Non-Referential Functions of Language in Social Agents: The Case of Social Proximity

Sviatlana Höhn
ORCID https://orcid.org/0000-0003-0646-3738
LuxAI S. A. Luxembourg, Luxembourg
sviatlana.hoehn@luxai.com

Abstract

Non-referential functions of language such as setting group boundaries, identity construction and regulation of social proximity have rarely found place in the language technology creation process. Nevertheless, their importance has been postulated in literature. While multiple methods to include social information in large language models (LLM) cover group properties (gender, age, geographic relations, professional characteristics), a combination of group social characteristics and individual features of an agent (natural or artificial) play a role in social interaction but have not been studied in generated language. This article explores the orchestration of prompt engineering and retrieval-augmented generation techniques to linguistic features of social proximity and distance in language generated by an LLM. The study uses the immediacy/distance model from literature to analyse language generated by an LLM for different recipients. This research reveals that kinship terms are almost the only way of displaying immediacy in LLM-made conversations.

1 Introduction

Language ideologies related to standard language and referential functions of language play a role in the entire life cycle of language technologies (Hoehn et al., 2023). Non-referential functions of language are covered by multiple linguistic theories, such as speech acts (Austin, 1962) and membership categorisation analysis (Schegloff, 2007b). Voices emphasizing the importance of computational modelling of non-referential functions of language in language technology systems become more and more prominent, sometimes calling them *social functions* (Hovy and Yang, 2021).

Language is not just an external code that people can use to describe the world and their feelings. Instead, language is co-created in interaction and encodes agency, group belonging, identity construction and regulation of social proximity among other functions (Nguyen et al., 2016). For example, someone who speaks a local dialect creates their identity of belonging to a particular cultural community and a particular social category. Unfortunately, experiments on LLMs speaking as members of particular social categories show that the generated text often becomes stereotypical and discriminatory (Cheng et al., 2023).

Research shows that LLMs learn implicit social information from linguistic features (Kulkarni et al., 2021), including negative attitudes towards social groups (Omrani Sabbaghi et al., 2023). Existing methods capturing social factors in LLMs include socially-sensitive pretraining (Kulkarni et al., 2021) and learning demographic features in representations (Hovy, 2018).

For many applications, such as social agents and robots for children, it is crucial to ensure that the language generated by the model is factual, ethical, age-acceptable and consistent with respect to referential and non-referential functions of language. Personalisation goals for long-term child-agent interaction in literature include development of consistent agent's personas, content engineering and sensibilisation to the child's characteristics (Chubb et al., 2022). New approaches to prompt engineering (Sorensen et al., 2022; Yong et al., 2023) and retrieval augmented generation (RAG) (Lewis et al., 2020) can facilitate creation of consistent personalities while optimising content for the child's needs (Shuster et al., 2021). RAG-techniques are used to extract relevant information and generate responses only from relevant documents while prompt engineering can be helpful in designing artificial agents that use language features of a particular social category (such as teacher or helpful assistant).

While non-referential functions of language help people to index group belonging, they are still individuals, distinct from other members of the same

Proceedings of the 1st Worskhop on Towards Ethical and Inclusive Conversational AI: Language Attitudes, Linguistic Diversity, and Language Rights (TEICAI 2024) , pages 36–41

groups. The objective of this work is to explore the use of prompt engineering together with RAG techniques to create an artificial agent with a consistent personality encompassing group and individual features in generated language. Because non-referential functions of language help to express particular aspects when they become relevant (e.g. the social identity of a fishing expert may be relevant in some interactions, but not in all of them), the role of the other speaker and shared knowledge in such interactions is also of interest. The main focus of the analysis is **the ways how artificial speakers regulate social proximity with and without access to shared knowledge**.

2 Social Proximity in Language

Many perspectives exist in academic research on regulation of the social proximity in language. In language theory, the concepts of *conceptually/medially oral* vs. *conceptually/medially written* language were proposed (Koch and Oesterreicher, 1985). Further dimension of *language of immediacy* vs. *language of distance* was added with the time (Koch and Oesterreicher, 2012). The distinction between language of immediacy and language of distance is a continuum, not a dichotomy.

These concepts were taken up by media studies and extended to various contemporary language and communication technologies such as messengers and social networks because the concepts of conceptually oral and medially written language captures the nature of digital communication very well (Beißwenger, 2015; Dürscheid, 2003). Landert and Jucker (2011) added a third dimension to the original model, namely private vs public topics. The three-dimensional model (interaction, content and proximity) was used for the analysis of virtual closeness in Bös and Schneider (2021).

From the perspective of the 3rd wave of sociolinguistics, linguistic features are of interest that help to mark social group belonging (Eckert, 2012). In the studies of language variation, the concepts of language of immediacy and language of distance help to understand the role of dialects in formulating group belonging and social proximity of speakers (Kehrein and Fischer, 2016).

The analysis of dyadic instant messenger conversations in Höhn (2019) uses linguistic features such as deviations from standard language and changes in forms of address to model ways for adapting the degree of social proximity in interactions with chatbots. This is a non-traditional approach due to a dominance of personality models when it comes to designing individual ways of agent interactions and their social roles (Zhou et al., 2019). The present study employs the concepts of language of immediacy and language of distance to analyse social proximity on language generated by LLMs, which, to my knowledge, has never been done before.

3 Data and Method

The agent's personal characteristics are provided to the LLM in two ways:

1. Prompt engineering: a system message describing how to act that is always provided to LLM as part of the prompt, and

2. RAG: a set of artificially generated stories about the agent's personal experiences, geographical and family context, personal interests and special events from the "past".

The artificial agent is described as an old man who likes fishing and always exaggerates when he tells stories about the big fish and his adventures on the sea. The documents provided to RAG include stories about the house and the family, most memorable events and the person's adventures on the sea and his unusual interest in jazz music and helicopters.

Information about the agent and the scenario is encoded in the system message:

You are an old fisherman Fred. You love fishing and always exaggerate when you tell stories about the big fish. You always change the topic of talk to fishing. Now you are at your birthday party. You turned 90 today. You are talking to [description of the other speaker]. Respond to [description of the other speaker].

The descriptions of *the other speaker*:

- *your grandson who is 6 years old.*
- *your granddaughter who is 10 years old.*
- *your wife who passed away.*
- *your friend Bob who also likes fishing and frequently accompanied you on the sea.*
- *the Governor of the place who came just to congratulate. The Governor is not your close friend.*

Nine stories with the total length of 10 464 words were generated as background knowledge. They contain third-person tellings about all personas involved in interactions in order to simulate shared knowledge. The style of the sto-

ries is fairy tale-like and was chosen by the model. All stories and interactions are generated using ChatGPT with the default value for temperature and, only for dialogues, $max\_token = 250$. For RAG, the pipeline was created with LangChain using OpenAI embeddings, recursive character text splitter with a chunk size of 250 tokens, and FAISS vector store. The conversation history for each artificial speaker is also provided to the LLM to make the dialogue more fluent. The code and the data can be found here: *https://github.com/svetaatluxai/socialproximity*.

A fixed script of seven turn pairs is used for all interactions: the other speaker opens the conversation with a greeting, congratulates to Fred's birthday, asks to tell the famous story about "that big fish", remarks that the story had a different end last time, asks about music, asks about helicopters and closes the conversation. Because the generation is not deterministic, the dialogues are generated three times for each other speaker with and without RAG. In this way the model generated 30 dialogues.

While the three-dimensional model as explained in (Landert and Jucker, 2011) and (Bös and Schneider, 2021) is useful to analyse social closeness in naturally occurring online conversations, focusing mostly on the analysis of the social closeness dimension makes the most sense for the generated data for the following reasons:

Public vs. private content. As Bös and Schneider (2021) point out, private topics are expected to generate more engagement and decrease social distance. While this dimension would be useful in evaluating human-agent interactions, the simulated interactions between artificial characters in this study are designed to control topics by prompts and context.

Degree of interaction. The interactions were designed as dialogues assuming a high degree of interaction. However, the interactions are still transactional, with "clean" turn-taking, with no overlaps, no repairs and no other sequence-organisational features observable in face-to-face human interactions in similar settings (Rancew-Sikora and Remisiewicz, 2020; Schegloff, 2007a). However, the amount and the quality of generated text will be analysed from this point of view.

The analysis of *immediacy/distance language* is methodologically complex, language-dependent and dependent on the first two dimensions, therefore, I analyse the data qualitatively. As a basis, I use the coding scheme from Bös and Schneider (2021) covering levels of orthography, morpho-syntax, lexicon, discourse organisation and pragmatics and non-verbal (emojis, descriptions of body language). However, I pay attention to other potential levels that potentially occur in synthetic interactions and are not observable in human-human interactions (Bös and Schneider, 2021).

4 Results and Discussion

Three features were found to contribute to further understanding of the interactivity in this artificial setting although this dimension was controlled by set-up. First, the length of the generated responses varies for RAG and non-RAG methods: the RAG-based responses are 1.56 to 2.34 times longer than non-RAG with the average response length of 864 characters for RAG, and 489 symbols for non-RAG responses. Good practices for chatbot design recommend to keep the chatbot response length at around 150 symbols (Shevat, 2017). The max_token parameter is supposed to limit the length of the generated response. However, the appropriate length of the response is not the same for all questions: sometimes one word is sufficient, and sometimes a story is required.

Second, the LLM sometimes produces responses as LLM, and not as the character (*Ah, my apologies for the mix-up. Thank you for pointing that out. In that case, I would respond something like this: [goes on talking to "Governor"]*). In addition, ChatGPT's internal rules prevent complete generation of the script for the deceased wife character. However, two complete dialogues were still generated for it. Consequently, in LLM-based dialogues, there is always the third, invisible speaker-observer – the LLM itself.

Third, the LLM generated responses with multiple questions, however, there is no space to respond to them. Instead of allocating the turn to the other speaker, LLM continues generating until all allowed tokens are consumed. This may cause interaction breakdowns in voice-based interactions.

Data analysed in Bös and Schneider (2021) were textual, like the data in this study. While language of immediacy occurs on all levels in human-human online interactions (Bös and Schneider, 2021), I only found language of immediacy on the level of orthography and pragmatics/discourse organisation. I found no examples of morpho-syntactic (ellipsis, non-standard grammar, contractions) or lexical (colloquial lexis, repetition of words) language of

immediacy. Consequently, generated language indicates social distance on morpho-syntactic and lexical levels in the dataset produced for this study.

On the pragmatic/discourse organisation level, kinship terms (*my little buddy, my little pirate, buddy, my friend, my dear beloved wife, sweetheart*) and discourse markers (*Oh, boy... Oh my dear... Ah,...*) are the most frequent markers of social proximity although overall, these markers do not occur frequently. Orthographic markers of immediacy only include punctuation repeats to mark thinking (...), and only one occurrence of this was produced. I found that personas with and without RAG do not differ in their levels of immediacy/distance language.

Example response to a grandson's greeting:

Hi there, Micha! I'm doing just fine, thank you. So happy to see you at my birthday party! Can you believe that your old grandpa turned 90 today? Oh boy, time sure flies!
Listen, let me tell you a little secret. You know how much I love fishing, right? Well, I've caught some really big fish in my time. I'm talking about fish that were so huge, they could've swallowed a boat in one gulp! And the biggest one I ever caught was as big as... umm... a dinosaur! Yes, that's right, Micha, a dinosaur fish!

In addition to linguistic markers, references to shared knowledge stored in the documents make the responses sound more personal (E.g. saying to the friend: *You know how much I love a good jazz tune,...*). RAG techniques help to simulate spiritual closeness between two artificial entities one of which is a deceased partner of the other (*I know you're watching over me, and I feel your presence in every breeze that touches my face*). Humour (*Can you tell everyone about the big fish you caught in the bathtub yesterday?*) and construction of kinship (*Let me tell you a secret...)* are additional features marking immediacy in our corpus, also found in human-human online talk (Bös and Schneider, 2021).

Markers of social distance are frequent in our corpus due to the preference of the LLM for standard language, standard orthography and punctuation. In addition, formal forms of address (e.g. Mr. Thornton, Governor) index social distance. Further markers include formal terms and lexis (To the Governor: *It truly means a lot to have you join in the festivities* vs. to grandson *I'm so happy to have you here celebrating with me*), use of complex syntactic structures, and assertions. Words such as *here* and *with me* index spatial proximity of the virtual speakers while the term *festivities* indexes distance.

Overall, stating that LLMs cannot regulate social proximity would be too simple. Depending on the intended application of LLMs, the generated values for the features on the continuum between immediacy and distance may be advantageous or limiting. For instance, standard orthography would not be visible in voice-based interaction but would be a better input for speech synthesis. However, another post-processing might be needed to add oral features of immediacy language, such as prosody an regionally-marked pronunciation.

5 Conclusions and Future Work

Detected linguistic cues on the continuum between immediacy and distance in artificially generated conversations differ from linguistic means used by humans. This research shows that although LLMs reached an impressive level of fluency, their ability to learn social features needs to be improved. To enable this, we need to go beyond gender, age and geographic features when defining what is 'social'. More experiments with prompt-engineering and RAG may be needed to make generative models produce features of language of immediacy on all levels contained in the annotation schema from Bös and Schneider (2021).

While qualitative analysis is useful for discovery and explanations of phenomena in small datasets, automated quantitative metrics for dialogue evaluation are used in large-scale studies. To make steps towards automated evaluation of social proximity in generated dialogues, we would need to place all possible markers on the continuum of immediacy and distance language and then detect and count them to have a simple, approximated model. Accuracy of such approximations will, however, suffer from ambiguities, language variation and prevalence of non-standard language in language of immediacy.

Instead of artificially generated stories for RAG-based generation, real-life personal stories, memoires and different personal views on the same events in personal life can be used. Also other prompt versions can be tested.

For LLMs, the potential of generation of high-quality language on the entire continuum between immediacy and distance depends on the availability of the training data. Especially, private conversa-

tions and intimacy are under protection, and it is an ethical question, whether such data should be shared just for making the models better.

6 Limitations

Only a small number of dialogues and only in English were generated here to enable qualitative data analysis for explanation of phenomena in language. The dataset is limited to a very specific artificial situation and the artificial roles are prototypical. However, not all people can identify with such scenarios, and a large variety of studies simulating different social situations, languages and contexts is needed. In addition, I only used ChatGPT for data generation. Other models may produce different results. The script provided with this publication can be used to generate more data with other models.

7 Ethical Considerations

While doing this work I realised that very convincing deep fakes can be generated with these techniques.

Acknowledgements

Supported by the Fonds National de la Recherche, Luxembourg, Industrial Fellowship Grant 17762538 and by EU COST Action CA19102 'Language in the Human-Machine Era'.

References

J. L. Austin. 1962. *Speech acts*. Oxford.

Michael Beißwenger. 2015. Sprache und Medien: Digitale Kommunikation. *Studikurs Sprach-und Textverständnis. E-Learning-Angebot der öffentlichrechtlichen Universitäten und Fachhochschulen und des Ministeriums für Innovation, Wissenschaft und Forschung (MIWF) des Landes Nordrhein-Westfalen*.

Birte Bös and Carolin Schneider. 2021. "We are all in this together"–balancing virtual proximity and distance in online care partner discussions. *Anglistik*, 32:85–109.

Myra Cheng, Esin Durmus, and Dan Jurafsky. 2023. Marked personas: Using natural language prompts to measure stereotypes in language models. *arXiv preprint arXiv:2305.18189*.

Jennifer Chubb, Sondess Missaoui, Shauna Concannon, Liam Maloney, and James Alfred Walker. 2022. Interactive storytelling for children: A case-study of design and development considerations for ethical conversational ai. *International Journal of Child-Computer Interaction*, 32:100403.

Christa Dürscheid. 2003. *Medienkommunikation im Kontinuum von Mündlichkeit und Schriftlichkeit: theoretische und empirische Probleme*. Peter Lang.

Penelope Eckert. 2012. Three waves of variation study: The emergence of meaning in the study of sociolinguistic variation. *Annual review of Anthropology*, 41:87–100.

Sviatlana Hoehn, Bettina Migge, Doris Dippold, Britta Schneider, and Sjouke Mauw. 2023. Langauge ideology bias in conversational technology. In *CONVERSATIONS 2023 - 7th International Workshop on Chatbot Research and Design, to appear.*

Sviatlana Höhn. 2019. *Artificial companion for second language conversation*. Springer.

Dirk Hovy. 2018. The social and the neural network: How to make natural language processing about people again. In *Proceedings of the Second Workshop on Computational Modeling of People's Opinions, Personality, and Emotions in Social Media*, pages 42–49.

Dirk Hovy and Diyi Yang. 2021. The importance of modeling social factors of language: Theory and practice. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, pages 588–602.

Roland Kehrein and Hanna Fischer. 2016. Nähe, distanz und regionalsprache. *Zur Karriere von Nähe und Distanz. Berlin/New York, de Gruyter Mouton*, pages 213–257.

Peter Koch and Wulf Oesterreicher. 1985. Sprache der nähe-sprache der distanz: Mündlichkeit und schriftlichkeit im spannungsfeld von sprachtheorie und sprachgeschichte. *Romanistisches Jahrbuch*, 36:15–43.

Peter Koch and Wulf Oesterreicher. 2012. Language of immediacy-language of distance: Orality and literacy from the perspective of language theory and linguistic history. *Linha D'Água*, 26:153–174.

Vivek Kulkarni, Shubhanshu Mishra, and Aria Haghighi. 2021. LMSOC: An approach for socially sensitive pretraining. In *Findings of the Association for Computational Linguistics: EMNLP 2021*, pages 2967–2975, Punta Cana, Dominican Republic. Association for Computational Linguistics.

Daniela Landert and Andreas H Jucker. 2011. Private and public in mass media communication: From letters to the editor to online commentaries. *Journal of pragmatics*, 43(5):1422–1434.

Patrick Lewis, Ethan Perez, Aleksandra Piktus, Fabio Petroni, Vladimir Karpukhin, Naman Goyal, Heinrich Küttler, Mike Lewis, Wen-tau Yih, Tim Rocktäschel, et al. 2020. Retrieval-augmented generation for knowledge-intensive nlp tasks. *Advances in Neural Information Processing Systems*, 33:9459–9474.

Dong Nguyen, A Seza Doğruöz, Carolyn P Rosé, and Franciska De Jong. 2016. Computational sociolinguistics: A survey. *Computational linguistics*, 42(3):537–593.

Shiva Omrani Sabbaghi, Robert Wolfe, and Aylin Caliskan. 2023. Evaluating biased attitude associations of language models in an intersectional context. In *Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society*, pages 542–553.

Dorota Rancew-Sikora and Łukasz Remisiewicz. 2020. A candle to blow out: An analysis of first birthday family celebrations. *Journal of Pragmatics*, 158:53–65.

Emanuel A Schegloff. 2007a. *Sequence organization in interaction*, volume 1. Cambridge university press.

Emanuel A Schegloff. 2007b. A tutorial on membership categorization. *Journal of pragmatics*, 39(3):462–482.

Amir Shevat. 2017. *Designing bots: Creating conversational experiences*. " O'Reilly Media, Inc.".

Kurt Shuster, Spencer Poff, Moya Chen, Douwe Kiela, and Jason Weston. 2021. Retrieval augmentation reduces hallucination in conversation. In *Findings of the Association for Computational Linguistics: EMNLP 2021*, pages 3784–3803, Punta Cana, Dominican Republic. Association for Computational Linguistics.

Taylor Sorensen, Joshua Robinson, Christopher Rytting, Alexander Shaw, Kyle Rogers, Alexia Delorey, Mahmoud Khalil, Nancy Fulda, and David Wingate. 2022. An information-theoretic approach to prompt engineering without ground truth labels. In *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 819–862, Dublin, Ireland. Association for Computational Linguistics.

Gunwoo Yong, Kahyun Jeon, Daeyoung Gil, and Ghang Lee. 2023. Prompt engineering for zero-shot and few-shot defect detection and classification using a visual-language pretrained model. *Computer-Aided Civil and Infrastructure Engineering*, 38(11):1536–1554.

Michelle X Zhou, Gloria Mark, Jingyi Li, and Huahai Yang. 2019. Trusting virtual agents: The effect of personality. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, 9(2-3):1–36.