Leroi-Gourhan's translation history is somewhat chequered (see discussion in Schlanger, 2024). Unlike their rapid translation in Latin languages, *Le Geste et la parole* only appeared in English with a thirty years delay (in 1993) and *Évolution et techniques* not at all. This differs from Leroi-Gourhan's archaeological and art historical titles, mostly translated into English or American in the couple of years following their publication.



recognised and acknowledged contributions mostly concerned his studies of Franco-Cantabrian upper-Palaeolithic parietal art, for which he proposed a chronology based on stylistic criteria and a structuralist interpretation. In this spirit, a (partial) Russian translation of his *Religions de la préhistoire* (1964) was published in Leroi-Gourhan, 1971. However, so far as prehistoric stone tools assemblages and cultural sequences were concerned, the readily recognised and quoted authority was undoubtedly François Bordes (1919-1981). An expert experimental flintknapper in his own right, the Bordeaux-based prehistorian had reached early on the conclusion that prehistoric civilisations are best characterised by the form of the tools they produced, and not their techniques of manufacture. Bordes's highly influential “morphological typology,” whose statistical rigour was found appealing also by Soviet scholars, cast prehistoric tool types as reflections of the mental images held by their ancient makers. This contrasted with the “functional typology” advanced by Semenov and his GAIMK colleagues, who aimed to study and to understand prehistoric tools as the outcomes of the technical productive processes in which they were implicated.¹⁴

Now it so happens that Leroi-Gourhan's interests and publications, from the very onset of his scientific career, were clearly set within this later functionalist and dynamic perspective; yet this was a contribution that Semenov and Soviet archaeology at large somehow appear to have missed. There is no doubt that Leroi-Gourhan's pioneering studies of ethnographic and prehistoric technology – starting with his contributions to the “elementary forms of human activity” – can shed much light on the more tangible and infrastructural dimensions of human existence. The same goes for his incipient formulation and theorisation of the “*chaîne opératoire*” (partly reinforced, as we saw above, by Semenov's notions of “operations” and “blows” as presented in 1958), an approach that has since contributed decisively to the development of technology and material culture studies in archaeology, anthropology and the social sciences worldwide. More broadly, while Russian and Soviet scholars became acquainted with his more ostensibly “bourgeois” publications on Prehistoric art and religion, they have apparently overlooked or failed to engage with (and, to this day, to translate) *Le Geste et la parole* – undoubtedly one of the boldest and most comprehensive twentieth-century contributions to what is known in Russian scholarship as “anthropogenesis.” For all the differences between their philosophical and their scientific orientations, and of course the century long accumulation of knowledge in between, Leroi-Gourhan's opus clearly stands well on par with Friedrich Engels' inspired conjectures in the *Dialectics of Nature*, notably regarding the role of the erect posture and the liberation of the hand, as both instrument and product, in “The part played by labour in the transition from Ape to Man” (Engels, 1883).¹⁵

¹⁴ On this debate see Bordes, 1967 (in reaction to Semenov's 1964 book) and Semenov's 1968 reply.

¹⁵ It may be mentioned that these dimensions of Leroi-Gourhan's work are beginning to be taken on board by Russian anthropology, for example Krutkin, 2019.



CONCLUSIONS: PINCEVANT, *TERRAIN D'ENTENTE* OR NO-MAN'S-LAND?

If there is a place that could have served as an ideal meeting grounds, a site of mutual recognition and collaboration, it is surely Pincevent – an unimposing floodplain on the banks of the Seine some 80 km south of Paris, where quarrying works led in early May 1964 to the discovery of extraordinarily well-preserved campsites of Magdalenian reindeer hunters dating to some 12,000 years ago. The material disposition and research potential of this site proved to be in several respects similar to those which Leroi-Gourhan had been reviewing in the Russian literature. This familiarity no doubt contributed to his decision to dedicate henceforth most of his energy and institutional resources to the excavation of Pincevent – thereby establishing what is probably the longest running and continuously productive research programme in Palaeolithic archaeology worldwide.¹⁶

Yet, so far as Prehistoric interpretations as such were concerned, this may well have been something of a *rendez-vous manqué*. While the meticulous horizontal peeling or *décapage* of the archaeological layers carried out by Leroi-Gourhan and his teams – known as planimetric excavations – made it possible to recover and document the “living floors” of these ancient “men of every day,” this empirical approach actually left little scope, at least initially, for an inquiry into the modes of production and of reproduction of the socio-economic formations at hand, let alone the study of their “material culture” within the scope of “proletarian archaeology.” Likewise, while the very first scientific presentation of the site's preliminary excavations happened to take place at an anthropological congress held in August 1964 in Moscow, this temporary translocation to Soviet lands seems to have left little repercussions in either country (Leroi-Gourhan, 1970; and Alymov, 2022 on the 1964 congress).

In fact, rather than through some theoretical impacts on archaeological interpretations, the Russian-cum-Soviet affinities of Pincevent manifested themselves at a different level, at the intersection of heritage policies and politico-nationalist ideologies. As just noted, it was upon its partial destruction by the surrounding gravel quarry that the site of Pincevent came to light: the ensuing urgency to protect it from further devastation and to secure its long-term availability for scientific research brought questions of heritage management and legislation to the fore. On these issues, as Leroi-Gourhan knew well, French heritage policies paled in comparison with the protective regulations and centralised initiatives enacted by Soviet archaeology since the 1930, which, as paleolithician Pavel Boriskovski (1965) explained to his French audience in 1964, compelled “the entrepreneurs of large scale works to deduct from their annual budget a sum dedicated to the archaeological research and excavations made necessary by these works, prior to their execution” (p. 8-9). In these post-war years of the *Trentes glorieuses*, rich in infrastructure projects in both rural and urban areas, France was clearly lacking behind in terms of organisation, legislation and personnel as well (Demoule & Schnapp, 2024; Négri & Schlanger, 2024). With this comparison in mind, Leroi-Gourhan recorded his frustration in an early report on Pincevent:

¹⁶ On the history of research at Pincevent, see in particular Soulier, 2018, p. 463-481, and Soulier, 2021; also, among numerous other publications, Karlin & Julien, 2012; Ballinger et al., 2014; Valentin et al., 2015.



Given the tidal wave that the development of major works represents for archaeology, it is timely to point out the scientific inadequacy of many excavations where the gathering of material 'documents' takes precedence over observation. France has shown through a number of works that it is not inferior to the best foreign teams, but five or six insufficiently staffed and equipped prehistoric excavation teams cannot, despite their quality and zeal, cope with thousands of bulldozers. (...) it is to be hoped that this current impetus in archaeology will result in the broadest possible adaptation of research methods to the needs and pace of modern life. (Leroi-Gourhan & Brézillon 1964, pp. 63-64)¹⁷

It is not clear exactly how or by whom – it would make perfect sense to conjecture Leroi-Gourhan's personal involvement – were the merits of Soviet heritage management policies brought to the attention of the relevant officials, and indeed of the highest realms of the Republic. Through an admixture of emulation and competition, equally designed to enhance the archaeology of France and the benevolent reputation of the French state, we find already in November 1964 the novelist and former adventurer André Malraux, now Minister for Cultural Affairs in the government of Général de Gaulle, declaring in his budget parliamentary speech:

“This expression of a national will, essential for us – not nationalistic, but national – concerns first and foremost our heritage. This year, we had to draw up an excavation program. (...) In France, we have discovered the site of Pincevent, the largest Palaeolithic site in the world, *with the exception of the Russian sites*. In just three days, we achieved what we had set out to do [to protect the site].”¹⁸ To make sure that “the Russian sites” in question are understood not just in terms of their purely geographical localisation, let alone their specialised scientific potential, but also in all their political and ideological dimensions as well, Malraux added in another parliamentary speech a couple of years later: “The awakening of archaeology is a fact, given that Pincevent is the largest prehistoric site *on this side of the Iron Curtain*.”¹⁹

¹⁷ “Par contre, devant le raz de marée que représente pour l’archéologie le développement des grands travaux, il est à peine temps encore de souligner l’insuffisance scientifique de nombreuses fouilles où le rassemblement de “documents” matériels prime l’observation, de dire aussi que la France témoigne par un certain nombre de travaux du fait qu’elle n’est pas au-dessous des meilleures équipes étrangères, mais que cinq ou six équipes de fouille préhistoriques insuffisamment étoffées et équipées ne peuvent pas, malgré leur qualité et leur mordant, faire face à des milliers de bull-dozers. (...) Mais il est à souhaiter que l’élan actuel de l’archéologie se traduise par une réadaptation aussi large que possible des méthodes de recherches aux besoins et au rythme de la vie présente.”

¹⁸ “Cette expression d’une volonté nationale, essentielle chez nous – non pas nationaliste, mais nationale – concerne d’abord notre patrimoine (...). En France, on a découvert le gisement de Pincevent, le plus grand gisement paléolithique du monde, exception faite des gisements russes. En trois jours, ce qui était demandé a été obtenu”, André Malraux, Loi de finances pour 1965, “Présentation du budget des affaires culturelles”, Assemblée nationale, séance du 7 novembre 1964, p. 4992 (emphasis added).

¹⁹ “Le réveil de l’archéologie est un fait, puisque Pincevent est le plus grand site préhistorique de ce côté du rideau de fer”. André Malraux, Loi de finances pour 1967 “examen des crédits du ministère des affaires culturelles”, Assemblée nationale, séance du 27 Octobre 1966, p. 3975 (emphasis added).



REFERENCES

- Alymov, S. (2022). Cold War Anthropology from Both Sides of the Iron Curtain. *Revue d'histoire des sciences humaines*, 40, 85-109. <https://doi.org/10.4000/rhsh.6853>
- Audouze, F. (2002). Leroi-Gourhan, a Philosopher of Technique and of Evolution. *Journal of Archaeological Research*, 10, 277-306. <https://doi.org/10.1023/A:1020599009172>
- Audouze, F. & Schlanger, N. (2004). Introduction: singularité et diversité [Introduction: singularity and diversity]. In F. Audouze & N. Schlanger (Eds.), *Autour de l'homme: contexte et actualité d'André Leroi-Gourhan* (pp. 17-29). Editions APDCA.
- Ballinger, M., Bignon-Lau, O., Bodu, P., Debout, G., Dumarçay, G., Hardy, M., Julien, M., Karlin, C., Malgarini, R., Orliac, M., Peschaux, C., Soulier, P., & Valentin, B. (2014). *Pincevent (1964–2014): 50 années de recherches sur la vie des Magdaléniens* [Pincevent (1964–2014): 50 Years of Research on the Life of the Magdalenians]. Société préhistorique française.
- Barbe, N. & Bert, J.-F. (Eds.). (2011). *Penser le concret. André Leroi-Gourhan, André-Georges Haudricourt, Charles Parain* [Thinking concretely. André Leroi-Gourhan, André-Georges Haudricourt, Charles Parain]. Creaphis.
- Bergson, H. (1907). *L'Évolution créatrice* [Creative Evolution]. Presses universitaires de France.
- Bertrand, F. (2002). *L'Anthropologie soviétique des années 20-30. Configurations d'une rupture* [Soviet Anthropology of the 20s and 30s. Configurations of a Rupture]. Presses universitaires de Bordeaux.
- Bordes, F. (1967). Considérations sur la Typologie et les techniques dans le Paléolithique [Considerations on Palaeolithic Typology and Techniques]. *Quartär*, 18, 25-55.
- Boriskovski, P. I. (1965). À propos des récents progrès des études paléolithiques en U.R.S.S. [About Recent Advances in Paleolithic Studies in the U.S.S.R.]. *L'Anthropologie*, 6, 5-30.
- Bounak, V. V. (1958). L'origine du langage [The Origin of Language]. In *Les Processus de l'hominisation: colloques internationaux du CNRS, Paris 19-23 mai 1958* (pp. 99-111). CNRS.
- Bulkin V. A., Klejn, L. S., & Lebedev, G. S. (1982). Attainments and Problems of Soviet Archaeology. *World Archaeology*, 13(3), 272-295.
- Davis, R. S. (1983). Theoretical Issues in Contemporary Soviet Paleolithic Archaeology. *Annual Review of Anthropology*, 12, 403-428. <https://doi.org/10.1146/annurev.an.12.100183.002155>
- Demoule, J.-P., & Schnapp, A. (2024). *Qui a peur de l'archéologie ? La France face à son passé* [Who is Afraid of Archaeology? France Facing its Past.]. Les Belles Lettres.
- Efimenko, P. P. (1949). Iz materialov paleoleticheskogo poseleniya Kostenki 1 [From the Materials of the Paleolithic Settlement of Kostenki 1]. *Sovietskaia archeologia*, 11, 113-126.
- Engels, F. (1883). *Dialectics of Nature*. Progress Press.
- Gellner, E. (Ed.) (1980). *Soviet and Western Anthropology*. Columbia University Press.



- Gouarné, I. (2013). *L'Introduction du marxisme en France. Philosoïétisme et sciences humaines, 1920-1939* [The Introduction of Marxism in France. Philosoïetism and Human Sciences, 1920-1939]. Presses universitaires de Rennes.
- Gouarné, I. & Kirtchik, O. (Ed). (2022). Réformer (avec) les sciences sociales en Union soviétique [Reforming (with) the Social Sciences in the Soviet Union]. *Revue d'histoire des sciences humaines*, 40. <https://doi.org/10.4000/rhsh.6744>
- Haudricourt, A.-G. (1987). *La Technologie, science humaine. Recherches d'histoire et d'ethnologie des techniques* [Technology, A Human Science. Historical and Ethnological Studies of Techniques]. Maison des Sciences de l'Homme.
- Haudricourt, A.-G. & Dibie, P. (1987). *Les Pieds sur terre* [With Feet on the Ground]. Métailié.
- Haudricourt, A.-G. & Bertrand, F. (2002). Entretien (décembre 1994) [Interview (December 1994)]. In Bertrand, & F. (Ed.), *L'Anthropologie soviétique des années 20-30. Configurations d'une rupture* (pp. 245-255), Presses universitaires de Bordeaux.
- Karlin, C. & Julien, M. (2012). Les Campements de Pincevent: Entre archéologie et anthropologie [The Pincevent Camps: Between Archaeology and Anthropology]. In N. Schlanger, & A. C. Taylor (Eds.), *La Préhistoire des autres: Perspectives archéologiques et anthropologiques* (pp. 185–200). La Découverte.
- Klejn, L. S. (2012). *Soviet Archaeology. Trends, Schools and History*. Oxford University Press.
- Klejn, L. S. (2017). Archaeology in Soviet Russia. In L. Lozny (Ed), *Archaeology of the Communist Era. A Political History of Archaeology of the 20th Century*, Springer, pp. 59-99.
- Krutkin, V. L. (2019). “Human Corporeality and the Technologies of Light Writing,” Forum 38: Body and Technology. *Forum for Anthropology and Culture*, 15, 32-42.
- Lecourt, D. (1977). *Proletarian Science? The Case of Lysenko* [B. Brewster, Trans.; L. Althusser, Intr.]. NKB and Humanities Press.
- Leroi-Gourhan, A. (1936). L'Homme et la nature [Humans and Nature]. In P. Rivet (Ed.), *Encyclopédie française permanente. Vol. 7, L'Espèce humaine: Peoples et races. Part I, “L'Humanité d'aujourd'hui”. Section A, “Les Formes élémentaires de l'activité humaine”* (pp. 7.10.3–16, 7.12.1–4). Comité de l'Encyclopédie française édition.
- Leroi-Gourhan, A. (1943). *L'Homme et la matière. Vol. 1., Évolution et techniques* [Humans and Matter. Vol. 1., Evolution and Techniques]. Albin Michel.
- Leroi-Gourhan, A. (1945). *Milieu et techniques. Vol. 2, Évolution et techniques* [Milieu and Techniques. Vol. 2, Evolution and techniques]. Albin Michel.
- Leroi-Gourhan, A. (1952). Homo faber... Homo sapiens. *Revue de Synthèse*, 30, 79-102.
- Leroi-Gourhan, A. (1953). Origine et diffusion de la connaissance scientifique [The Origins and Dissemination of Scientific Knowledge]. *Structure et Évolution des Techniques*, 33-34, 1-16.
- Leroi-Gourhan, A. (1954). Traduction de Matériaux sur le peuplement paléolithique de Kostienki I, par P.P. Efimenko, *Sovietskaia archaeologia*, 11, 1949, pp. 113-126 [Translation of Materials on the Paleolithic Settlement of Kostienki I, by P.P.



- Efimenko, Sovietskaia archaeologia, 11, 1949, pp. 113-126]. *Bulletin de la Société préhistorique française*, 51, 322-325.
- Leroi-Gourhan, A. (1957). Recension d'articles intéressant la préhistoire, parus récemment dans les périodiques soviétiques [Review of Articles Concerning Prehistory, Recently Published in Soviet Periodicals]. *Bulletin de la Société préhistorique française*, 54, 723-725.
- Leroi-Gourhan, A. (1958a). Recension de travaux intéressant la préhistoire parus récemment en URSS (Comptes rendus d'ouvrages) [Review of Works Concerning Prehistory Recently Published in the USSR (Book Reviews)]. *Bulletin de la Société préhistorique française*, 55, 481-482.
- Leroi-Gourhan, A. (1958b). Intervention de M. Leroi-Gourhan après la communication de M. Bounak ("L'Origine du langage") [Intervention by Mr. Leroi-Gourhan Following the Communication of Mr. Bounak ("The Origin of Language")]. In *Les processus de l'hominisation: Colloques internationaux du CNRS, Paris 19-23 mai 1958* (p. 110). Collectif, Éditions du CNRS.
- Leroi-Gourhan, A. (1961). Préhistoire et archéologie soviétique [Soviet Prehistory and Archaeology]. *Cahiers du monde russe et soviétique*, 2(2), 262-272.
- Leroi-Gourhan, A. (1964). *Le geste et la parole. Vol. 1: Technique et langage* [Gesture and Speech. Vol. 1: Technics and Language]. Albin Michel.
- Leroi-Gourhan, A. (1965). *Le geste et la parole. Vol. 2: La Mémoire et les Rythmes* [Gesture and Speech. Vol. 2: Memory and Rhythms]. Albin Michel.
- Leroi-Gourhan, A. (1970). Le campement-atelier magdalénien de Pincevent (près Montereau, Seine-et-Marne) [The Magdalenian Camp-Workshop of Pincevent (near Montereau, Seine-et-Marne)]. In *VII International Congress of Anthropological and Ethnographic Sciences* (vol. 5, pp. 372-374). Nauka.
- Leroi-Gourhan, A. (1971). Religions de la préhistoire [The Religions of Prehistory] (B. A. Frolov, Trans.). In *Art primitive* (pp. 81-90). Novossibirsk.
- Leroi-Gourhan, A. (1982). *Les racines du monde* [The Roots of the World]. Belfond.
- Leroi-Gourhan, A. (1993). *Gesture and Speech* [A. Berger, Trans.]. MIT Press.
- Leroi-Gourhan, A. (2024). *André Leroi-Gourhan on Technology. A Selection of Writings from the 1930s to the 1960s* (N. Schlanger, Ed. & Intr.). Bard Graduate Center.
- Leroi-Gourhan, A. (2024). The Origin and Dissemination of Scientific Knowledge. *Technology and Language*, 5(2), 101-115. <https://doi.org/10.48417/technolang.2024.02.08>
- Leroi-Gourhan, A. & Mazon, A. (1950). Ethnographie et géographie (Chroniques, publications: Russe) [Ethnography and Geography (Chronicles, Publications in Russian)]. *Revue des études slaves*, 26, 202-203.
- Leroi-Gourhan, A. & Brézillon, M. (1964). Le site Magdalénien de Pincevent (Seine et Marne) [The Magdalenian Site of Pincevent (Seine et Marne)]. *Bulletin de l'Association française pour l'étude du quaternaire*, 1(1), 59-64.
- Levitt, J. (1979). A Review of Experimental Traceological Research in the USSR. In B. Hayden (Ed.), *Lithic Use-wear Analysis* (pp. 27-38). Academic Press.
- Longo, L. & Skakun, N. (Eds.). (2005). *The Roots of Use-wear Analysis: Selected papers of S.A. Semenov*. Museo Civico di Storia Naturale di Verona.



- Lozny, L. R. (Ed). (2017). *Archaeology of the Communist Era. A Political History of Archaeology of the 20th Century*. Springer. <https://doi.org/10.1007/978-3-319-45108-4>
- Miller, M. O. (1956). *Archaeology in the U.S.S.R* [Research Program on the U.S.S.R., Research Monographs, No. 3, Praeger Publications in Russian History and World Communism, No. 25]. Frederick A. Praeger.
- Mongaït, A. L. (1951). Krizis burzhuaznoi arkheologi [The Crisis of Bourgeois Archaeology]. *Kratkiye soobshcheniya IIMK*, 40, 3-15.
- Mongaït, A. L. (1961). *Archaeology in the USSR*. Pelican books (Original work published 1955).
- Négri, V. & Schlanger, N. (Eds.). (2024). *1941, Genèse et développements d'une loi sur l'archéologie* [1941. The Genesis and Developments of a Law on Archaeology]. La documentation française.
- Philipps, P. (1988). Traceology (Microwear) Studies in the USSR. *World Archaeology*, 19(3), 349-356.
- Plotkin, V. & Howe, J. E. (1985). The unknown tradition: continuity and innovation in soviet ethnography. *Dialectical Anthropology*, 9, 257-312.
- Rivet, P. (1936). Ce qu'est l'ethnologie [What is Ethnology]. In P. Rivet (Ed.), *Encyclopédie française permanente. Tome VII. L'Espèce humaine: peuples et races. Section A: Formes élémentaires de l'activité humaine* (p. 7.1-16). Comité de l'Encyclopédie française.
- Schlanger, N. (2023). *L'invention de la technologie. Une histoire intellectuelle avec André Leroi-Gourhan* [The Invention of Technology. An Intellectual History with André Leroi-Gourhan]. Presses universitaires de France.
- Schlanger, N. (2024). Reading Technology - Introducing Leroi-Gourhan. In *André Leroi-Gourhan on Technology. A Selection of Writings from the 1930s to the 1960s* (pp. 1-54). Bard Graduate Center.
- Semenov, S. A. (1957). *Pervobytnaia tekhnika: opyt izucheniia drevneishikh orudiï i izdeliï po sledam raboty* [Primitive Technique: The Experience of Studying Ancient Tools and Products in the Footsteps of Work]. Institut istorii materialnoi kultury, Izdatelstvo Akademii nauk SSSR.
- Semenov, S. A. (1964). *Prehistoric Technology. An Experimental Study of the Oldest Tools and Artefacts from Traces of Manufacture and Wear*. Cory, Adams & Mackay.
- Semenov, S. A. (2005). Production and function of stone tools. In L. Longo & N. Skakun (Eds.), *The Roots of Use-wear Analysis: Selected papers of S.A. Semenov* (pp. 32-41). Museo Civico di Storia Naturale di Verona. (Original work published 1970)
- Semenov, S. A. (2005). The Experimental Method of Studying Prehistoric Technology. translated In L. Longo & N. Skakun (Eds.), *The Roots of Use-wear Analysis: Selected papers of S.A. Semenov* (pp. 134-139). Museo Civico di Storia Naturale di Verona. (Original work published 1965)
- Soulier P. (2018). *André Leroi-Gourhan (1911-1986). Une vie* [André Leroi-Gourhan (1911-1986). A Life]. CNRS Éditions.



- Soulier, P. (2021). *Pincevent (1964–2019). Cinquante-cinq années d'ethnologie préhistorique* [Pincevent (1964–2019). Fifty-Five Years of Prehistoric Ethnology]. Société préhistorique française.
- Teilhard de Chardin, P. (1955). *Le Phénomène humain* [The Human Phenomenon]. Édition du Seuil.
- Thompson, M. W. (1961). Translator's Preface. In A. L. Mongaït, *Archaeology in the USSR* (pp. 15-32). Pelican books. (Original work published 1955)
- Valentin, B., Olive, M., Valla, F., Audouze, F., de Beaune, S., Soulier, P., Chambon, P., & Bodu, P. (2015). Ou en est l'ethnologie préhistorique? [Where is Prehistoric Ethnology?]. In Soulier, P. (Ed.) *André Leroi-Gourhan, l'homme tout simplement : mémoires et postérités d'André Leroi-Gourhan* (Travaux de la maison René-Ginouvès, 20), pp. 173-186. Editions de Boccard.
- Vasiliev, S. (2004). La Notion de sol d'habitat dans la préhistoire soviétique [The Notion of Living Floor in Soviet Prehistory]. In F. Audouze & N. Schlanger (Eds.), *Autour de l'homme: contexte et actualité d'André Leroi-Gourhan* (pp. 275–83). Editions APDCA.
- Vasiliev, S. (2011). La Préhistoire russe et François Bordes: les influences réciproques [Russian Prehistory and François Bordes: Reciprocal Influences] In F. Delpech, & J. Jaubert (Eds.), *François Bordes et la Préhistoire* (pp. 71-84). Editions du Comité des travaux historiques et scientifiques.

СВЕДЕНИЯ ОБ АВТОРЕ / THE AUTHOR

Nathan Schlanger
schlanger1@gmail.com
ORCID 0000-0002-2029-4929

Натан Шлангер
schlanger1@gmail.com
ORCID 0000-0002-2029-4929

Статья поступила 14 января 2024
одобрена после рецензирования 3 мая 2024
принята к публикации 12 июня 2024

Received: 14 January 2024
Revised: 3 May 2024
Accepted: 12 June 2024





Contributed paper



<https://doi.org/10.48417/technolang.2024.02.12>

Research article

Doing Things with Words and Things: A Social Pragmatist View on the Language–Technology Analogy

Werner Rammert  

Technical University of Berlin, 17. Juni St., 135, 10623 Berlin, Germany

werner.rammert@tu-berlin.de

Abstract

This paper claims to show that the making of technology and the material agency of technical objects can be analyzed analogously to the making of meaning through words and speech acts. It proposes the development of a more comprehensive view on the making and working of technology that connects the social pragmatist approach of technical practice and symbolic interagency (Kant, Dewey, Mead) with the linguistic concept of pragmatics and speech acts (Peirce, Wittgenstein, Austin). Both, speech acts and technical acts can be considered as two modes of meaning-making in the social construction of reality. Furthermore, the paper exhibits some parallels between the objectification processes of language and technology. It emphasizes how both evolve from early stages of signs and tools in practical contexts to encoded collections of grammatical rules and technological tools later on. Doing things with concrete things (technology) reveals two different modes of “efficacy” (Jullien). There is implicit experienced efficacy in the language of directed material forces and causes and also an explicit ascribed efficacy in the language of instituted ends–means relations. The text explores the analogy between language and technology through the lenses of semantics, syntax, pragmatics, and grammar. It emphasizes the importance of such an extended pragmatist/pragmatics approach in the face of new technologies that exhibits a high degree of self-activity, more modes of intra-action between physical and digital objects, and a growing interactivity at interfaces with human actors and environmental factors. They can be more appropriately understood, conceptualized, and also designed as sociotechnical constellations of distributed agencies between people, machines, and programs.

Keywords: Technical practice; Material agency; Speech acts; Linguistic pragmatics; efficacy; Meaning; Interactivity; Digital objects

Citation: Rammert, W. (2024). Doing Things with Words and Things: A Social Pragmatist View on the Language–Technology Analogy. *Technology and Language*, 5(2), 154–177. <https://doi.org/10.48417/technolang.2024.02.12>



This work is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/)



УДК 1:62

<https://doi.org/10.48417/technolang.2024.02.12>

Научная статья

Делать вещи с помощью слов и вещей: Прагматический взгляд на аналогию языка и технологии

Вернер Раммерт (✉)

Берлинский технический университет, ул. 17 Июня, 135, 10623 Берлин, Германия

werner.rammert@tu-berlin.de

Аннотация

Статья претендует на то, чтобы показать, что создание технологий и материальное воздействие технических объектов можно анализировать аналогично созданию смысла через слова и речевые акты. Она предлагает развитие более всестороннего взгляда на создание и работу технологий, который соединяет социально-прагматический подход к технической практике и символическому взаимодействию (Кант, Дьюи, Мид) с лингвистической концепцией прагматики и речевых актов (Пирс, Витгенштейн, Остин). И речевые акты, и технические действия можно рассматривать как два способа создания смысла в социальном конструировании реальности. Кроме того, в статье показаны некоторые параллели между процессами объективации языка и технологий. В нем подчеркивается, что и то, и другое развивается от ранних стадий знаков и инструментов в практическом контексте к закодированным наборам грамматических правил и технологических инструментов на более позднем этапе. Делание вещей с помощью конкретных вещей (технологий) раскрывает два разных режима “эффективности” (Жюльен). В языке направленных материальных сил и причин присутствует неявная переживаемая эффективность, в языке установленных отношений “цель-средства” – явно приписываемая эффективность. В тексте аналогия между языком и технологией исследуется через призму семантики, синтаксиса, прагматики и грамматики. Подчеркивается важность такого расширенного прагматического подхода перед лицом новых технологий, который демонстрирует высокую степень самоактивности, большее количество способов взаимодействия между физическими и цифровыми объектами, а также растущую интерактивность при взаимодействии с людьми и факторами окружающей среды. Они могут быть более адекватно поняты, концептуализированы, а также спроектированы как социотехнические совокупности распределенной агентности между людьми, машинами и программами.

Ключевые слова: Техническая практика; Материальная агентность; Речевые акты; Лингвистическая прагматика; Эффективность; Смысл; Интерактивность; Цифровые объекты

Для цитирования: Rammert, W. Doing Things with Words and Things: A Social Pragmatist View on the Language–Technology Analogy // Technology and Language. 2024. № 5(2). P. 154-177. <https://doi.org/10.48417/technolang.2024.02.12>



This work is licensed under a [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/)



TECHNICAL PRACTICE: SPEAKING THROUGH THINGS IN THE LANGUAGE OF EFFICACY¹

Technical knowledge is based on practical action and seeks more effective action. It differs from scientific knowledge in that it is ultimately not about recognizing and explaining abstract and isolated effects but about designing and creating concrete operations, objects, and interactions between them that have to function efficaciously in a sociotechnical context. This fact applies to the matching of screws and threads and their relations to national standardization systems as well as to the combination of partial material and digital technologies into complete global manufacturing, traffic, or communication systems. Technical practices require scientific knowledge about things and their causal links, for example, about the hardness of steel, its interaction with concrete, or the loss of information during signal transmission via the copper or fiber-optic cable. However, the use of this abstract knowledge depends on specific situations and is embedded in the thousand-fold rule-based knowledge of the respective technical sciences.

In the *first* instance, this pragmatist view of technical knowledge as technical practices that have been created, objectified, and compiled as rules and tools in the arsenals of technology and then work when enacted in use relations suggests the idea of developing a *pragmatics of technical acts* in analogy to the linguistic concept of pragmatics.² This pragmatist approach to technical knowledge would be analogously concerned with the capability, functionality, and efficacy of *technical acts* in certain cases and contexts in comparison to the competency, intentionality, and successful communication of the latter pragmatics of *speech acts*. The *second thrust* of this article can be summarized as follows: Even natural sciences make sense of the nature of things with ‘words’ when they use alphanumerical signs, visual representations, calculations, and lawlike formulations to give them a meaning that can be communicated and contested. The technological sciences, however, create new things, and, when doing things with *things*, they are at the same time doing *words* with things. They make significant statements with things; they invent new techniques and capabilities of effective expression, and they weave new sociotechnical textures by constructing media and technical infrastructures. Through things, the technological sciences are speaking in the language of “efficacy.”³

To propose a combined pragmatism/pragmatics approach to the ‘question concerning technology’ (Heidegger, 1977 <1954>) actually seems to carry owls to Athens since the productive making of and practical doing with things can be regarded as the

¹ This text is a translation of a modified and supplemented version of the key lecture “The Pragmatics of Technological Knowledge or: How to Do WORDS with Things” at the conference “Technological Knowledge” organized by the German Academy of Technical Sciences (Kornwachs, 2010). I thank Alfred Nordmann for his critical comments and Stephan Elkins from *Sociotrans* for final proof reading.

² The founders of a philosophy of language that induced the development of linguistic concepts like semiotics and pragmatics are Peirce (1970), Wittgenstein (1953), and Austin (1962).

³ Cf. the comparisons of the Western and Chinese concepts of “efficacy” as a direct strategic act of an autonomous agent (subject) or as an indirect transformation of different agencies over a long time (Jullien, 2004; Hui, 2016).



core of the modern technological sciences.⁴ Three ideas and interests, however, give this new undertaking of a social pragmatist theory of technology its particular appeal and justification:

First, the paper is intended to counteract the overemphasis on the semantics of technical knowledge that accompanied the sweeping scientification of the engineering arts. The drive to win the reputation of the natural sciences also strengthened the belief in the perfect programmability of complex technical systems and displaced the non-explicit knowledge that is objectified in techniques and technological artifacts, stored in the habits of the respective communities of practice, and incorporated in the specific technical practices themselves.⁵

I argue that the linguistic concept of pragmatics will redirect the focus back to the practical activities during various phases and places when the stored technical knowledge is enacted. This process comprises a) the designing, planning, constructing, and programming of specific functionalities; b) the development and testing of technically effective devices and user-friendly interfaces; and c) the integration of these functionalities, devices, and interfaces into existing production facilities and company structures, into large technical infrastructures, and into social, ecological, and economic contexts of use.

Second, the reasoning presented here seeks to overcome the self-imposed limitations of a dualistic view⁶ of technology and society. According to this dualistic view, the making of technical things has nothing to do with social action and meaning-making; and vice versa, social action and doing things with words has nothing to do with doing things with real technical things. This paper argues that the social pragmatist view makes it possible to understand “the technical construction as part of the social construction of reality” (Rammert, 2006/2024, chapter 3). In consequence, the meaning (“words”) of a new technology (“things”) develops from three kinds of social interaction: when people or communities are doing things with words a) by naming, classifying, and negotiating the attribution of meaning to things; b) by experimenting with things, producing artifacts, installing physical and digital infrastructures; and c) by using and repurposing things and thereby creatively and constantly changing both how they work and what they mean.

Third, I argue that the combined view of philosophical pragmatism and the linguistic concept of pragmatics might also facilitate the interdisciplinary collaboration between technical and social scientists, which is urgently needed in light of interactive “digital objects” (Hui, 2012), “artificial communication” with “generative language systems” (Esposito, 2017), “human-brain interfaces” that translate neural signals into sentences (Nicolelis, 2001), and “social robotics” (Breazeal et al., 2016; Muhle, 2023).

⁴ Other concept transfers can be found under the names “material semiotics” (Latour, 1987; Law, 2009), “new materialism” (Barad, 2007; Lemke, 2015), “performance theory” (Mackenzie & Millo, 2003), or “narratology” (Czarniawska, 2004; Joerges & Czarniawska, 1998).

⁵ Cf. Hickmann (2001) on the difference of techniques, technologies, and technosciences and Rammert (1999a) on the three modes of technization: habitualization of human corporal and cognitive activities, mechanization of physical operations, and algorithmizing of symbol processing.

⁶ Cf. the critique by Rorty (1991), also Rammert (1997) and Schulz-Schaeffer (2000).



Thanks to modern information, communication, and sensor technologies, actual technical systems and environments are designed as more or less ‘autonomous agencies’ (Rammert, 2008a). They are equipped with software programs of ‘intelligent behavior’ and personally tailored to their users to interact with them. The elements are ‘loosely’ rather than ‘tightly coupled’ (Weick, 1988; 1995). Sociological models of social interaction and social institutions lend themselves to the construction of such complex cyber-physical or “socionic systems.”⁷ And vice versa, if nearly all kinds of social agency in society are increasingly distributed between different entities such as people, machines, and programs, then the social scientific analysis as well as the institutional design of society’s infrastructures require the knowledge and experience of the engineering and planning sciences.

The paper is divided into two parts. The next section introduces some problems of the nature and status of technical knowledge and the idea to approach them with a fresh view that combines philosophical pragmatism with the linguistic concept of pragmatics and draws on an analogy between speech acts and technical acts. First, I ask how the pragmatic aspects of technical knowledge were repressed and what problems the technical sciences have encountered as a result. This is followed by outlining the social pragmatist perspective on technology and the linguistic concept of pragmatics to show how they could be used to address both the problems of doing technology and of material agency. The third section exposes four characteristics of the actual technostructure that challenge the technological sciences to adapt their knowledge: self-activity, complexity, heterogeneity, and ubiquity. The fourth section uses the question of “technology-in-action” (Rammert, 2006; 2008a) to demonstrate that the approach to combine the pragmatist perspective on technology with the linguistic concept of pragmatics is especially fruitful for analyzing and shaping the actions and interactions of technical objects in various sociotechnical constellations. The article closes with a short summary and some conclusions on the pragmatics of technical skills, as they result from the previous considerations on sociotechnical agency. These considerations relate to the practice of engineering education as well as to the projects of interdisciplinary research and development.

PROBLEMS WITH THE NATURE OF TECHNICAL KNOWLEDGE AND THE PRAGMATISM/PRAGMATICS APPROACH

Technical Sciences: From the Recognition to the Adaptation Problem

Two problems affect the technical sciences with regard to the value and usefulness of their knowledge: first, the problem of recognition, which has led to a reinforced scientification and a suppression of the pragmatic character of technical knowledge, and second, the problem of adapting to quickly changing structural requirements. Addressing these two problems suggests a return to the pragmatic core of the technical sciences and

⁷ See the early inter-disciplinary DFG research program “Socionics: Inquiry and Design of Artificial Societies” (1999–2005) that connected computer science and sociology (Malsch, 1998; Rammert, 1998; 2012).



a fine-grained analysis of the pragmatics of technical agency.

The recognition of technical knowledge as socially highly valued and later as scientifically founded knowledge is an old problem that has persisted from Antiquity to modern science at the end of the 19th century. In ranking forms of knowledge, technical knowledge always remained in last place. In ancient society, the skills of master builders, lawmakers, and surveyors were considered necessary and useful for the domestic, urban, and military economy. Nevertheless, it was not granted the status of well-founded knowledge. Excellence was sought and seen in philosophical knowledge. From medieval to modern times, this hierarchy of forms of knowledge was largely maintained at the universities: Religion ranked before philosophy, philosophy before the natural sciences, even when the latter slowly detached themselves from theology and philosophy after the Enlightenment. The technical and engineering arts ranked at the bottom of the status pyramid. Technical knowledge was passed on outside the universities at special military, mining, and craft schools and at the academies of architecture, construction, and fine arts. The trend has reversed since the end of the 19th century, when the technical arts began forming into technological disciplines and transforming into sciences based on axiomatic theories. They formalized practices, simulated processes, and mathematized relations, following the example of physics. This ‘catch-up scientification’ visibly advanced the technical arts to academic sciences, also in terms of social recognition. In the industrial modern age, the secular belief in technical and scientific progress has ranked far ahead of religious beliefs. Engineering schools and construction academies were renamed and upgraded to the status of ‘technical universities’ with the right to award doctorates and habilitations. At present, technical knowledge no longer has any problems of recognition in the system of sciences. A look at the current technologies of the computer sciences, genetic engineering, or material sciences as well as at the high reputation of institutes and institutions of technology (MIT, CalTech, EIT, acatech) even suggests that the science-based production of technologies has risen to become the dominant model of research.

However, the successful solution to the problem of recognition through consistent scientification of technology has caused some barely noticed side effects: The orientation towards the model of natural science has pushed the pragmatic basis of technical knowledge into the background. Abstract knowledge has displaced the value of concrete experience; mathematized relationships have become more important than case-related knowledge of rules; and system designs purged of context ignore the art of ‘piecemeal engineering’ and the intuitive handling of “not well-structured problems” (Star, 1989) that, in complex sociotechnical constellations, cannot be neatly separated into parts. As long as technical systems can be relatively easily encapsulated and delimited from the natural and social environments and as long the subsystems can be integrated internally in a stable hierarchical manner, abrupt changes in the environmental dynamics only resulted in isolated problems at most. However, the more permeable the boundaries between the systems become and the more the technologies themselves approximate the model of human agency, the more urgent it becomes to return to the constructive and active aspect of technical knowledge and its relationship to the pragmatic conditions that make it successful. Otherwise, the one-sided view of technical knowledge will lead to systematic problems in adapting the technological sciences to the future emerging digitized technostructures of society.⁸

⁸ Cf. Nassehi (2024) for an operational and system theoretical approach to the digital society.



The View of Philosophical Pragmatism

Technical knowledge is generated ‘in pragmatic terms’ (“in pragmatischer Hinsicht”) and has to prove its efficacy in a practical context. The focus is on ‘shaping the world’ (“Weltgestaltung”), not explaining the world, and this kind of knowledge must be applied to ‘real people’ (“wirkliche Menschen”) and ‘for use in the world’ (“zum Gebrauch für die Welt”) and has to take into account the changing states of the world and the purposes of life, as Immanuel Kant’s (1778/1998) *Anthropology from a Pragmatic Point of View* (*Anthropologie in pragmatischer Hinsicht*) already stated. It is also remarkable that he mentioned technique/technology (“Technik”) in the context of the terms “habits,” “methods and machines and among these the distribution of work.” If one sees technical knowledge mainly under the aspects of perfecting it toward the scientific ideal and achieving complete causal explanation and total mastery, it is easy to lose sight of the gap between ‘pure’ semantics and ‘dirty’ pragmatics: The purified parts of technological knowledge, which of course have contributed to the progress of the modern technological sciences, must ultimately be incorporated into the messy end–means relationships of concrete environments.⁹ As *real* constructs, they must be fitted into the variety of different technologies that actually exist and operate locally. As *practical* constructs, they are by definition blended with economic, political, and cultural purposes. All aspects, practical skills and technical experience, implicit knowledge and experimental testing, are therefore not residues of the imperfection of technical knowledge that need to be remedied; on the contrary, they are and always will be necessary components of the technological sciences in spite of any scientification.

The pragmatist view was systematically developed in American philosophy. Philosophical pragmatism, as elaborated by Charles Peirce, John Dewey, and George Herbert Mead, was essentially a response to continental rationalism and idealism. It begins with the *primacy of practice* in dealing with questions of thought, knowledge, and cognition. It emphasizes *solving concrete problems*, such as how certainty can be established under specific conditions, over abstract questions of truth. Dewey’s pragmatism especially refers to the *principle of experimental action* when solving theoretical and practical problems (Dewey, 1925, chapter 1; Hickman, 1990, p. 60).

The proponents of pragmatism thus turn the tables of knowledge around: They take the approach of practical and everyday problem solving and the model of laboratory and engineering sciences as a model for philosophical thinking. They use the pragmatics of technical knowledge as a basis and model for solving problems of scientific, linguistic, and social theory. Truth, knowledge, and meaning are based on practices of ‘making true,’ of solving concrete problems, and of observing the interactive behavior between bodies. The concept of an object is not defined in an essentialist way but on the basis of operations to be carried out and as the epitome of the consequences that result when the object is acted upon in a certain way (Peirce, 1907/1970 and Dewey, 1929, chapter V). The meaning of a word can be derived neither from an intrinsic sense of the word nor from its structural position; rather, it emerges from the context of cooperation and physical

⁹ Cf. Mitcham (1994) for a detailed history of philosophical thinking about technology and a critique of Bunge’s (1966) outdated view on “Technology as Applied Science.”



interaction in which the participants both expect and observe what consequences a gesture or a sign has for themselves and the other (Mead, 1934/1968, p. 155). Just as an *interobjective* order of causes and effects is inferred from the observed interaction between two physical objects in experimental systems (Mead, 1932; Rammert, 1999b), George H. Mead's social pragmatist theory derives the emergence of an *intersubjective* meaningful order from the interaction among two bodies equipped with senses and organs (Mead, 1934/1968, p. 72; Joas, 1985).

This pragmatist perspective¹⁰ – despite all the differences between the respective elements under observation – not only places the production of technical artifacts and the generation of moral facts on a common basis, thereby rendering them comparable in terms of identity and difference, but also draws attention to the trying, tinkering, and testing that is necessary to transform technological knowledge into concrete devices, real machines, and operational digital programs. It examines the experimental trial and error between intended effects and observed adversities, between explicit means and implicit possibilities that act like a “mangle of praxis” (Pickering, 1995). In an ongoing process of “re-configuration” (Suchman, 2007), these experimental interactivities mutually create the respective competencies of people and the respective capacities of technology of an epoch. They form the glue between artifacts and knowledge that make up our society and its technologies. Pragmatism thus enables a new perspective to overcome the dualism of technology and society. It also intensifies the cooperation of the technical and social sciences, which are fighting on different fronts for the same cause: to advance “the substance of things hoped for” (Rorty, 1999, p. 27). Pragmatist inquiry is always guided by the demand for practical solutions in order to develop methods and programs as well as to create new combinations, insights, and instruments. Technical knowledge that arises in the form of technical rules and tools refers to the aspects of both the cause–effect relations between interacting entities in terms of efficacy and the intended end–means relationship in terms of expediency.¹¹

The Linguistic Concept of Pragmatics

Pragmatics is first and foremost an area of linguistic language theory. While *semantics* deals with the meaning of words and *syntax* with their position in sentence structure, *pragmatics* is concerned with meaning-making through practical use in various contexts. Since Wittgenstein's (1953) philosophy of normal language, scientific interest has shifted from the rules of grammar to the practices of speaking. The meaning of words cannot be understood without understanding their practical use in certain contexts and under certain circumstances. People do not apply the rules of grammar in the well-structured way that computer programs do.¹² Even generative artificial intelligence programs work with statistically measured relations to other words or pictures instead of

¹⁰ For a similar idea for blending the pragmatist theory of action with the analytical theory of social mechanisms, cf. Gross (2009).

¹¹ For similar interpretations, cf. Hubig's, version of pragmatism (2006) and the ones that have been combined with hermeneutical and phenomenological views (Ihde, 1990 and Verbeek, 2005).

¹² Cf. Rammert et al. (1998, “Knowledge Machines”) on the limits and failures of knowledge-based expert systems.



following grammatical rules. Therefore, they need human assistance for training, debugging, and abolishing errors in semantics. When people use words in everyday life, they are producing meanings and patterns that can be reconstructed as rules only in retrospect.

John L. Austin’s lecture on “Words and Deeds” (1952–54) was groundbreaking for the theory of language because he moved this aspect of doing when uttering words to the center of the entire humanities and social sciences. The title of his legendary book “How to Do Things with Words” (Austin, 1962) could also be translated as ‘making things happen with words.’ In the case of performative utterances, when speech is intended to achieve something, uttering words becomes an action. Sentences not only *say* something; they also *do* something. Such “speech acts” do something by giving it a name, for instance, they can marry two people, or they can sentence someone to some form of punishment. Their function is not merely to state what *is* the case but also to practically bring about what *will be* the case. These “performative speech acts” neither describe nor represent reality; instead, they construct reality with all its social consequences, for spouses, heirs, debtors, or prisoners.

This performative aspect of ‘doing things with words’ is actually self-evident if we compare the pragmatics of language with the pragmatics of technology. One aspect of the analogy, however, could shed new light on the theory of technical action and knowledge. We can see it when we slightly modify Austin’s title into ‘How to Do *Words* with Things.’ ‘Saying something with things’ means that, whenever things are made or something is done with real things, words are also formed, statements are made, and above all new meaning is created. Similar to speech acts, *technical acts* create a universe of technical forms of meaningful expression. A language of technology emerges with its own grammatical rules, its own dictionary that collects the semantical translations, and its own syntax of functioning technical combinations. This means that technical agency cannot just be understood as a meaningless ‘how to do things with things,’ which is dressed up only subsequently by attaching cultural meaning as one would put on an item of clothing. It implies much more, namely,

- that technical agency makes things and meaningful statements *simultaneously*,
- that it *objectively* and *meaningfully* orients, restricts, or enables the actions of others, and
- that the *real texture* and *architecture* of technical systems co-shape the “constitution” of society and the culture of an era much more than the attributed symbolic meanings.¹³

Pragmatics as a concept and program for the analysis of technical knowledge and practice thus always implicates two aspects at the same time:

First, one should examine technology as “objectified culture” (Simmel, 1900/1983, p. 96) and technical acts, artifacts, and systems as a form of human expression that is similar to speech acts, words, and language structures. The philosopher and anthropologist Ernst Cassirer had already regarded technology as a fourth “symbolic

¹³ Cf. Langdon Winner (1980), who uses an analogy between the political and technical “constitution,” and Bruno Latour (2005), who speaks of “assemblies” of both people and things.



form” alongside myth, logos, and art. He argued that through the form of technology people express themselves in the” medium of efficacy” (“Medium des Wirkens”; Cassirer, 1930/2012, p. 24¹⁴; “language of efficacy”, Rammert, 1999a, p. 276). A pragmatics of technical knowledge then examines technology as a ‘technical act’ from the point of view of the generation and purposeful use of artifacts in specific contexts and under historical circumstances. It is only the use of a thing or a configuration of things in a certain constellation that ultimately turns it into an expedient means and meaningful tool. In this wake, technical rules – in which the knowledge of efficacy and finality is inscribed – can gradually establish themselves as a set of rules similar to grammar. The pragmatist view on technical construction can learn from the pragmatics of language that technical acts can also be organized in concrete expressive and objectified forms in a similar way that speech acts and syntactic constellations of words can. Another analogy exists between linguistic grammar and the “technological archive” (Groys, 1997): Both are collections and frameworks of successful practical solutions that are generalized in terms of rules and tools to fall back on. They are themselves also ever-changing storage and memory places that are the basis for the construction of new acts and artifacts, different configurations, and extended sets of rules and tools when new problem situations require creating and testing new answers.¹⁵

Second, one might find new ways of expressing expectations by doing things and thereby create new possibilities for action through technical construction in the same way that thoughts arise when speaking.¹⁶ This similarity challenges us to inquire and test existing technical acts and technological designs in practice and to revise them until they work (Dewey, 1938, chapter VI). It is crucial to adapt them to the technological, economic, ecological, and social conditions that will be encountered in future situations of use. Methods have to be provided that are sensitive to the interdependencies of these conditions to develop innovations that are not only technologically up to date, ecologically sustainable, and economically profitable but also widely accepted by people.¹⁷

Looking at “technical construction as part of the social construction of reality” (Rammert, 2006/2024, chapter 3), we can characterize the gradual development of the objectification of both knowledge (language) and technical action (technology) in an analogous way:

At the *first* stage of development, language begins with “indices” (Schütz & Luckmann 1974, pp. 326-331; Berger & Luckmann, 1966/1969, p. 38) for something which – in the practical context of *interaction* – turns traces, events, or gestures into meaningful signs and symbols. Accordingly, technology begins with “samples” and

¹⁴ See the entire quote: “All mental handling of reality is bound to this double act of ‘grasping’ – ‘comprehending’ reality in linguistic-theoretical thought and ‘gripping onto it’ through the medium of efficacy. This is true for both mental and technological forming” (Cassirer 1930/2012, p. 24).

¹⁵ Cf. Leroi-Gourhan’s (1964/1993) seminal work on the similarity of relations between techniques, language, and the arts and the parallel developments of “Gesture and Speech” in history.

¹⁶ See Heinrich von Kleist (2009) “On the Gradual Formation of Thoughts while Speaking.”

¹⁷ On the implications of this pragmatist view for “Innovation Society Today,” cf. Passoth & Rammert (2018; 2023).



“problems,” which in the practical context of “inquiry,” tinkering and “experimental interactivity” (Dewey, 1938, chapter I and III; Rammert, 1999b, p. 291) turns things, forms, and forces into effective technical means, operations, and installations. They remain substantively tentative, temporally episodic, and spatially localized.

At the *second* stage, the uses of signs are transformed into “products.” They are decoupled from their initial places and purposes of origin but remain tied to typical situations as significant signs and symbols. In the case of technology, the designs then acquire the status of ‘tried and tested installations’ for specific contexts. Their descriptions are collected in books on technology and machines as detailed procedures, effective artifacts, or useful probes. In their function as exemplary solutions to problems, they are detached from their places of origin in mills, mines, or irrigation systems. However, at the stage of simple or combined machines, they still require embedding in the professional and customary repertoire of mechanics or engineers.

At a *third* stage, the language elements are completely detached from their initial contexts. They form their own “sign system,” a well-structured grammar of language with explicit rules of use and a semantic collection of a vocabulary. As in the case of language, a universalized archive of technology emerges in a similar fashion. Its corpus consists of all the specific rules that originate from successful technical solutions. It combines them to a ‘state of the art’ body of decontextualized principles and effective “schemata of technization” (Rammert, 1999a, p. 277).

This parallel modeling of linguistic and technical action opens up a research program that analyzes technology – analogous to language – under the aspects of semantics, pragmatics, and syntax in a more refined and systematic way. The research program of pragmatics in particular identifies the single ‘technical act’ as the smallest unit of investigation in the overall context of technical action. This concept allows a more fine-grained analysis of the making *of* technology, the making *with* technology, and the *active* participation of technology in the respective situations.¹⁸

At the same time, the grammatical perspective allows us to understand technology as an arsenal of objectified technical acts as well as an archive of encoded schemata of mechanization. The use of this recorded and stored potential – this is also a lesson from linguistic pragmatics – always requires “enactment” in practical situations.¹⁹ Just as every uttered sentence cannot be derived from the rules of grammar alone but requires considering the situation, every act of technical construction remains a more or less appropriate utterance only under specific contextual conditions, despite any recourse to existing sets of rules or states of the art of technological expertise. The more manifold the contexts become to which technical action must relate, the more they differ (such as economic efficiency, sustainability, security, data protection), and the greater the range of settings in which such action is institutionally integrated (economy, politics, science, law, aesthetic design, etc.),²⁰ the more strategically important such a program geared

¹⁸ Cf. Rammert (1997) on “new rules” and Rammert (2008b, p. 344) on the theory and method of “technography.”

¹⁹ Cf. Weick (1995), Schulz-Schaeffer (2000, p. 64), and Orlikowski (2002) on the concepts of enactment and sense-making.

²⁰ Cf. the contributions on “intelligent objects” in Herzog and Schildhauer (2009) and on future



toward a detailed analysis of technical knowledge becomes that is based on technical acts and experiential situations.

TECHNOLOGY IN ACTION: CHARACTERISTICS OF CHANGING TECHNOSTRUCTURES AND THE CONSEQUENCES FOR THEORIZING

So far, we have dealt with the production and use of technical artifacts by pointing out the experimental and meaning-expressing character of the technical act. Now we shift our focus from the generation, design, and installation of technology to the functionality and *agency* of the technical objects themselves. The new social pragmatist view in conjunction with the pragmatics approach also facilitates the analysis of changes in technological objects and their external activities and interactions. For a long time, technical objects could safely be understood as passive and fixed means that silently and constantly fulfill their function as unchanging components. Current technostructures of production, transportation, and communication, however, exhibit four conspicuous changes:

The *self-activity* of technical objects has increased significantly and in more diverse ways than in the traditional forms of mechanization and automation. The objects operate as “agents” and are designed as bundles of agencies or “autonomous” systems (Wooldridge & Jennings, 1995, p. 116). Concepts of instrumentality are no longer sufficient for dealing with them. The higher level of autonomous activity also requires an analytical concept that is better capable of taking into account the reciprocity between humans and technology. For this purpose, I proposed the concept of “interactivity” (Rammert, 2008a, p. 71) to describe the relation between these two sides.

The *complexity* of the objects’ technical configuration has grown rapidly. The number of material components and their coupling relations – for instance, in a car – has not only been multiplied; the number of functional systems for ignition, stabilization, computer-mediated optimization, as well as for connection with control and information systems in the environment has also increased. The focus has already shifted from the simple and single artifact and its modular aggregation to a multi-layered and complicatedly nested cyber-technical system.

The *heterogeneity* of the elements that must be considered in the design of such highly complex systems requires a more comprehensive and particularly more diverse conceptualization and modeling: Such designs must not only pay attention to the proper functional alignment of material systems and systems of human action but also align digital systems and environment-sensitive systems in a functional relationship that takes their respective characteristics and different activities into account.

The *ubiquity* of technical services is taking a leading role: Technical infrastructures are increasingly being set up in such a way that all artifacts can be accessed everywhere in the world. from anywhere at any time. The cell phone and “ubiquitous computing”

constellations that can be expected to change in the wake of emerging artificial intelligence technologies (Lee & Chen, 2021).



(Greenfield, 2006; Fleisch & Mattern, 2005) are examples of this shift from stationary to mobile technology use.

The technological sciences already feature some important conceptual responses to these changes, for instance, a shift from artifact-centered to process- and systems-oriented science (in chemical engineering, mechanical and plant engineering, or in architecture and urban design). Above all, the greater self-activity of objects now suggests an additional widening of the perspective on technology that concentrates on the activities and interactions of objects in heterogeneous sociotechnical constellations.

Objects in Action: Dimensions and Levels of Self-Activity

Compared to purely mechanical processes, technical objects are increasingly achieving higher levels of autonomous activity. The latest information and communication technologies indicate a qualitative change in the four dimensions of motorics, action, sensorics, and informatics. Technical objects are being transformed from stationary devices into mobile agents, from hard-wired artifacts into programmable agents. Sensor technologies transform what was once a stubborn device into reactive agency. And the methods and means of computer science can elevate a technological object from the status of a passive instrument to the role of pro-active agent.²¹ Taken together, these tendencies toward increased self-activity suggest that the instrument and machine concept of technical means should be replaced by a more appropriate concept: that of technical agents and systems of agencies.

The concept of pragmatics can also be applied to how we assess the activity levels of technical agency. Initial analyses of the new production, programming, and information technologies have proven the following classification of activity levels to be useful for the time being (Rammert, 2012, p. 96):

1. passive: objects that are moved or changed in any way from the outside in order to have an effect (tool, index card, price tag)
2. active: object combinations that perform certain operations in the four dimensions automatically (drilling machine, punch card, barcode scanner)
3. re-active: object combinations with feedback mechanisms for simple adjustments (sensor-controlled drilling machine, help agent, RFID checkout)
4. pro-active: distributed systems that coordinate themselves by means of mutual coordination (autonomously cooperating rescue robots, multi-agent systems, ‘smart objects’; ‘Internet of things’)
5. trans-active: intelligent systems that independently reflect on and change means–ends relations with regard to their own actions, others’ actions, and the aggregate results of action (so far only human teams or hybrid sociotechnical constellations).

As long as the activities fall into the first two categories, the traditional concepts are sufficient for analysis and modelling. Only once the qualitative changes mentioned above take hold do we see a shift toward the more sophisticated levels of classification. A social pragmatist concept of distributed agency makes the differences between the

²¹ Cf. Wooldridge & Jennings (1995, p. 117) and Rammert (1998, p. 99).



levels visible and their consequences addressable.²²

Objects, Intra-Action, and Interactivity: Increases in Internal Complexity and Heterogeneity at the Interfaces

Once we have become familiar with the social pragmatist view and its specification along the lines of the linguistic concept of pragmatics, two further consequences become apparent: the first one concerns the change in the internal structure of the technical systems and the second one the interchanges at the external boundaries, the so-called interfaces of the technical systems where technical acts and human acts intersect.

The history of the engineering sciences is no longer just about single artifacts or simple machines such as levers, wheels, and inclined planes. Its extended objects are the effective combination of individual artifacts into complex machines and ultimately their optimal configuration into large plants as well as into entire production and distribution facilities. The standard view of technical systems actually includes the whole hierarchical nesting of technical infrastructure systems with large technical production systems and how these combine with drive, processing, transportation, and control subsystems. The latter are again made up of smaller subsystems and further elements.

However, when the fixed relationships of such systems are set in motion by the increasing self-activities of individual objects and especially by the activities of programmed and mobile digital objects, the encapsulated hierarchical systems of machinery are transformed into *open systems* with internally interacting subsystems. Tightly coupled process systems are then transformed into *loosely coupled networks* of various tightly coupled subsystems with buffers and variable possibilities for reacting to dynamic environments (Perrow, 1984). The more objects, the more dimensions of object activity; and the more different types of objects and their relationships come into play, the more appropriate it is to use the vocabulary of action and interaction for describing the individual and overall behavior of such technical objects and systems. The concepts of instrumental function and fixed structure can no longer capture to a sufficiently differentiated degree the internal complexity of interrelated multiple agencies.

If technical objects gain greater scope for autonomous behavior, the requirements for *interfaces*, especially those between people and technology, must change too. When people use tools or operate machines, we can describe this as *instrumental use relation*. The shape of a handle or the design of levers and cranks reflect the intended mode of employing the technical device. When machines and systems are set up for intended processes, when even the computers controlling them are variably programmed, and when digital objects, such as search engines, become active and search relatively independently in files ('re-active'), then the direct instrumental relationship of use is transformed into a *mediated instructional-triggering* relationship.

If technical systems enter into a dialog with the user asking him or her for input or offering help on their own initiative ('pro-active') and these objects become active as communicative agents, then the interface develops into an *interactive-communicative*

²² Cf. Rammert & Schulz-Schaeffer (2003) and chapter 7 on "Technology and Agency" in Rammert (2024) and Schulz-Schaeffer & Rammert 2023 for a revised version.



relationship that ties in with the signs, sounds, gestures, and other media that constitute interpersonal interaction. Writing and reading, speaking and listening, pointing and following, touching and triggering within the context of expedient symbolic framings and iconographies each form the activities between the two sides in a specific way. The greater the degree of self-activity of the technical elements, the more the relationship between people on one side and physical or digital objects on the other shifts from instrumentality to interactivity. In comparison to interpersonal interactions and internal technical intra-actions, *Interactivity* refers to the particular form of relations between humans and technology. A para-social communicative process emerges by the reciprocal nature of the activities that engenders a sense of trust and by the capacity of the technical media to show reactions and give answers in real-time and in dialogical form.

This technical mediation of control and communication also means that the interface is no longer tied to the local workstation of a machine nor to the location of a plant or an archive. Via the various media, the interface can provide access to the functionalities of production, ordering, and information systems from nearly any location. These are the new characteristics that were referred to above when speaking of the shift from the locality to the *ubiquity* of technology use (Greenfield, 2006). In addition, the increase in design possibilities apparently strengthens the performativity and connectivity of the interface with nearly all gadgets anytime and anyplace. These shifts from instrumentality to agency, from material to digital processing, and from activity to interactivity are increasingly developing dynamics of their own and require a more detailed approach when we analyze the agency and inter-agency of material and digital objects in such “synthetical situations” (Knorr Cetina, 2009). This is because technical acts take place as intra-actions within a cyber-physical system; furthermore, technical acts transform machine language into pictures, texts, speech, and sounds when they react to human speech acts via voice or keyboard.

Distributed Activities in Sociotechnical Constellations

These developments will change the nature of technical design in two directions: First, the object of design will grow in size and complexity. Second, it will increase in the variety and heterogeneity of the entities and relations involved. These trends require an approach that is also capable of decomposing complexities into smaller units and distinguishing between diverse relations of inter-agency.²³

Let us briefly look at automobiles as a case in point. For a long time now, the design of cars has not been limited to a focus on engines and mechanical engineering only. Rather, it has involved the design of an entire vehicle system including mechanical, hydraulic, electrical, and electronic subsystems. The latter are equipped with interfaces to external information and communication systems. In view of the increase in the number of human–machine interfaces and their change from directly transmitted instructions (turning the steering wheel, pressing the brake pedal) to computer-mediated, video-assisted, or even the automated control systems of autonomous driving, the relationship of interactivity with material and digital objects as well as the *intra-actions* between them

²³ Cf. Law & Mol (2002) for a similar view on heterogeneity and complexity but a different approach.



has become more important. This has given rise to a broader view of the whole driver–vehicle system. The more interactions take place between driver, vehicle, and technical infrastructures through radio contacts, relays, and traffic control systems, the more appropriate the concept of large technical systems (Mayntz & Hughes, 1998) seems to be to properly comprehend the driver–vehicle environment. If we also include other technical systems of mobility – say, the mobility behavior of people and the programs of traffic design and management – we arrive at a highly abstract but differentiated view of the entire *sociotechnical constellation* of traffic with all its interactivities between people and material or digital objects, the intra-actions between the latter, and all its interdependencies between the various installations and institutions of the entire traffic system.²⁴

Some things will, of course, not change. Technical knowledge will continue to be applied in a proven manner to the design and optimization of machines, equipment, and systems. It will also be extended to the design of electronic gadgets, computer programs, and architectures of intelligent systems. And it will continue to be applied to the design of human–machine interfaces. However, in view of the differences between material objects (“hardware”), digital objects (“software”), and human actors (“humans”), it is no longer sufficient only to aggregate the bodies of technical knowledge, for instance, of “mechanical,” “electrical,” “software,” and “human engineering” in a modular way. As technical complexity increases, so does the need for a multi-level system engineering that deals with the interactive coupling of the various parts and levels. As the sociotechnical interactivity between humans and technologies is changing, this requires us to adopt a more comprehensive view that connects the two modes of meaning-making through technical acts and through speech acts.

The social pragmatist approach of distributed agency (Rammert, 2024) can contribute to this endeavor when it asks the overarching question: How do we distribute the activities between the various agencies of humans, machines, and programs to ensure efficient, sustainable, and safe mobility? This view then leads to more detailed questions: For example, how much of keeping a distance or hitting the brakes should be left to the car driver, delegated to the automatic cruise control, or a traffic guidance system that is connected with the environment via sensors and video cameras? Or how should the transparency at the interfaces be designed: in favor of the individual user of a central traffic control via cameras, or in favor of opportunities for external intervention at the interfaces?

If we approach these questions from the particular perspective of pragmatics, we gain a more differentiated view of these activities and their distribution. For instance, a common response to human shortcomings in the operation of technology is to replace human with technical agency, say, to avoid accidents. The irony in this is that it induces a trend toward increasing the complexity of technical systems via ever more sophisticated automation while this increasing complexity makes these systems more vulnerable to human shortcomings, be it in their design or operation. A pragmatist approach would break with this blind reflex of responding to insufficiencies by increasing complexity in

²⁴ Cf. Rammert (2002 and 2024, chapter 8) on distributed intelligence among traffic systems.



favor of a more reflexive response by adopting a more dynamic and distributive approach that examines the respective shares of agency and balances intervention and automation depending on the specifics of the situation or the function of the task that the system is supposed to perform.

Such an analysis would allow us to more accurately assess the advantages and disadvantages of the respective agencies and make better use of a differentiated understanding in shaping the constellation. For instance, the disadvantage of human proneness to error is paired with the advantage of the human capacity for high flexibility in dynamic and opaque situations. By contrast, the stability and predictability of material systems is the basis for their security and efficiency, while it is also the source of catastrophes in the case of unexpected deviation or interference. Digital systems, on the other hand, are excellently suited for the construction and simulation of different designs of single objects and variants of entire production systems at almost no cost; yet, they are mostly blind to resistance from the environment or to the wear and tear of physical material systems, while they also come with often underestimated risks of abusive manipulation by unauthorized persons or criminal users.

Thus, the development of a social pragmatist view on technical action and knowledge that unites productive pragmatism with linguistic pragmatics not only proves to be an epistemic advancement but also to be a suitable approach to finding practical answers to the questions posed above. Finding such solutions requires constructive and coordinated action and the interweaving of technical, ecological, and social facts. This means that the existing skills from the relevant disciplines must be brought together, not only in theory but also, and above all, in practice by actively trying out and mixing the different methods, tools, and findings as well as developing a common view by translating between the specific languages and by co-constructing sociotechnical solutions that work in society.

THE PRAGMATICS OF TECHNICAL AGENCY

The pragmatist perspective has shown that overcoming the problems of heterogeneity and complexity requires continuous action and mutual adaptation. The black-boxing and fixation of technical agency only works to a limited extent. This is because ‘technology in action’ means – as I have argued – that technical objects actually feature higher levels and higher grades of agency. This agential capacity changes the objects themselves as well as their relations with each other and their relations with people at the interfaces. The complexity of the technical systems that make up the subject matter of the technological sciences is growing. The task is no longer mainly limited to the construction of individual machines, specific plants, or special kinds of buildings; it is rather increasingly extended to include the cyber-physical design of complete energy, production, and distribution systems and the integrative planning of urban and regional districts. To come to terms with complexity of this magnitude, the engineering sciences of the future will have to place greater emphasis on participative planning processes and multi-optional experimental engineering.

If the object of technological development is no longer just the spatially,



temporally, and materially fixed object but instead increasingly the mobile, self-active, and interactive object, then this activation of technologies also requires a different understanding of systems. The idea of blending the perspective of philosophical pragmatism on doing technology with the linguistic concept of pragmatics seems to be fruitful to grasp these shifts toward more autonomy, hybridity, and inter-agency of sociotechnical constellations. Through the pragmatics of technical agency, one first and foremost gains a new dimension for the analysis of technology beyond materialist approaches that emphasize the determinative nature of productive forces and neglect the dimension of meaning-making in shaping it. At the same time, the pragmatics approach also goes beyond cultural approaches that overstate the force of speech acts, discourses, and attributions of meaning and underrate corporally experienced and inscribed agency.²⁵ The core concept of technical agency concentrates on the observation, recognition, and theorizing of single acts, sequences of acts, and configurations of acts that simultaneously produce and reproduce meaning through the language of efficacy. This being the case, an appropriate modeling of technological systems must include the systems' capacities to change themselves via *intra-action* within the technical system and by means of *inter-activities* with people at its interfaces.

The whole complex of material, digital, and human inter-agencies takes on the character of a sociotechnical constellation that places different demands on the control and management of these heterogeneous systems. In the future, engineering sciences will therefore have to enrich themselves with logistical, linguistic, and social-scientific knowledge on the management of complex systems and on the governance of mixed systems to be able to balance the technical, legal, economic, and cultural aspects appropriately.

If the point is not only the effective design of the *material* objects and their intra-actions but also the design of the *symbolic* objects ranging from agent programs to architectures of distributed intelligent systems, then completely new mixed constellations will arise in which the activities of the various material, symbolic, and human agents will not only simply intra-act in a purely mechanical way; rather, they will influence each other via interactive relationships between humans and technology.

The appropriate management of this heterogeneity of sociotechnical constellations requires the technological sciences to have an extended concept of *socio-technical design* and a competence for *reflexive innovation* in which the distributed agencies and the expected interactions are already included in the processes of developing and testing technology instead of improving it after the fact through ergonomics, technology assessment, or user responses. Technical knowledge has to be enriched with psychological, aesthetic, linguistic, and sociological knowledge if it is to be appropriately understood as experimental knowledge of action in order to become fit for the specific task of designing interfaces of human–technology interactivity. In terms of the pragmatics approach, this means that human speech acts have to be translated into both digital and material acts and, vice versa, technical acts in turn into meaningful speech acts via observable material behavior or signs on a screen.

²⁵ Cf. Rammert & Schubert (2023) on the human and technical embodiments of the social.



There is no identity or symmetry between human programs and technical programs as, for instance, actor–network theory assumes with its reference to Greimas’ semiotics when stating that anything that can be in a speaker or subject position of a sentence can be the actor (‘actant’) both in a sentence and in material reality.²⁶ According to our pragmatics approach, doing things with things and doing things with words exhibit differences in their modes of efficacy (e.g., forcing things and people to do things versus influencing their inclinations or dispositions to do things. This approach refers to technical acts and speech acts as the smallest analytical units and enables a more comprehensive analysis of the meaning-making in different sequences, situations, and constellations.

We come full circle at the end. We began by looking at technology from the pragmatist point of view of doing things by applying skillful techniques, tools, or technological rules to achieve effects. From the linguistic perspective of pragmatics, I defined this doing of things as ‘technical acts’ that create constellations of efficacy and simultaneously give rise to processes of meaning-making. Then I demonstrated that the products – the technical objects themselves – actually become increasingly pro-active and, especially, interactive with users. The more the objects actively participate in bringing about the desired overall action, like traffic navigation or knowledge distribution, and the more humans enter into a relationship of meaningful interactivity with them, the more the efficacy and the comprehensibility of sociotechnical constellations are created through coordinated and distributed interagency between people, programs, and machines.

It might become apparent in the future that the new technologies will be made to function only through interaction between the technical acts of the manufacturers and the technical acts of the users, as is already sometimes the case in the niches of open forums on the Internet. Just as the music industry, the film industry, and the press are currently reconfiguring themselves, we might see the renewal of the design and development of entire technical systems as well. Regardless of where such transformation takes place now and in the future, both the pragmatics of technical agency and the levels and grades of autonomous action given to technical objects will be of particular importance. Blending the social pragmatist perspective with the linguistic concept of pragmatics, we are well equipped to observe and practically accompany the production and use of technology as an endeavor that combines meaning-making with words as well as with technical things at the same time. This will hopefully turn out to be a fruitful approach to study the new constellations once even the doing of words becomes the object of AI-generated texts and chats.

REFERENCES

- Austin, J. L. (1962). *How to Do Things with Words*. Clarendon Press.
Barad, K. (2007). *Meeting the Universe Halfway. Quantum Physics and the Entanglement of Matter and Meaning*. Duke University Press.

²⁶ Latour’s definition of an ‘actant’ (Latour, 1987, p. 84) refers to the Greimas & Courtès’ (1979/1983) “Semiotic and Language Dictionary”.



- Berger, P. L. & Luckmann, T. (1969). *The Social Construction of Reality: A Treatise in the Sociology of Knowledge*. Anchor Books. (Original work published 1966)
- Breazeal, C., Dautenhahn, K., & Kanda, T. (2016). Social Robotics. In B. Siciliano & O. Khatib (Eds.), *Springer Handbook on Robotics* [2nd ed.] (pp. 1935–1972). Springer. https://doi.org/10.1007/978-3-319-32552-1_72
- Bunge, M. (1966). Technology as Applied Science. *Technology and Culture*, 7, 329–347.
- Cassirer, E. (2012). *Form and Technology*. A. S. Hoel & I. Folkvord (Eds.), *Ernst Cassirer on Form and Technology* (pp. 15-53). Palgrave Macmillan. https://doi.org/10.1057/9781137007773_2 (Original work published 1930)
- Czarniawska, B. (2004). Metaphors as Enemies of Organizing, or the Advantages of a Flat Discourse. *International Journal of Sociology of Language*, 166, 44–65. <https://doi.org/10.1515/ijsl.2004.014>
- Dewey, J. (1925). *Experience and Nature*. Open Court Publishing Co.
- Dewey, J. (1929). *The Quest for Certainty*. Minton, Balch and Co.
- Dewey, J. (1938). *Logic. The Theory of Inquiry*. Henry Holt and Co.
- Esposito, E. (2017). Artificial Communication? The Production of Contingency by Algorithms. *Zeitschrift für Soziologie*, 46(4), 249–265. <https://doi.org/10.1515/zfsoz-2017-1014>
- Fleisch, E. & Mattern, F. (Eds.). (2005). *The Internet of Things. Ubiquitous Computing and RFID in Practice*. Springer. <https://doi.org/10.1007/3-540-28299-8>
- Greenfield, A. (2006). *Everyware. The Dawning Age of Ubiquitous Computing*. New Riders.
- Greimas, A. J. & Courtès, J. (1983). *Semiotic and Language: An Analytic Dictionary*. Indiana University Press. (Original work published 1979)
- Groys, B. (1997). Technik im Archiv. Die dämonische Logik technischer Innovation [Technology in the Archive. The Demonic Logic of Technical Innovation]. In W. Rammert & G. Bechmann (Eds.), *Technik und Gesellschaft. Jahrbuch 9* (pp. 15–32). Campus.
- Gross, N. (2009). A Pragmatist Theory of Social Mechanisms. *American Sociological Review*, 74, 358-379.
- Herzog, R., & Schildhauer, T. (Eds.). (2009). *Intelligente Objekte* [Smart Objects]. Springer acatech.
- Heidegger, M. (1977). The Question concerning Technology. In *The Question concerning Technology and other essays* (pp. 3–35) Harper Row.
- Hickman, L. (1990). *John Dewey's Pragmatic Technology*. Indiana University Press.
- Hickman, L. (2001). Tuning Up Technology. In *Philosophical Tools for Technological Culture* (pp. 8–43). Indiana University Press.
- Hubig, C. (2006). *Die Kunst des Möglichen I. Technikphilosophie als Reflexion von Medialität* [The Art of the Possible I. Philosophy of Technology as a Reflection of Mediality]. Transcript.
- Hui, Y. (2012). What is a Digital Object? *Metaphilosophy*, 43(4), 380–395. <https://doi.org/10.1111/j.1467-9973.2012.01761.x>
- Hui, Y. (2016). *The Question concerning Technology in China*. Urbanomic Media Ltd.
- Ihde, D. (1990). *Technology and the Lifeworld. From Garden to Earth*. Indiana



- University Press.
- Joas, H. (1985). *G. H. Mead: A Contemporary Re-Examination of His Thought*. MIT Press.
- Joerges, B. & Czarniawska, B. (1998). The Question of Technology, or How Organizations Inscribe the World. *Organization Studies*, 19(3), 363–385. <https://doi.org/10.1177/01708406980190030>
- Jullien, F. (2004). *A Treatise on Efficacy. Between Western and Chinese Thinking*. University of Hawaii Press.
- Kant, I. (1998). *Anthropology from a Pragmatic Point of View*. Reclam. (Original work published 1789)
- Kleist, H. von (2009). On the Gradual Formation of Thoughts while Speaking]. In *Selected Prose of Heinrich von Kleist*. (pp. 310–316). Archipelago Books.
- Knorr Cetina, K. (2009). The Synthetic Situation: Interactionism for a Global World, *Symbolic Interaction*, 32, 61–87. <https://doi.org/10.1525/si.2009.32.1.61>
- Kornwachs, K. (Ed.) (2010). *Technologisches Wissen*. Springer acatech.
- Latour, B. (1987). *Science in Action. How to Follow Scientists and Engineers through Society*. Harvard University Press.
- Latour, B. (2005). *Reassembling the Social. An Introduction to Actor-Network-Theory*. Oxford University Press.
- Law, J. (2009). Actor Network Theory and Material Semiotics. In B. S. Turner (Ed.), *The New Blackwell Companion to Social Theory* (pp. 141–158). Malden. <https://doi.org/10.1002/9781444304992.ch7>
- Law, J. & Mol, A. (2002). *Complexity. Social Studies of Knowledge Practices*. Duke University Press.
- Lee, K.-F. & Chen, Q. (2021). *AI 2041. Ten Visions for Our Future*. WH Allen.
- Lemke, T. (2015). Varieties of Materialism. *BioSocieties*, 10(4), 490–495. <https://doi.org/10.1057/biosoc.2015.41>
- Leroi-Gourhan, A. (1993). *Gesture and Speech* (“Le Geste et la Parole. Vol. 1 Technique et Langage”). MIT Press. (Original work published 1964)
- MacKenzie, D. & Millo, Y. (2003). Constructing a Market, Performing Theory: The Historical Sociology of a Financial Derivatives Exchange. *American Journal of Sociology*, 109, 107–145. <https://doi.org/10.1086/374404>
- Malsch, T. (Ed.). (1998). *Sozionik. Soziologische Ansichten über künstliche Sozialität* [Socionics. Sociological Views on Artificial Sociality]. Edition sigma.
- Mayntz, R., & Hughes, T. P. (Eds.) (1998). *The Development of Large Technical Systems*. Campus and Westview Press.
- Mead, G. H. (1932). The Physical Thing. Supplementary Essay 2. In *The Philosophy of the Present*. (pp. 119–139). Open Court.
- Mead, G. H. (1968). *Geist, Identität und Gesellschaft* [Mind, Identity and Society]. Suhrkamp. (Original work published 1934).
- Mitcham, C. (1994). *Thinking Through Technology. The Path between Engineering and Philosophy*. University of Chicago Press.
- Muhle, F. (Ed.). (2023). *Soziale Robotik. Eine sozialwissenschaftliche Einführung* [Social Robotics. A Social Science Introduction]. De Gruyter.



- Nassehi, A. (2024). *Patterns. Theory of the Digital Society*. John Wiley & Sons.
- Nicolelis, M. (2001). Actions from Thoughts. *Nature*, 409, 403–407.
<https://doi.org/10.1038/35053191>
- Orlikowski, W. J. (2002). Knowing in Practice: Enacting a Collective Capability in Distributed Organizing. *Organization Science*, 13(3), 249–273.
<https://doi.org/10.1287/orsc.13.3.249.2776>
- Passoth, J.-H. & Rammert, W. (2018). Fragmental Differentiation and the Practice of Innovation. Why Is There an Ever-increasing Number of Fields of Innovation? In W. Rammert, A. Windeler, H. Knoblauch, & M. Hutter (Eds.), *Innovation Society Today* (pp. 35-63). Springer VS. https://doi.org/10.1007/978-3-658-19269-3_3
- Passoth, J.-H. & Rammert, W. (2023). Fragmental Differentiation as a Diagnosis of Society: What Is Behind the Increasing Orientation Towards Innovation, Granularity and Heterogeneity? In C. Schubert & I. Schulz-Schaeffer (Eds.), *Berlin Keys to the Sociology of Technology* (pp. 129–159). Springer Nature.
https://doi.org/10.1007/978-3-658-41683-6_6
- Peirce, C. S. (1970). “How to Make Our Ideas Clear” and “What Pragmatism is”. In H. S. Thayer (Ed.), *Pragmatism. The Classic Writings* (pp. 79-101). Mentor Book. (Original work published 1907)
- Perrow, C. (1984). *Normal Accidents: Living with High-Risk Technologies*. Basic Books.
- Pickering, A. (1995). *The Mangle of Practice: Time, Agency, and Science*. University of Chicago Press.
- Rammert, W. (1997). New Rules of Sociological Method: Rethinking Technology Studies. *British Journal of Sociology*, 48(2), 171–191.
- Rammert, W. (1998). Giddens und die Gesellschaft der Heinezmännchen. Zur Soziologie technischer Agenten und Systeme Verteilter Künstlicher Intelligenz [Giddens and the Society of Good Folk. On the Sociology of Technical Agents and Systems of Distributed Artificial Intelligence]. In T. Malsch (Ed.), *Sozionik* (pp. 91–128). edition sigma.
- Rammert, W. (1999a). Relations that Constitute Technology and Media that Make a Difference. Toward a Social Pragmatic Theory of Technicization. *Philosophy & Technology*, 4(3), 271–290.
- Rammert, W. (1999b). Weder festes Faktum noch kontingentes Konstrukt: Natur als Produkt experimenteller Interaktivität [Neither Fixed Fact nor Contingent Construct: Nature as a Product of Experimental Interactivity]. *Soziale Welt*, 50(3), 281–296.
- Rammert, W. (2002). Verteilte Intelligenz im Verkehrssystem. Interaktivitäten zwischen Fahrer, Fahrzeug und Umwelt. [Distributed Intelligence in the Traffic System. Interactivity between Driver, Vehicle, and Environment]. *Zeitschrift für wirtschaftlichen Fabrikbetrieb*, 97, 404–408.
- Rammert, W. (2006). Technik in Aktion: Verteiltes Handeln in soziotechnischen Konstellationen [Technology in Action: Distributed Agency in Sociotechnical Constellations]. In W. Rammert, & C. Schubert (Eds.), *Technographie. Zur Mikrosoziologie der Technik* (pp. 163–195). Campus.
- Rammert, W. (2008a). Where the Action Is: Distributed Agency between Humans, Machines, and Programs. In U. Seifert, J. H. Kim, and A. Moore (Eds.), *Paradoxes*



- of Interactivity. Perspectives for Media Theory, Human–Computer-Interaction, and Artistic Investigations* (pp. 62–91). transcript.
- Rammert, W. (2008b). Technographie trifft Theorie. Forschungsperspektiven einer Soziologie der Technik [Technography Meets Theory. Research Perspectives for a Sociology of Technology]. In H. Kalthoff, S. Hirschauer, & G. Lindemann (Eds.), *Theoretische Empirie. Über die Relevanz qualitativer Forschung* (pp. 341–367). Suhrkamp.
- Rammert, W. (2012). Distributed Agency and Advanced Technology. Or: How to Analyze Constellations of Collective Inter-Agency. In J.-H. Passoth, B. Peuker, and M. Schillmeier (Eds.), *Agency without Actors? New Approaches to Collective Action* (pp. 89–112). Routledge.
- Rammert, W. (2024). *Distributed Agency. A Social Pragmatic View on Human-Technology Interaction*. Springer VS. (Original work published 2007)
- Rammert, W., Schlese, M., Wagner, G., Wehner, J., & Weingarten, R. (1998). *Wissensmaschinen. Die soziale Konstruktion eines technischen Mediums: Das Beispiel der Expertensysteme*. [Knowledge Machines. The Social Construction of a Technical Medium: The Example of Expert Systems]. Campus.
- Rammert, W. & Schulz-Schaeffer, I. (2003). Technik und Handeln. Wenn soziales Handeln sich auf menschliches Verhalten und technische Abläufe verteilt. [Technology and Agency. When Social Action is Distributed across Human Behavior and Technical Processes]. In W. Rammert & I. Schulz-Schaeffer (Eds.), *Können Maschinen handeln? Beiträge zum Verhältnis von Mensch und Technik* (pp. 11–64). Campus.
- Rammert, W. & Schubert, C. (2023). Technical and Human Embodiments of the Social. In Schubert, C. & Schulz-Schaeffer, I. (Eds.), *Berlin Keys to the Sociology of Technology* (pp. 95–126). Springer VS. https://doi.org/10.1007/978-3-658-41683-6_5
- Rorty, R. (1991). Inquiry as Recontextualization. An Anti-Dualistic Account of Interpretation. In *Objectivity, Relativism and Truth. Philosophical Papers* (pp. 137–163). Cambridge University Press.
- Rorty, R. (1999). *Philosophy and Social Hope*. Penguin.
- Schulz-Schaeffer, I. (2000). *Sozialtheorie der Technik* [Social Theory of Technology]. Campus.
- Schulz-Schaeffer, I. & Rammert, W. (2023). Technology, Action and Practice: The Concept of Gradual Action Revisited. In Schubert, C. & Schulz-Schaeffer, I. (Eds.), *Berlin Keys to the Sociology of Technology* (pp. 37–68). Springer VS. https://doi.org/10.1007/978-3-658-41683-6_3
- Schütz, A. & Luckmann, T. (1974). *Structures of the Life-World* [Vol. 1]. Northwestern University Press.
- Simmel, G. (1983). *Die Arbeitsteilung als Ursache für das Auseinandertreten der subjektiven und objektiven Kultur* [The Division of Labour as the Cause of the Divergence Between Subjective and Objective Culture]. Suhrkamp. (Original work published 1900)
- Star, S. L. (1989). The Structure of Ill-Structured Solutions: Boundary Objects and



- Heterogeneous Distributed Problem Solving. In M. Huhns & L. Gasser (Eds.), *Distributed Artificial Intelligence* (Vol. 2 pp. 37–54). Pitman.
- Suchman, L. A. (2007). *Human-Machine Reconfigurations. Plans and Situated Actions*. Cambridge University Press.
- Verbeek, P.-P. (2005). *What Things Do. Philosophical Reflections on Technology, Agency, and Design*. Pennsylvania State University Press.
- Weick, K. E. (1988). Enacted Sensemaking in Crisis Situations. *Journal of Management Studies*, 25, 305–317.
- Weick, K. E. (1995). *Sensemaking in Organizations*. Sage.
- Winner, L. (1980). Do Artifacts Have Politics? *Daedalus*, 109, 121–136.
- Wittgenstein, L. (1953). *Philosophical Investigations*. Macmillan.
- Wooldridge, M. & Jennings, N. (Eds.). (1995). Intelligent Agents: Theory and Practice. *The Knowledge Engineering Review*, 10(2), 115–152.

СВЕДЕНИЯ ОБ АВТОРЕ / THE AUTHOR

Вернер Раммерт,
werner.rammert@tu-berlin.de
ORCID 0000-0001-7554-2902

Werner Rammert,
werner.rammert@tu-berlin.de
ORCID 0000-0001-7554-2902

Статья поступила 8 ноября 2023
одобрена после рецензирования 2 марта 2024
принята к публикации 30 апреля 2024

Received: 8 November 2023
Revised: 2 March 2024
Accepted: 30 April 2024



Содержание

ChatGPT и голоса разума, ответственности и регулирования

[Елена Середкина и Юнмоу Лю](#)

ChatGPT и голоса разума, ответственности и регулирования 1-10

[Владимир Иванович Аршинов и Максим Францевич Янукович](#)

Нейросеть как отелесненный наблюдатель сложности:
Энактивный подход 11-25

[Владимир Иванович Шалак](#)

Избавление от иллюзий ИИ на примере ChatGPT 26-39

[Ребека Перес Леон](#)

Общаются ли языковые модели? Коммуникативное намерение и
референт с точки зрения Дерриды 40-56

[Анна Александровна Карташева](#)

Диалог с LLM как аутокоммуникация 57-66

[Александр Внутских и Сергей Комаров](#)

Lebenswelt, цифровая феноменология и модификации
человеческого интеллекта 67-79

[Андрей Юрьевич Алексеев и Екатерина Алексеевна Алексеева](#)

Персонологический функционализм GPT-ассистентов 80-99

Дискуссия: Андре Леруа-Гуран – Технология, язык, человек

[Андре Леруа-Гуран](#)

Происхождение и распространение научных знаний 101-115

[Ксавье Гуше](#)

Леруа-Гуран и объект технологии 116-124

[Симон Аврора](#)

“Une Véritable Syntaxe”. Некоторые примечания о Леруа-Гуране и
структурной лингвистике 125-135

[Натан Шлангер](#)

Между одобрением и отрицанием: Взаимодействие Андре Леруа-
Гурана с Россией” 136-152

[Вернер Раммерт](#)

Делать вещи с помощью слов и вещей: Прагматический взгляд на
аналогию языка и технологии 154-177