

Being a Generative Syntactician in the Age of Generative AI: A Coordinated Wh-Questions Case Study

Barbara Citko (University of Washington)

Recent years have witnessed the excitement about the usefulness of Large Language Models like ChatGPT in a wide variety of domains, as well as concerns about it (Bender *et al* 2021). In this talk, I address the claim that ChatGPT can be used as a native speaker informant/consultant as a source of grammaticality judgements. This is the question Haider (2023) asks (i.e., 'Is ChatGPT a grammatically competent informant?'), and answers in the affirmative. He examines the bot's performance on well-studied syntactic phenomena such as the *that-trace* effect or the Subject Condition and concludes that in general it 'will at least perform no worse than the average human informant'. Wilcox *et al* (2023) test various models', including ChatGPT's, performance on wh-island constraints, and Piantadosi (2023: 1) concludes more broadly that 'large language models have attained remarkable success at discovering grammar'.

I compare GPT-4's performance with human performance in light of these claims, where by 'human performance' I narrowly mean native speaker judgments reported in the relevant literature. I focus on wh-questions with coordinated wh-pronouns, illustrated in (1a), which belong to a family of constructions I have worked on for over a decade (Citko 2013, Citko & Gračanin Yuksek 2013, 2016, 2021, submitted), which in addition to coordinated wh-questions, includes coordinated free relatives (1b) and coordinated sluices (1c).

- (1) a. **What and where** does Mary sing?
- b. John sings **whatever and wherever** Mary sings.
- c. I know Mary sang something, but I forget **what and where**.