xtitle

AI and language change & variation

st. schwarz

2025-12-06

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 prompts[1] 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?
    3. what are the rules (historic evidence) for adaptation?

---

[1]which could very well be adapted to our research question

References