Social Identities in Conversation: An Analysis of Human and AI-Generated Dialog

The research reported here looks at whether AI can simulate human identities in conversational interaction. Recent work has shown that LLMs can simulate social roles in conversations (Ye & Gao, 2024), yet these studies ignore how such simulations realize discourses and patterns of language use. To address this gap, the study adopts a variationist corpus-based approach centered on language use. The corpus is composed of two subcorpora: the human subcorpus consists of conversations from the the British National Corpus 2014, and the AI-generated subcorpus comprises conversations produced by GPT. Each AI-generated conversation was designed to mirror one single human conversation: prompts specified the topic of the corresponding BNC conversation, the socio-demographic traits to be simulated (e.g. race, gender, age), and a sequence of turn-level summaries of the human dialog, to ensure that each synthetic conversation corresponded to a version of the human conversations. Additionally, the AI was prompted with a generic, unprofiled prompt. The corpus, comprising 4900 texts, was analyzed using Lexical Multidimensional Analysis (Berber Sardinha & Fitzsimmons-Doolan, 2025), geared toward detecting discourses in a corpus. The corpus was processed by annotating all texts for POS and lemmatizing each word. Lemma frequencies were calculated and used for keyword analysis across subcorpora and conditions. The resulting keyword counts were entered into a factor analysis, which identified dimensions based on patterns of lexical co-occurrence. The dimensions were interpreted through analysis of texts with high scores on each factor, each corresponding to a particular discourse and style enacted in the conversations. To assess variation between the human and AI conversations and social profiles, ANOVAs were conducted on dimension scores. The study provides an empirical account of how LLMs linguistically construct social identities in interaction, thereby throwing light on the representational limits of AI-driven human modeling. The dimensions will be introduced and exemplified in the presentation.

References

Berber Sardinha, T., & Fitzsimmons-Doolan, S. (2025). *Lexical Multidimensional Analysis: Identifying Discourses and Ideologies*. Cambridge University Press.

Ye, F. T.-f., & Gao, X. (2024). *Simulating Family Conversations using LLMs: Demonstration of Parenting Styles*. https://arxiv.org/abs/2403.06144v1