

xtitle

AI and language change & variation

st. schwarz

2025-12-13

index

abstract B: idee 2

language is developing and changing. there are many factors that influence language change, but i want to focus on one that may become relevant in todays language development: the change of language by constant and increasing use of AI tools that :communicate: with us as partners one may consider :real: or human-equivalent.

inspiration

now what are we experiencing if communicating with an artificial intelligence? first comes to mind the seemingly :natural language: adressed to us. one may feel as if talking to a human when asking questions and getting a response. studies prove that a significant amount of us show behaviour towards the AI that one would expect humans show only towards each other. that leads to the first question:

if we hold the AI as a human communication partner, could its behaviour towards us (here: their language) influence the way we talk/act viceversa? can people learn from an AI *how* to talk and what would they learn in this case? what is the language *taught* here specifically? do we adapt to patterns or linguistic markers common for *AI speech*?

AI speech: wtf

we can assume as common ground that the language used (here: output) by LLMs seems rather neutral, deprived of features deviating from the norm. its rather easy to understand, doesnt contain irony or sarcasm very often nor hyperbolic sentence structures (if not explicitly prompted) and could be very well used in a textbook for learners. it may be considered *universal* in aspects of transferability into other languages. it uses to not contain any specific vocabular or non-standard phrases. the syntax and grammar seems to follow the corresponding rules as the models are trained on large corpora of natural language. if we would (and we will do that) analyse a corpus of LLM outputs we very probably will find that in any feature it complies with the average feature matrix of any language compared. so if one language goes like SPO with having an average wordcount of 5wds/phrase and an average wordlength of 5 chars then the LLM certainly will show the same features for output in that language. no magic so far.

But: what if learners or people with deficient language skills begin to sync their output with the artificial language in their chatverlauf? simple like: beginning a response firstly with an appreciation of the :very interesting question: whatever the other may have asked? we're already heading that way...

There may also be tiny (oberlehrerhafte) standardisations of our own speech peculiarities (idiosyncrasies) we are confronted with which we are kind of nudged to relativate if always sending them into a black hole.

methods

first to do would be to create (or search for) a corpus of AI generated output. to use an existing corpus would prevent biasing which on the other hand could be interesting to explore i.e. we could by building (generating) a corpus ourselves on the basis of certain dedicated prompts¹ force the AI to generate phenomena of interest to our research question. where we get into medias res...

focusing questions

1. does a generative AI generally produces output that is in any aspects of interest for linguistic research?
2. how will users prime the output?
3. how are users adapting their own production to the output?
 1. is there any consistency concerning this adaptation?
 2. is there then societal adaptation of AI produced language?

¹which could very well be adapted to our research question

3. what are the rules (historic evidence) for adaptation?

going deeper

corpus creation

as proposed in **?@sec-methods** a secure way to building a corpus of AI speech - which we need to explore phenomena - is to archive LLM output to dedicated prompts. we will use an open source model provided with llama which can run on a local machine and is addressed via an API by script.

constraints

we will form prompts following these directives lead by our research question: 1. correction of mistakes to devise actual “knowledge” of model concerning standardisation and normative aspects 2.

capacities

to arrive at a research question, maybe discarding above

what i would like to...

- AI chat queries analysis
- analyse linguistic knowledge capacities of AIs
- proofread responses, factchecking

and what i actually can do

- statistics
- automated prompting using APIs
- automated response processing
- masked prompt generation

annotations

Q: BSI ([2025](#))

Table 1: annotations: revolution

id	annotations
1	Die Grenze zwischen Werkzeug und Gegenüber, zwischen Medium und Partner, wird dabei zunehmend durchlässig. Der Sprachgebrauch selbst wird zum Ort der Aushandlung: zwischen individueller Intention und algorithmischer Interpretation. Dieses neue Spannungsfeld erzeugt – so die zentrale These dieser Studie – neue Sprachgewohnheiten, veränderte semantische Ökonomien und möglicherweise sogar emergente Sprachformen, die sich jenseits etablierter kultureller und linguistischer Ordnungssysteme bewegen
2	Es entsteht eine hybride Sprachpraxis, die Merkmale von Alltagssprache, Programmiersprache, ökonomischer Instruktion, metaphorischer Verdichtung und affektiver Andeutung kombiniert – ein Phänomen, das wir im weiteren Verlauf dieser Studie als Promptlinguistik bezeichnen werden.
3	Promptlinguistik
4	Nutzer beginnen zunehmend, in einer Art Promptlogik zu denken und zu sprechen. Diese Logik folgt nicht der klassischen Satzstruktur oder narrativen Kohärenz, sondern ist geprägt von zielgerichteter Instruktion, semantischer Kompression, elliptischen Formulierungen und reduzierter Kontextualisierung
5	Der Kommunikationsstil verschiebt sich von einem syntaktisch vollständigen, sozialen Text hin zu einem effizienzorientierten, oft kryptisch anmutenden Sprachgebrauch
6	Diese Entwicklung lässt sich als Kognitionskompression begreifen – ein psycholinguistischer Prozess, bei dem sprachliche Komplexität reduziert wird, um mentale Entlastung, Output-Beschleunigung und Interaktionsökonomie zu erreichen
7	Prompten wird zur „sprechenden Skizze“ – nicht mehr zur Mitteilung, sondern zur Initialzündung algorithmischer Leistung.
8	Prompting ist in vielen Fällen nicht mehr klar als Englisch, Deutsch oder Spanisch identifizierbar, sondern als transkultureller, maschinenadaptierter Idiolekt
9	eröffnet zugleich neue Perspektiven auf eine mögliche Globalgrammatik der Mensch-MaschineInteraktion.
10	zielt die vorliegende Studie darauf ab, den Wandel sprachlicher und kommunikativer Muster im Kontext der Nutzung generativer KI systematisch zu erfassen und differenziert zu analysieren
11	Nutzergruppe von 112 Heavy AI Usern, die über einen Zeitraum von mindestens drei Monaten eine besonders intensive Interaktion mit generativen Sprachmodellen –vorrangig ChatGPT – aufwiesen
12	Entsteht durch die Mensch-Maschine-Interaktion ein emergentes, transkulturelles „Meta-Idiolekt“– eine neue Ebene symbolischer Ordnung zwischen Natur-, Kunst- und Programmiersprachen?
13	Ziel ist es, ein erstes kognitiv-semiotisches Mapping der neuen Sprachlandschaft im Zeitalter generativer KI zu entwerfen
14	Menschen streben danach, bei gleichbleibender Ergebnisqualität möglichst wenig kognitive Energie aufzuwenden (Kahneman, 2011).

Table 1: annotations: revolution

id	annotations
15	Erwartungshaltung, dass die Maschine semantische Lücken „mitdenkt“ , also kontextuelle Implikaturen und Nuancen selbstständig ergänzt
16	kompensatorisches semantisches Vertrauen
17	Diese Sprachveränderung steht nicht im Gegensatz zur Intelligenz des Users, sondern ist Anpassung an das System: Sie basiert auf gelerntem Promptingverhalten, das durch Erfahrung mit dem Modell bestärkt wird. Die Sprache wird maschinen-kompatibel – nicht mehr menschlich-vollständig, sondern systemadäquat
18	zeigen, dass bei intensiver Nutzung generativer KI die durchschnittliche Promptlänge signifikant abnimmt, während die semantische Dichte steigt
19	kognitive Prozesse nicht auf das Gehirn beschränkt sind, sondern sich über Werkzeuge und Umweltstrukturen erweitern können
20	Generative KI kann unter dieser Perspektive als externes semantisches Kognitionsorgan verstanden werden – ein „sprachlicher Cortex außerhalb des Körpers“ , der jedoch aktiv am Denkprozess teilnimmt.
21	Verarbeitungsflüssigkeit (Processing Fluency).
22	kognitive Bevorzugung von flüssiger Information wirkt sich nachweislich auf Urteile, Entscheidungen und Gedächtnisprozesse aus
23	kein Zeichen sprachlicher Degradation, sondern Ausdruck einer anpassungsfähigen Sprachintelligenz, die erkennt, dass komplexe Ästhetik nicht belohnt wird –während funktionale Direktheit zu besseren Ergebnissen führt.
24	sprachliche Verlagerung vom Ausdruck zum Effekt – das Ziel ist nicht die Formulierung als solche, sondern die Antizipation eines möglichst optimalen Outputs. Sprache wird zur strategischen Handlung, nicht mehr zur sozialen Äußerung.
25	Komplexe Idiome, Metaphern oder poetische Strukturen aus spezifischen Sprachkulturen funktionieren weniger zuverlässig – während reduzierte, direktive, entkontextualisierte Sprache überdurchschnittlich effektiv ist
26	Symbolische Interaktionismus als soziologische Theorieperspektive geht davon aus, dass soziale Wirklichkeit nicht objektiv vorliegt, sondern in sozialen Interaktionen hergestellt, verhandelt und symbolisch codiert wird
27	Bedeutung nicht in Dingen selbst liegt, sondern im Prozess der Bedeutungszuweisung durch Akteure.
28	Sprache
29	Sprache

References

- BSI, Brand Science Institute. 2025. “Wie KI Unsere Sprache Verändert – Eine Empirische Studie.” <https://www.bsi.ag/cases/104-case-studie-wie-ki-unsere-sprache-veraendert---eine-empirische-studie.html>.
- “Critical AI.” n.d. Accessed December 8, 2025. <https://www.dukeupress.edu/critical-ai>.
- Milička, Jiří, Anna Marklová, and Václav Cvrček. 2025. “AI Brown and AI Koditex: LLM-Generated Corpora Comparable to Traditional Corpora of English and Czech Texts.” <https://doi.org/10.48550/arxiv.2509.22996>.
- “Studie Zeigt: KI Lernt Sprachregeln Beim Lesen.” 2025. *FAU*. <https://www.fau.de/2025/11/news/forschung/studie-ki-lernt-sprachregeln-beim-lesen/>.