

How does the language used **TO REFER** to one's in-group or out-group systematically vary relative to changes in state-of-the-world?

Through large-scale tagging of parallel comments from fans of both teams in NFL games, we find several **LINEAR VARIATIONS** in the form of the referent *grounded in live in-group win probabilities* — fans are less likely to refer to their (in-group) team the more likely they are to win.

Do *they* mean 'us'? Interpreting Referring Expression variation under Intergroup Bias

Venkata S Govindarajan, Matianyu Zang, Kyle Mahowald, David I. Beaver, Junyi Jessy Li

In prior work, we have shown that studying intergroup bias can enrich our understanding of how social relationships shape meaning in communication. Here we demonstrate how **the form used to refer** to one's in-group or out-group exhibit systematic variation relative to changes in state-of-the-world. We introduce a new dataset of comments from live game threads on subreddits dedicated to fandom for teams in the National Football League (NFL). Comments are aligned in time and *grounded in live win probabilities* (WP) of each team derived from the live score and other game features.



FIGURE 1: Our parallel corpus of intergroup language, grounded in live win probabilities.

Formulating our interest in reference form as a **tagging task** guides our annotation schema and modeling approach. We annotate 1499 comments — commenters refer to the in/out-group using names, nicknames, metonymy relations, and pronouns (*they/them, we/us*); the pronouns *we/us* covers every possible member of the in-group and is the most **inclusive**.

Few-shot training sample for Llama3

Comment: [SENT] Defense getting absolutely bullied by a dude that looks like he sells solar panels
Parent comment: None
In-group: Jets
Out-group: Bears
Live score: Jets 7 - 3 Bears
Target: [SENT] [IN] getting absolutely bullied by [OUT] that looks like [OUT] sells solar panels.
Explanation: The commenter is probably talking about the in-group, since 'Defense' is said without qualification, and the description of the offensive player is disparaging ('he sells solar panels'). 'Defense' should be tagged [IN] since it refers to in-group, and 'a dude' and 'he' should be tagged [OUT] since it refers to an out-group offensive player.

To understand if there are clear trends in referential expression form *over changes in WP*, we need a large sample of tagged comments. We fine-tune Llama-3-8b over our annotations with instructions, few-shot examples, and chain-of-thought explanations, obtaining an F1-score of **71%** on our expert test set. We use this fine-tuned model to tag and analyze **over 100,000 comments from our raw dataset**.

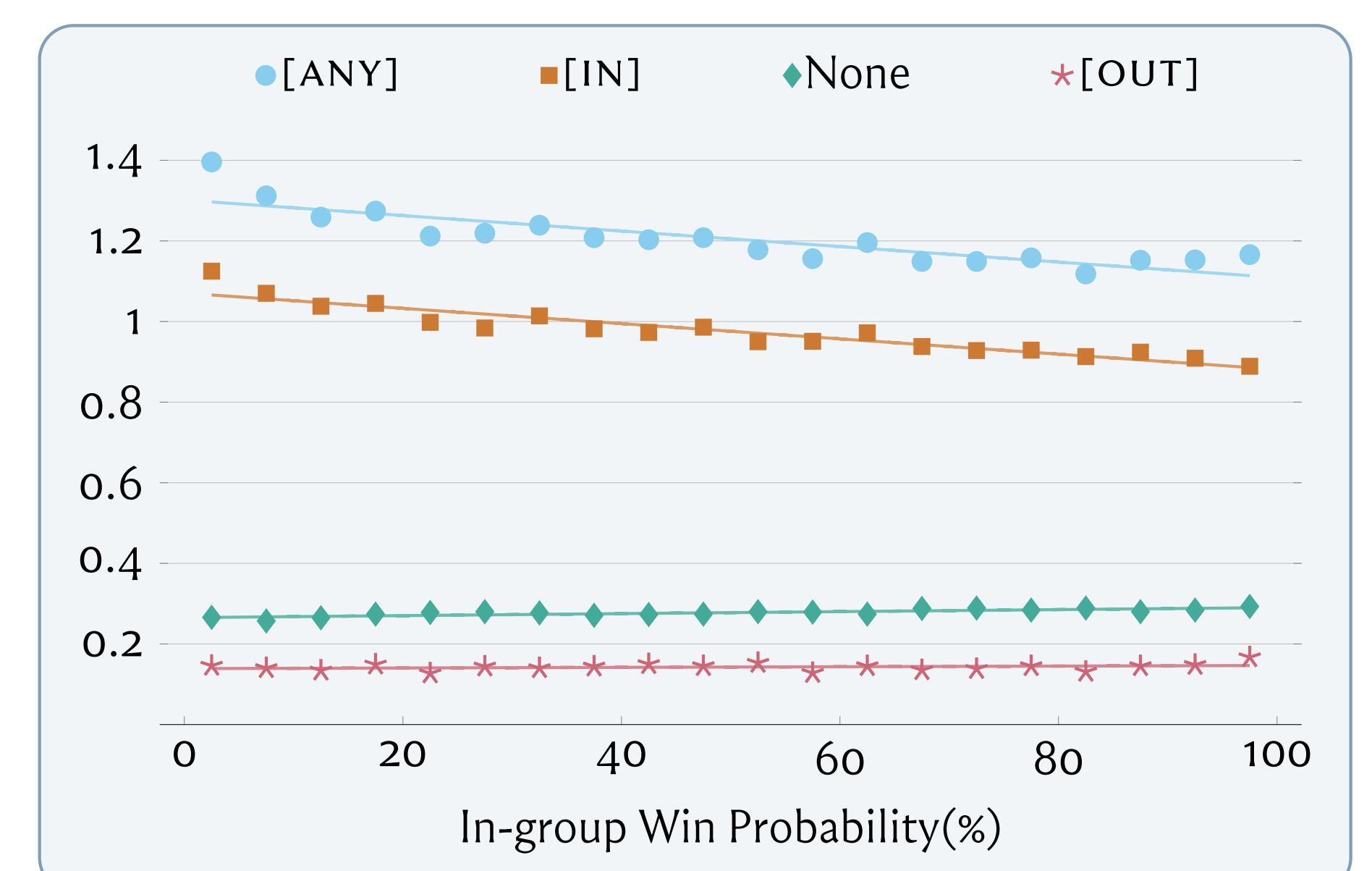


FIGURE 2: Absolute frequency of various reference expressions over WP windows.

- ① The more likely the in-group is to win, the less likely fans refer to the in-group — references to the out-group stay stable (Figure 2).
- ② Within referring comments (Figure 3), we see that **fans refer less to the in-group using *we/us* the more likely they are to win.**

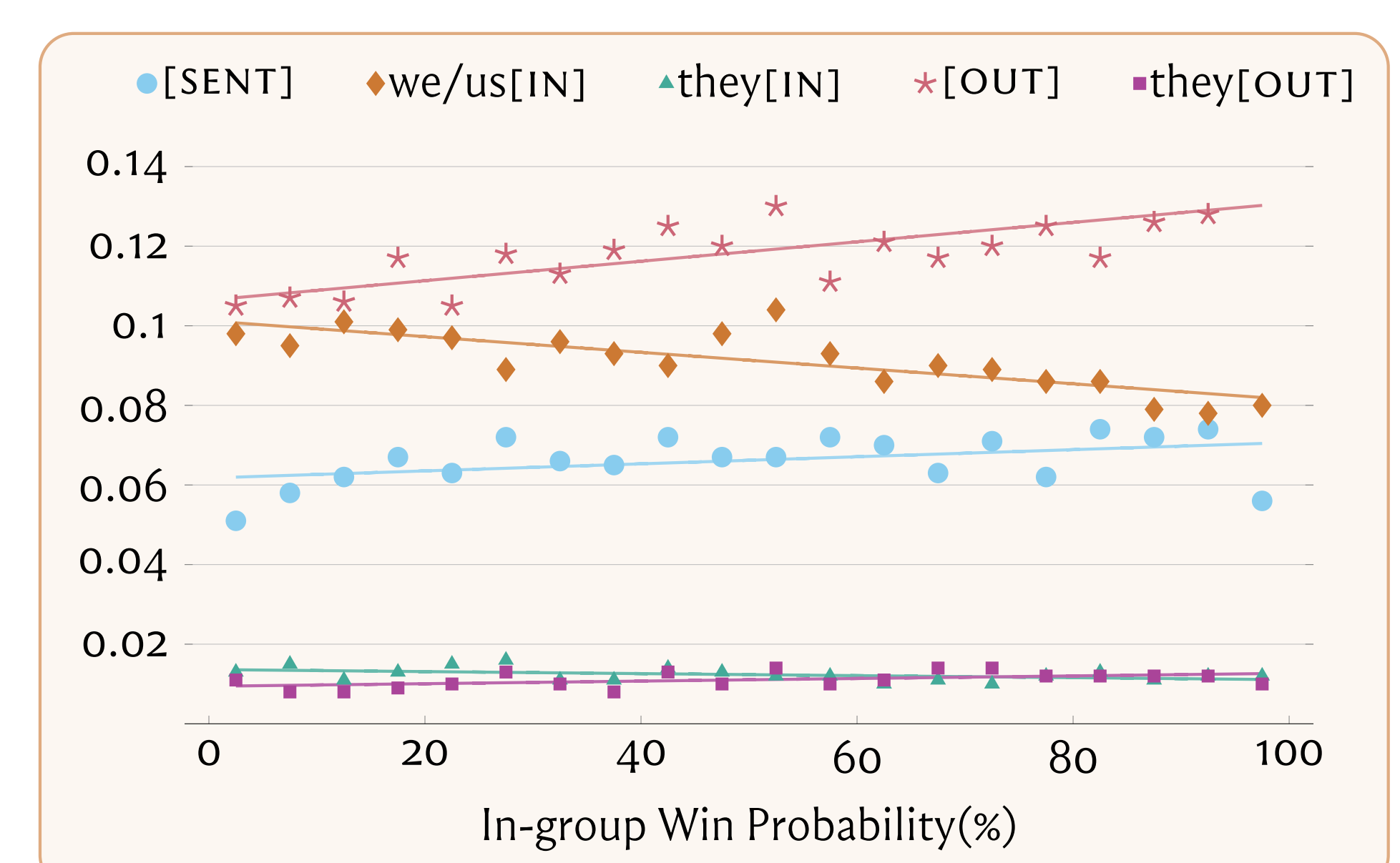


FIGURE 3: Normalized frequency of various reference expressions over WP windows.