Large Language Models in Large Language Games

Chris Cousens, The University of Glasgow

When I receive an email making me a promise, giving me an order, or offering me a job, have I in fact been promised, ordered, or offered? Until recently, this would have likely been easily answered in the affirmative. We have become quite accustomed to the performance of online speech acts.

The recent proliferation of large language models might seem to undermine this. We can no longer be (as) confident that an email was created by a human 'speaker'. As LLM technology is increasingly embedded into the software and digital platforms used in the workplace (where speech acts have elevated moral significance) we are faced with a dilemma. Either we stop responding to online promises, orders, and offers *as* promises, orders, and offers, recognising illocutionary force only in the utterances whose author we *know* to be human. Or we respond to the 'promises', 'orders', and 'offers' of LLM-generated 'speech' as fully-fledged speech acts. Fisher (2024) argues that the integration of LLM and human speech threatens our ordinary speech practices. I agree—but what to do about it?

If we adopt the restrictive response, we preserve received views in speech act theory (invoking Austin 1962; Strawson 1964). Speech acts are often taken to require *intention* from the speaker, which LLMs seem to lack. The predominant alternative view takes speech acts to require *uptake* from the audience—but this is typically cached out as the recognition of speaker intention (Lance and Kukla 2013; McDonald 2022)! So, speech acts require (some sort of) speaker intention, we can no longer be confident that this is present for online speech acts, so we can no longer confidently give uptake to emailed speech acts. Speech act theory remains secure, but at the cost of our ability to promise, order, and offer online.

If we adopt the expansive response and allow that LLMs can perform speech acts, we reject a key component of traditional speech act theory regarding speaker intention (and the nature of uptake). Large language models would be no mere 'stochastic parrots' (Bender et al 2021). This may also entail that speech acts can be performed unintentionally by human speakers. We would then need an alternative conception of illocutionary force to explain this.

I will suggest a potential solution, grounding the force of an utterance-token in its contribution to the 'conversational score' (see Lewis 1979), rather than the intention of the speaker or uptake of the audience. This provides another reason to adopt the speech-act theoretic framework I have developed in earlier work (Cousens 2023, 2024). It preserves our ordinary online communicative practices, but at the cost of some received views in speech act theory—a cost I think we should be willing to pay. And it does so without making the concessions of other attempts to allow for AI to perform speech acts (e.g. Green and Michel 2022).

References

Austin, J. L. (1962). *How to do things with words*. Oxford University Press.

Bender, E., Gebru, T., McMillan-Major, A., & Mitchell, M. (2021). On the dangers of stochastic parrots: Can language models be too big? *FAccT '21: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 610-623, https://doi.org/10.1145/3442188.3445922.

Cousens, C. (2023). Tweet Acts and Quote-Tweetable Acts. *Synthese, 202*(187).

Cousens, C. (2024). Exercising Illocutionary Power, Or: How To Do Things With Other People's Words. In *Harmful Speech and Contestation*, ed. M. Popa-Wyatt, Palgrave Macmillan, 85-108.

Fisher, S. (2024). Something AI Should Tell You – The Case For Labelling Synthetic Content. *Journal of Applied Philosophy*, https://doi.org/10.1111/japp.12758.

Green, M., & Michel, J. (2022). What might machines mean? *Minds and Machines*, *32*, 323-338.

Lance, M., & Kukla, R. (2013). Leave the gun; take the cannoli! The pragmatic topography of second-person calls. *Ethics*, *123*(3), 456-478.

Lewis, D. (1979). Scorekeeping in a Language Game. *Journal of Philosophical Logic*, *8*, 339-359.

McDonald, L. (2022). Reimagining Illocutionary Force. *The Philosophical Quarterly*, *72*(4), 918-939.

Strawson, P. F. (1964). Intention and convention in speech acts. *Philosophical Review, 73*(4), 439-460.