Exploring Linguistic Style Matching in Online Communities: The Role of Social Context and Conversation Dynamics

Aparna Ananthasubramaniam,[*] **Hong Chen,**[*] **Jason Yan,**[*] **Kenan Alkiek,**[*] **Jiaxin Pei,**[*]
Agrima Seth,[*] **Lavinia Dunagan,**[*] **Minje Choi,**[*] **Benjamin Litterer**[*] and **David Jurgens**
School of Information, University of Michigan
{akananth,hongcc,jasonyan,kalkiek,pedropei,
agrima,laviniad,minje,blitt,jurgens}@umich.edu

Abstract

Linguistic style matching (LSM) in conversations can be reflective of several aspects of social influence such as power or persuasion. However, how LSM relates to the outcomes of online communication on platforms such as Reddit is an unknown question. In this study, we analyze a large corpus of two-party conversation threads in Reddit where we identify all occurrences of LSM using two types of style: the use of function words and formality. Using this framework, we examine how levels of LSM differ in conversations depending on several social factors within Reddit: post and subreddit features, conversation depth, user tenure, and the controversiality of a comment. Finally, we measure the change of LSM following loss of status after community banning. Our findings reveal the interplay of LSM in Reddit conversations with several community metrics, suggesting the importance of understanding conversation engagement when understanding community dynamics.

1 Introduction

Social influence can be subtle. When two persons converse, their interpersonal dynamics can lead to one person adopting the language of the other. For example, in settings where one person has higher status or power, the lower-status person may unconsciously begin mirroring the language of the other (Danescu-Niculescu-Mizil et al., 2012). This process has been described as *accommodation* (Giles and Ogay, 2007) or *linguistic style matching (LSM)* (Niederhoffer and Pennebaker, 2002) and can reflect the underlying influence that individuals have on each other (Chartrand and Bargh, 1999). Past work has primarily focused on how linguistic influence changes relative to the identities of the speakers. However, the larger social context in which a conversation happens also plays a role

---

[*]denotes equal contribution

in determining whether an individual may be influential. Here, we perform a large-scale study of linguistic influence to test how specific types of social context influence the level of accommodation.

Past work in the social sciences has studied accommodation to understand the influence and social power dynamics in specific settings, like job interviews (applicants and interviewers) (Willemyns et al., 1997) and academic context (students and faculty)(Jones et al., 1999). Also, LSM has been studied to understand group dynamics (Gonzales et al., 2010) and negotiations (Ireland and Henderson, 2014). Work in NLP has operationalized these theories to test accommodation theory in new domains. Typically, these works adopt some tests for measuring influence in language and have shown these measures correlate with known social differences. However, it is yet unknown how LSM occurs in conversations in online community platforms and differs by community dynamics.

Our work examines the larger context in which linguistic influence occurs. Using a large sample of 2.3 million conversations from Reddit and two measures of linguistic influence, we test how the level of linguistic influence correlates with conversational outcomes, such as conversation length and even the continued presence of a person in a community. Further, we examine how specific social and contextual factors influence the rates of linguistic influence. For instance, we discover that the controversy level of the parent comment can lead to different dynamics of style matching in the conversation threads.

This paper offers the following three contributions. First, we systematically compare complementary measures of accommodation, showing clear evidence of style accommodation in Reddit conversations. Second, we draw the relationships of several social factors that affect LSM, including levels of engagement, the popularity of the content, and tenure within a subreddit. Third, we demon-

strate the use of LSM to measure the loss of status through the banning of subreddits. We have released all code and data for full reproducibility.[1]

2 Accommodation and its Measurement

In this section, we discuss communication accommodation theory and associated sociolinguistic research to outline the accommodation of communicative behavior based on perceived social power dynamics. Subsequently, we explore the concept of linguistic style matching and methods adopted by researchers to quantify this phenomenon. We also investigate various factors that contribute to LSM variations and their strategic uses.

2.1 Accommodation Theory as Social Influence

When two individuals engage in social interaction, they may either converge or diverge in their communicative behavior. The Communication Accommodation Theory (CAT) suggests that the degree of convergence or divergence is affected by the relative social power between the interlocutors (Xu et al., 2018). Asymmetric convergence is more likely to occur in situations where there is a power imbalance between the interlocutors. Individuals with lower social power or status are more likely to adapt their communication style to align with those in higher or dominant positions (Muir et al., 2016). For instance, Puerto Ricans in New York City during the 1970s, who were perceived to hold less power than African Americans, adopted the dialect of African Americans to converge with their more powerful counterparts (Wolfram, 1974).

Social power has been often found to be an important determinant of degrees of accommodation (Giles et al., 1991; Ng and Bradac, 1993) and interactants of differential social power or social status can act in a complementary fashion (Street, 1991).

2.2 Linguistic Style Matching

Linguistic alignment is a pervasive phenomenon that occurs in human communication where interactants unconsciously coordinate their language usage. This coordination, described as convergence in the psycholinguistic theory of communication accommodation, involves aspects such as word choice, syntax, utterance length, pitch, and gestures (Giles et al., 1991). Linguistic style match-

ing (LSM) is a specific manifestation of linguistic alignment, wherein individuals unconsciously match their speaking or writing styles during conversations (Ireland et al., 2011). Unlike content accommodation, LSM focuses on stylistic accommodation, examining how things are communicated rather than what they communicate.

Individuals strategically negotiate their language style to decrease social distance, seek approval, and accommodate each other. LSM can also reflect the level of common understanding and conceptualization of the conversation topic between speakers. The degree of LSM can indicate social power dynamics as indicated by (Giles and Ogay, 2007). Empirical evidence from recent studies (Danescu-Niculescu-Mizil et al., 2012) showed that participants with less power (such as lawyers or non-administrative roles in Wikipedia) exhibit greater coordination in conversational behavior than participants with high power (such as justices or administrators). Additionally, Noble and Fernández (2015) identified a positive correlation between linguistic accommodation and social network centrality, which effect can be greater than the effect of power status distinction. Studies by Muir et al. (2016, 2017) further show that individuals in a lower position of power tend to accommodate their linguistic style to match that of their higher-power counterparts during face-to-face communication as well as computer-mediated communication.

The variance in LSM can be attributed to various social and psychological factors and can be triggered for different purposes. Linguistic alignment may signal likability and agreement, relate to seeking approval or arise from social desirability. Higher levels of accommodation in social behaviors are found to be associated with increased feelings of affiliation, liking, and successful interpersonal relationships (Bayram and Ta, 2019). Thus, linguistic alignment can be strategically employed to establish relationship initiation and stability (Ireland et al., 2011), increase group cohesion, and task performance (Gonzales et al., 2010), and assist in negotiations (Taylor and Thomas, 2008). Furthermore, alignment has been found to enhance persuasiveness, motivating listeners to adopt healthier practices (Cialdini, 2001) while in some cases like presidential debates, it has been perceived as more aggressive (Romero et al., 2015). The degree of matching may differ based on context and individual factors.

---

[1] https://github.com/davidjurgens/style-influence

(a) Accommodation, # Function Words
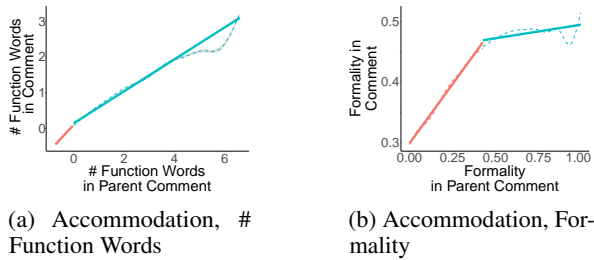
(b) Accommodation, Formality

Figure 1: Commenters on Reddit accommodate to the (a) # function words and b) formality of the comment they are replying to. Typically, the level of accommodation is higher when responding to posts with below-average (red) than above-average (blue) style.

3 Data

Reddit is a popular social media platform with a forum-based interface. It allows users to interact with dispersed individuals who share similar experiences or topics of interest. Our dataset to study LSM spans from July 2019 to December 2022 and includes 35M users and 500K subreddits.

Using the Pushshift Reddit Dataset which contains the full history of comments aggregated on a monthly basis (Baumgartner et al., 2020), we construct conversation threads from the comments and filter those that satisfy the following conditions: (1) the conversation chain consists of exactly two users; (2) the beginning of the conversation chain must be a root comment which does not have a parent comment; and (3) the lengths of a conversation chain must between 3 and 100. These conditions allow us to capture conversation dynamics between exactly two users without any interference. Our resulting dataset contains 16,893,013 conversation turns (or comments) across 2,305,775 conversation chains from 68,788 subreddits.

4 How should we measure linguistic influence?

Computational work has proposed multiple approaches for both what to measure and how to measure linguistic influence. In this section, we aim to build intuition for what the two measures of accommodation—using function words and formality—are operationalizing.

4.1 Linguistic Style Markers

Our study measures linguistic influence with two complementary style markers. We use the notation m to refer to a marker throughout.

Marker 1: Function Words   Function words (e.g. pronouns, prepositions, articles, and auxiliary words) are primarily employed unconsciously and frequently and incorporate social knowledge for comprehension and usage (Meyer and Bock, 1999; Ireland and Pennebaker, 2010). Prior computational studies of linguistic accommodation have measured linguistic influence by tracking the relative frequencies of function words across conversation turns (Danescu-Niculescu-Mizil et al., 2011; Babcock et al., 2014; Gonzales et al., 2010). Function words reflect *how* content is expressed, rather than what specific content is expressed (e.g., content words) and are thought to be a better proxy for unconscious language processing (Tausczik and Pennebaker, 2010). Here, we use the function words defined by the Linguistic Inquiry and Word Count (LIWC) lexicon (Pennebaker et al., 2001; Pennebaker and Chung, 2007).

Marker 2: Formality   Individuals adopt a specific register that is appropriate to their position in the social context, real or desired (Niederhoffer and Pennebaker, 2002). A commonly varied register is the level of formality used when speaking to another. The level of formality shown by a speaker is known to reflect the speaker's opinion towards a topic or their closeness to the listener (Hovy, 1987). Unlike function words, variation in formality often requires conscious processing to select the appropriate phrasing in a given circumstance. As a result, it offers a complementary view into how a speaker influences another through shifting the conversation towards a more formal or informal register.

Here, we measure formality using a supervised classification model. The model is a fine-tuned RoBERTa-based classifier (Liu et al., 2019) trained on the GYAFC (Rao and Tetreault, 2018) and Online Formality Corpus (Pavlick and Tetreault, 2016) datasets; we use the model available from the Hugging Face API[2]. Both datasets contain social media text and the reported model performance is high for both blogs and Q&A text (Spearman's $\rho>0.7$). Using this classifier, each comment's formality is measured on a continuous scale in [0,1].

Importantly, these style variables are related; function word frequency also changes in more formal contexts, where articles and prepositions typically become more common while pronouns and interjections become less common (Heylighen and

---

[2]https://huggingface.co/s-nlp/roberta-base-formality-ranker

Dewaele, 1999). Content word-based measures of style and function word counts are thought to capture the same latent style variables, i.e., they are interchangeable at a stylometric level (Grieve, 2023).

4.2 Measuring Linguistic Influence

At a high-level, linguistic influence (also referred to as LSM or accommodation in this paper) is measured by testing whether the value for some measure m of a comment made by user a is predictive of the value of m in the reply to that comment by user b. Therefore, one straightforward way to measure accommodation is with linear regression: $m_b \sim \beta_0 + \beta_1 m_a$ where β_0 reflects the baseline level of the measure (e.g., the average formality) and β_1 measures the level of accommodation (e.g., the average increase in formality associated with a 1-unit increase in the formality of the parent comment). However, as Xu et al. (2018) note, the characteristics of a comment are likely influenced by other unrelated factors such as the length of the comment or the number of turns in the conversation. Indeed, they show that unless one controls for such factors, linguistic influence may be overestimated. Therefore, we used a mixed-effects regression to control for comment a and b's length in tokens (fixed effects L_a, L_b), the number of replies $r_{b \to a}$ that b has made to a so far in the conversation. To capture individual and community-level variation, we include random effects to control for the effect of the subreddit s; these random effects let us control for differences in the norms of communities (e.g., some communities are more/less formal) to test for relative changes in m. Linguistic accommodation is modeled as

$$m_b \sim \beta_0 + \beta_1 m_a +$$
$$\beta_2 L_a + \beta_3 L_b + \beta_4 r_{b \to a} + (1|s)$$

where β_1 measures the level of accommodation.

4.3 Results

We first observe clear evidence of accommodation in both style markers: parent comments with more function words receive replies with more function words (Figure 1a), and more formal parent comments receive more formal replies (Figure 1b). For comments where we have the text of the original post, we observe accommodation even after controlling for the author and original post's style markers, suggesting that users may accommodate

(a) Time to Comment, # Function Words
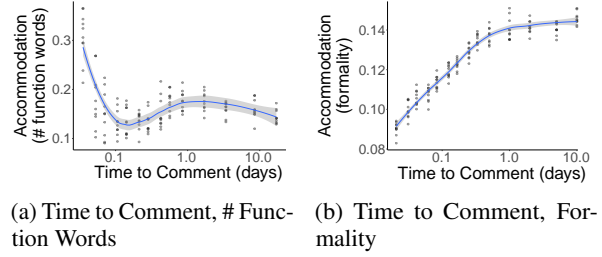
(b) Time to Comment, Formality

Figure 2: Commenters on Reddit are more likely to accommodate (a) # function words when they reply quickly (suggesting subconscious accommodation) and b) formality of the comment when they reply slowly (suggesting strategic accommodation).

to the style of the person they are interacting with in the comment thread. However, this effect plateaus when the parent comment has above-average levels of a style marker, suggesting a potential threshold for the impact of parent comment style on reply style. This attenuation of effect may be the result of several mechanisms, including regression to the mean or an author modulating their replies according to their own personal style (i.e., a more extreme parent comment may trigger greater modulation).

Second, the two style markers are almost perfectly uncorrelated, suggesting that they measure distinct constructs. In order to calculate the correlation between these two measures, we randomly sample 1,000 subsets of the conversation turns and calculate the extent of accommodation in function words and formality in that subset. The correlation between the function-word- and formality-based accommodation scores is -0.00171.

Third, accommodation in the two style markers seems to occur via fundamentally distinct psychological processes. Accommodation can occur either 1) through a subconscious priming mechanism, where the speaker instinctively repeats what they hear; or 2) through a more conscious, strategic act with communicative intent (Doyle and Frank, 2016). Figure 2 suggests that function-word-accommodation seems to be an unconscious form of relating to the audience, while formality-accommodation seems to be more intentional and strategic. Commenters exhibit greater accommodation in function words when they take less time to reply to the prior comment (2a) and greater accommodation in formality when they reply more slowly (2b). These results are consistent with prior work, suggesting that accommodation of function words occurs subconsciously (reflexively, takes less time) and builds on this work to show that accommoda-
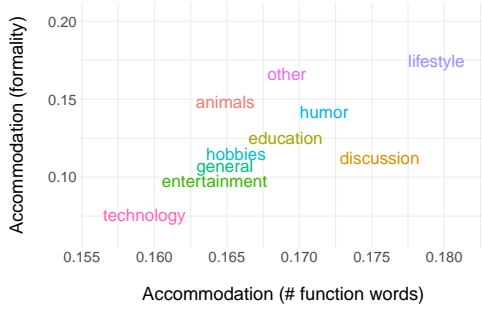
Figure 3: Level of accommodation in the number of function words (x-axis) and in formality (y-axis).
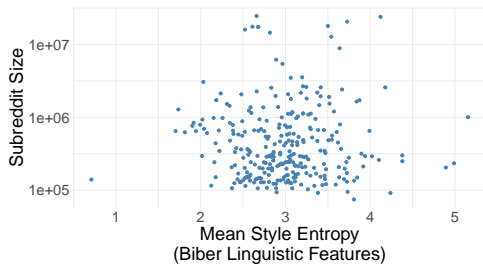

Figure 4: The mean Shannon Entropy of Biber's linguistic features (x-axis) is uncorrelated to the subreddit's number of subscribers (y-axis) ($p = 0.41$). Entropy is calculated using a random sample of comments in each subreddit.

(a) Comment Depth, Function Words

(b) Comment Karma, Function Words

(c) Comment Depth, Formality

(d) Comment Karma, Formality

Figure 5: Characteristics of the comment are non-linearly associated with accommodation, including comment depth (a,c) and comment karma (b,d).

tion in other style markers, like formality, occurs strategically (intentionally, takes more time).

Fourth, there is little variation in accommodation across subreddit characteristics. Figure 3 shows the levels of accommodation across ten different types of subreddits, using an existing taxonomy of popular subreddits.[3] While certain types of subreddits (e.g., lifestyle) tend to have higher levels of accommodation than others (e.g., technology, entertainment), most differences are only weakly significant ($p > 0.01$) with a small effect size. Moreover, Figure 4 shows the relationship between subreddit size and variation in linguistic style, for 300 subreddits sampled based on their number of subscribers. To calculate variation in linguistic style, we use Biber (1988)'s comprehensive set of linguistic features. Linguistic variation within each subreddit is estimated as the mean Shannon Entropy of each Biber tag frequency at the subreddit level. Despite expectations that larger communities may exhibit greater diversity in language use (Kocab et al., 2019), we find no relationship between community size and linguistic variation.
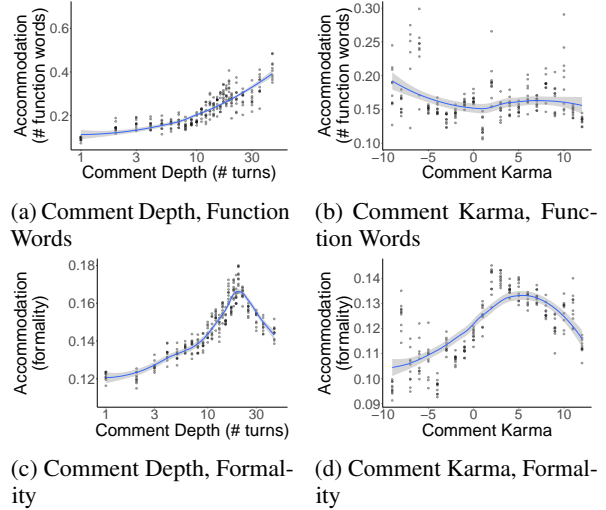
Overall, these findings point to the nuanced dynamics of LSM in online interactions, indicating that factors such as function word usage and formality in the parent comment are associated with the linguistic style and tone of replies.

5 What factors about a comment influence the degree of accommodation?

LSM can be affected by many factors and existing studies have pointed out the roles of not only linguistic characteristics but also the contextual factors affecting LSM (Niederhoffer and Pennebaker, 2002). In this section, we study the connection between LSM and a series of contextual factors where the comment is posted (i.e., comment depth) and the "success" of a comment (i.e., comment Karma and parent comment Karma).

5.1 Experimental Setup

To test for heterogeneity in the level of accommodation with respect to several covariates (e.g., depth, Karma), we run a mixed effects regression similar to Section 4.2, but include an interaction term to test whether accommodation changes significantly with respect to some covariate (say, Karma K):

$$m_b \sim \beta_0 + \beta_1 m_a + \beta_2 K + \beta_3 m_a * K + \\ \beta_4 L_a + \beta_5 L_b + \beta_6 r_{b \to a} + (1|b) + (1|s)$$

Here, β_1 measures the level of accommodation when $K = 0$ and β_3 measures the increase in accommodation when K increases by one point; if β_3

---

[3] https://www.reddit.com/r/ListOfSubreddits/wiki/listofsubreddits/

is significantly different from 0, then we have evidence that accommodation is heterogeneous with respect to Karma.

In order to visualize these effects, we fit the model in the above equation to estimate accommodation at different values of Karma. In order to appropriately represent uncertainty in this model, we sample 100,000 conversation turns at each value of Karma 10 times and use this to obtain 10 different estimates of accommodation for each value of the covariate. To visualize the association between Karma and accommodation, we plot Karma on the x-axis and the LSM estimates on the y-axis.

5.2 Results

As shown in Figure 5, various factors of comments are related to LSM.

Comment depth Comment depth reflects the position of a comment in the conversation tree. Deeper comments are usually posted in longer conversations and when the users are more engaged in the dialogue. As shown in Figure 5a and Figure 5c, comment depth is positively correlated with LSM. However, accommodation in formality drops off for very deep comments. LSM happen more when the comment is deeper in the conversation tree, suggesting that users tend to match not only the content but also the structural aspects of their language in response to their interlocutor. Such a trend could be due to greater investment in the conversation. When two users are involved in longer and deeper conversations, they are more likely to be engaged in the conversation, which may lead to higher subconscious but lower conscious LSM.

Comment Karma A key feature of Reddit is the ability for users to upvote or downvote comments, which determines the comment's karma - a measure of its popularity within the community. In figure 5, we observe several non-linear associations between karma, comment characteristics, and LSM. In terms of comment karma, users' LSM tends to remain relatively constant, except for cases where the comment has very high karma, which is associated with an increase in LSM. This finding implies that highly popular comments may foster greater linguistic alignment between users.

We also see that comments with low karma have lower levels of LSM than comments with high karma (Figure 5d), which makes sense since we'd expect users to respond better to comments whether the author is mirroring their interlocutor. Notably,

this upward trend reverses in comments with very high karma – which have lower levels of LSM than comments with lower levels of karma. The reversal of the LSM trend in comments with high karma warrants further exploration. One possible explanation for this phenomenon is that highly upvoted comments may exhibit unconventional linguistic styles that deviate from the norm, which could be seen as novel by the Reddit community. Another explanation may be that comments with high karma are more likely to be popular in larger, diverse communities where users may have a wider range of linguistic styles. Additionally, it is possible that comments with high karma receive a higher volume of comments and interactions, which may dilute the overall LSM score due to the presence of diverse linguistic styles from multiple interlocutors.

6 What effect does accommodation have on the conversation itself?

Linguistic accommodation is usually associated with positive social benefits (Taylor and Thomas, 2008; Gonzales et al., 2010). Here, we test whether linguistic accommodation is associated with two positive behaviors in social media: sustained conversation and length of participation in a subreddit.

6.1 Experimental Setup

We fit a linear regression on conversational dyads following the LSM measure in Section 4.2. Following the procedure from the prior section, we estimate the level of accommodation for comments around a particular covariate by sampling 100,000 conversation turns at or near the respective value of the covariate. Once again, we verify that differences between covariates are significant, by introducing interaction terms in the regression and testing for a statistically significance effect.

6.2 Results

Figures 6a and 6b compare the effect of alignment when conditioned on the total length of the conversation thread. For both functions words and formality, we observe from the fitted lines that accommodation is more likely to happen from longer conversations, but only up to a certain length of approximately 30-40. This suggests the possibility of LSM being an earlier indicator of how engaged the users will be in a conversation. On the other hand, the likelihood of accommodation in formality decreases when the conversation becomes longer

(a) Length, # Function Words

(b) Length, Formality

(c) Depth x Length, # Function Words
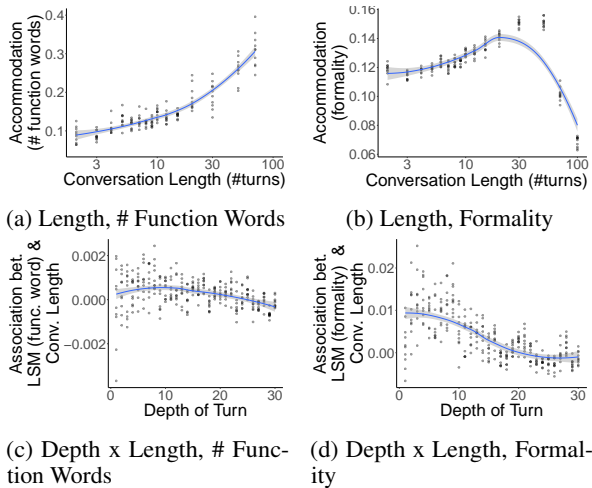
(d) Depth x Length, Formality

Figure 6: The number of turns in a conversation is associated with the level of accommodation in each turn: (a-b) Longer conversations (i.e., threads consisting of more conversation turns) are associated with higher accommodation, up to a threshold; for sufficiently long conversations, the association is negative for formality. (c-d) The effect of alignment on conversation length is stronger earlier in the conversation and weaker as more conversational turns occur.

than a certain threshold, which suggests that speakers may stop consciously trying to accommodate once the conversation becomes sufficiently long.

Figures 6c and 6d compare accommodation likelihoods at a given turn within a conversation. Interestingly, we can observe that LSM starts off highest at the beginning of a conversation and decreases as the number of turns increases. Combining the two results, we can conjecture that while the degree of LSM generally decreases within a conversation thread, the initial levels of LSM observed at the early stages of a conversation can indicate how engaged the speakers will be, which one can use to estimate the overall conversation length.

How does LSM differ by tenure and number of subsequent posts in a subreddit? Figure 7 shows that, for both style markers, users who have a longer tenure in the subreddit or who post more in the subreddit in the next month tend to display higher subconscious and lower conscious LSM. We consider these results as evidence of the "lifespan" of a user's engagement toward conversations held within that subreddit, and ultimately engagement toward the subreddit itself, which has been noted in prior work (Danescu-Niculescu-Mizil et al., 2013).
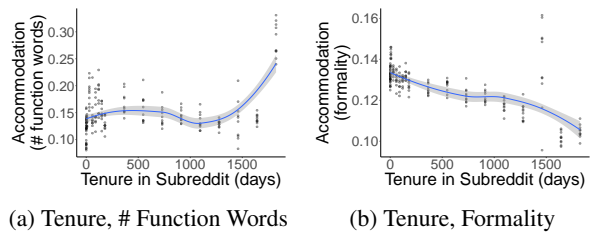

(a) Tenure, # Function Words

(b) Tenure, Formality

Figure 7: Alignment is associated with engagement: (a) stronger function-word LSM and (b) weaker formality LSM for higher tenure in subreddit.

7 What effect does the social context have on accommodation: Controversiality?

In this section, we examine whether LSM differs by social contexts that arise during conversations. Specifically, we focus on the controversy level of the parent comment. In contrast to non-controversial issues, controversial issues lead to competitive disagreement, where the goal of the groups involved in argumentation is to convince the opponent group(s) of the validity of one's point of view (Ilie, 2021). The arguments on controversial issues tend to invite strong emotions with negative affect (Mejova et al., 2014) and deteriorate the deliberation in the public sphere because interactions often turn uncivil (Doxtader, 1991).

7.1 Experimental Setup

Following the procedure from the prior section, we estimate the level of accommodation for comments at each covariate, separately for controversial and non-controversial comments. When a comment or post receives a substantial number of upvotes and downvotes, Reddit automatically designates it as controversial. The exact method used by Reddit to determine controversy remains private. However, the Reddit API offers a binary label indicating whether a comment is controversial or non-controversial (Koncar et al., 2021). Approximately 1.30% (n=218,899) of the comments in our sample are labeled as controversial.

We test that differences between conditions are significant with a three-way interaction term in the regression between the parent-comment style, the comment's Karma (or other covariates) and the comment's controversiality: $m_a \times K \times C$.

7.2 Results

Figure 8 reveals that LSM occurs differently in controversial and non-controversial comments. For both function words and formality, LSM is less
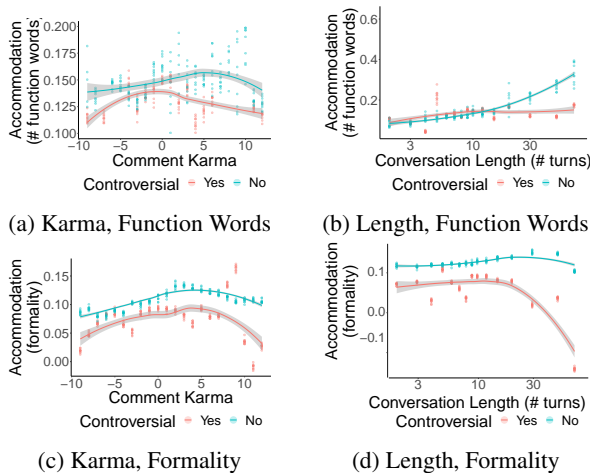
(a) Karma, Function Words   (b) Length, Function Words

(c) Karma, Formality   (d) Length, Formality

Figure 8: The associations between LSM and "success" are very different for controversial (red line) and non-controversial (blue line) comments: (a) Very-high-karma controversial comments tend to have lower, rather than higher, function-word alignment; and (c) overall lower formality-alignment. (b) Compared to shorter threads, longer controversial threads tend to have lower, rather than higher, function-word alignment and (d) formality alignment.

likely to occur in controversial rather than non-controversial comments when the conversation length is below a certain threshold (12-14). Interestingly, we see that this trend is strengthened as the conversation length increases. One possible explanation is that controversial comments generate more initial interest that promotes users to engage more in conversations. However, this initial effect is washed away as the conversation takes further turns, and the conversation is less likely to continue due to reasons such as incivility. Non-controversial comments, on the other hand, enjoy less of this initial boost and is more likely to carry on if the users have accommodated each other's language during their conversation.

With the addition of Karma, we can observe a more complex trend that plays out differently for each style marker. For function words, conversations in controversial comments have a nonlinear relationship that drops as the parent comment's Karma increases, whereas a weak positive correlation can be observed for non-controversial comments and levels of Karma. In contrast, for formality, LSM occurs most at comments with about 0-5 Karma and decreases for higher Karma for both controversial and non-controversial comments.

Overall, we observe that social contexts that are defined by the community platform such as Karma

or controversy have complex, nonlinear effects on how LSM occurs in conversations.

8 Loss of Status via Community Banning

Reddit bans specific subreddit communities as a result of policy violations, such as repeated posting of highly offensive content or lack of moderator oversight (Chandrasekharan et al., 2017). When users are highly active in such communities, the ban potentially results in a loss of status, as they are forced to find new communities to participate in. Here, we test the extent to which users change how they are linguistically influenced by others after such a ban. While prior work has studied how users change after *gaining* status (e.g., Danescu-Niculescu-Mizil et al., 2012), our unique setting allows us to perform a novel study of the potentially humbling effects of status loss. In addition, a study of the subreddit r/changemyview suggests that formality is (weakly) associated with more effective persuasion on Reddit (Dayter and Messerli, 2022); we hypothesize that users who recently experienced a ban may have multiple pragmatic reasons to accommodate more.

8.1 Experimental Setup

We test for changes to linguistic influence using a pseudo-causal difference-in-difference analysis (Lechner et al., 2011). Subreddit ban dates were determined by identifying all banned subreddits and then using the last date of a post in that subreddit. Our sample includes 1,024 subreddits banned between July 2019 and December 2022. We identify 16,686 users in our sample who made at least one comment in these subreddits in the 30 days before their ban. Each user from a banned subreddit is considered as treated and matched with a control user who did not participate in that subreddit.

Three analyses of the effect of the ban are performed, controlling for user-level and temporal factors. First, we estimate the effect of commenting in a banned subreddit, by comparing posts made in banned subreddits t months before the ban to posts made by the same users at the same time, in other subreddits. Second, using a difference-in-differences approach, we estimate the effect of banning a subreddit on authors' use of accommodation in (unbanned) subreddits they were active in for t months before and after the ban. This second analysis measures the spill-over effects of the ban on users' behaviors in other subreddits; the

(a) # Function Words

(b) Formality

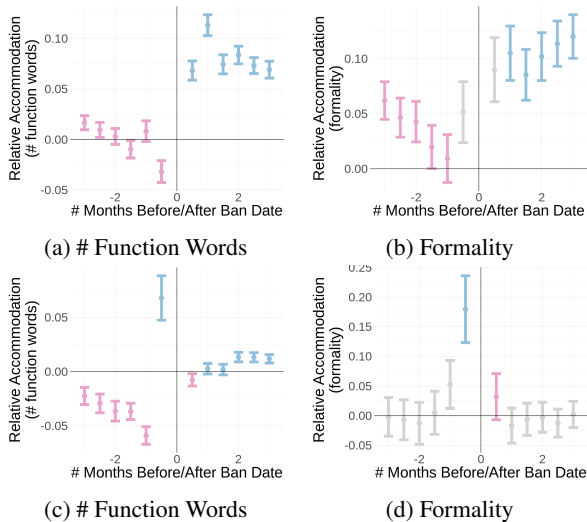(c) # Function Words

(d) Formality

Figure 9: When a subreddit is banned, (a,c) users who commented there tend to have higher LSM in other subreddits, (b) users tend to have higher LSM in function words in subreddits they migrate to, (d) LSM in formality tends to temporarily increases just before the subreddit is banned then returns to near-baseline levels in subreddits they migrate to.

difference-in-differences estimator uses users active in these subreddits at the same time, but not in a banned subreddit, as a control for temporal and subreddit-level effects. Third, we calculate the effect of the ban on commenting behavior in subreddits users migrated to (i.e., newly joined) after the ban was enacted. The difference-in-differences estimator compares accommodation in comments in the banned subreddits to comments in the subreddits these users migrated to; to isolate the effect of migration, the difference between the comments in the migrated and banned subreddits are compared against the spill-over effects in other subreddits that users were a part of during this time.

8.2 Results

Our results suggest that policy actions on Reddit, such as banning, have an effect on the level of accommodation by users. First, the level of subconscious accommodation tends to be lower in banned subreddits than other subreddits the users comment in during the 30 days before the ban (the effects are all below 0 in Figure 9c ($p < 2e - 16$).

Second, following the banning of a subreddit, users tend to change their LSM levels in other subreddits: Figure 9 shows that function-word-mirroring (Figure 9a) and formality-mirroring (Figure 9b) increase after a subreddit is banned. Our results suggest that users who had previously been

active in banned subreddits may have been making an effort to index agreeableness by accommodating (e.g., to avoid losing status in another community).

Third, changes in accommodation are initially amplified in subreddits that these users migrate to after their original community was banned. The comments left by these users in banned subreddits exhibit higher levels of accommodation than would be expected immediately before the ban and maintain higher subconscious accommodation in subreddits they migrated (Figures 9c and 9d $p < 2e - 16$). Since function-word mirroring is likely subconscious and formality-mirroring strategic (Section 4), our results suggest that users who had previously been active in banned subreddits may have, intrinsically, indexed agreeableness by accommodating (e.g., to gain status in their new community) but without making a conscious effort (e.g., because they were upset about the loss of a status). These users also increased LSM in the subreddit immediately before it was banned (e.g., perhaps to index agreeableness when warnings about the ban were issued).

9   Discussion and Conclusion

In this study, we performed a large-scale computational analysis on Reddit conversations to understand when LSM occurs and its effect on platform engagement. Overall, do our findings indicate that LSM frequently occurs in online conversations within Reddit, and that it exhibits complex nonlinear relationships with conversation metrics such as Karma, conversation lengths, or controversy scores, which suggests linguistic influence can affect conversation dynamics. Furthermore, we show that the degree of accommodation in conversations is related to greater levels of engagement both at conversation and platform levels. Our findings highlight the possibility of identifying LSM as an indicator of engagement and civil conversations and suggest ideas for building and maintaining online communities that promote constructive discourse.

In our experiments, we have assumed LSM as a unidirectional concept by measuring the exhibition of a particular style conditioned on the previous turn. However, LSM can occur in several different directions, such as the two speakers converging into a single style or even diverging to separate styles. While not in the scope of this study, the existence of such types of LSM in Reddit conversation threads can be studies in future research.

10 Ethical Considerations

This study was conducted only on observational data and did not require any human intervention. We did not use any information that could identify individuals or specific demographic groups, and all of our presented results were obtained through aggregation from millions of users and comments.

Acknowledgments

References

Meghan J Babcock, Vivian P Ta, and William Ickes. 2014. Latent semantic similarity and language style matching in initial dyadic interactions. *Journal of Language and Social Psychology*, 33(1):78–88.

Jason Baumgartner, Savvas Zannettou, Brian Keegan, Megan Squire, and Jeremy Blackburn. 2020. The pushshift reddit dataset.

A Burcu Bayram and Vivian P Ta. 2019. Diplomatic chameleons: Language style matching and agreement in international diplomatic negotiations. *Negotiation and Conflict Management Research*, 12(1):23–40.

Douglas Biber. 1988. *Variation across Speech and Writing*. Cambridge University Press.

Eshwar Chandrasekharan, Umashanthi Pavalanathan, Anirudh Srinivasan, Adam Glynn, Jacob Eisenstein, and Eric Gilbert. 2017. You can't stay here: The efficacy of reddit's 2015 ban examined through hate speech. *Proceedings of the ACM on Human-Computer Interaction*, 1(CSCW):1–22.

Tanya L Chartrand and John A Bargh. 1999. The chameleon effect: The perception–behavior link and social interaction. *Journal of personality and social psychology*, 76(6):893.

Robert B Cialdini. 2001. The science of persuasion. *Scientific American*, 284(2):76–81.

Cristian Danescu-Niculescu-Mizil, Michael Gamon, and Susan Dumais. 2011. Mark my words! linguistic style accommodation in social media. In *Proceedings of the 20th international conference on World wide web*, pages 745–754.

Cristian Danescu-Niculescu-Mizil, Lillian Lee, Bo Pang, and Jon Kleinberg. 2012. Echoes of power: Language effects and power differences in social interaction. In *Proceedings of the 21st international conference on World Wide Web*, pages 699–708.

Cristian Danescu-Niculescu-Mizil, Robert West, Dan Jurafsky, Jure Leskovec, and Christopher Potts. 2013. No country for old members: User lifecycle and linguistic change in online communities. In *Proceedings of the 22nd International Conference on World Wide Web*, WWW '13, page 307–318, New York, NY, USA. Association for Computing Machinery.

Daria Dayter and Thomas C. Messerli. 2022. Persuasive language and features of formality on the r/changemyview subreddit. *Internet Pragmatics*, 5(1):165–195.

Erik W Doxtader. 1991. The entwinement of argument and rhetoric: A dialectical reading of habermas' theory of communicative action. *Argumentation and Advocacy*, 28(2):51–63.

Gabriel Doyle and Michael C Frank. 2016. Investigating the sources of linguistic alignment in conversation. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 526–536.

Howard Giles, Nikolas Coupland, and IUSTINE Coupland. 1991. 1. accommodation theory: Communication, context, and. *Contexts of accommodation: Developments in applied sociolinguistics*, 1.

Howard Giles and Tania Ogay. 2007. Communication accommodation theory. In *Explaining Communication: Contemporary Theories and Exemplars*, chapter 16. Lawrence Erlbaum Associates, Inc.

Amy L Gonzales, Jeffrey T Hancock, and James W Pennebaker. 2010. Language style matching as a predictor of social dynamics in small groups. *Communication Research*, 37(1):3–19.

Jack Grieve. 2023. Register variation explains stylometric authorship analysis. *Corpus Linguistics and Linguistic Theory*, 19(1):47–77.

Francis Heylighen and Jean-Marc Dewaele. 1999. Formality of language: definition, measurement and behavioral determinants. Technical report, Center "Leo Apostel", Free University of Brussels.

Eduard Hovy. 1987. Generating natural language under pragmatic constraints. *Journal of Pragmatics*, 11(6):689–719.

Cornelia Ilie. 2021. Discussion, dispute or controversy? paradigms of conflict-driven parliamentary practices. *Journal of Language Aggression and Conflict*, 9(2):237–270.

Molly E. Ireland and Marlone D. Henderson. 2014. Language style matching, engagement, and impasse in negotiations. *Negotiation and Conflict Management Research*, 7(1):1–16.

Molly E Ireland and James W Pennebaker. 2010. Language style matching in writing: synchrony in essays, correspondence, and poetry. *Journal of personality and social psychology*, 99(3):549.

Molly E Ireland, Richard B Slatcher, Paul W Eastwick, Lauren E Scissors, Eli J Finkel, and James W Pennebaker. 2011. Language style matching predicts relationship initiation and stability. *Psychological science*, 22(1):39–44.

Elizabeth Jones, Cynthia Gallois, Victor Callan, and Michelle Barker. 1999. Strategies of accommodation: Development of a coding system for conversational interaction. *Journal of Language and Social Psychology*, 18(2):123–151.

Annemarie Kocab, Jayden Ziegler, and Jesse Snedeker. 2019. It takes a village: The role of community size in linguistic regularization. *Cognitive Psychology*, 114:101227.

Philipp Koncar, Simon Walk, and Denis Helic. 2021. Analysis and prediction of multilingual controversy on reddit. In *13th ACM Web Science Conference 2021*, pages 215–224.

Michael Lechner et al. 2011. The estimation of causal effects by difference-in-difference methods. *Foundations and Trends® in Econometrics*, 4(3):165–224.

Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov. 2019. Roberta: A robustly optimized bert pretraining approach. *arXiv preprint arXiv:1907.11692*.

Yelena Mejova, Amy X Zhang, Nicholas Diakopoulos, and Carlos Castillo. 2014. Controversy and sentiment in online news. *arXiv preprint arXiv:1409.8152*.

Antje S Meyer and Kathryn Bock. 1999. Representations and processes in the production of pronouns: Some perspectives from dutch. *Journal of Memory and Language*, 41(2):281–301.

Kate Muir, Adam Joinson, Rachel Cotterill, and Nigel Dewdney. 2016. Characterizing the linguistic chameleon: Personal and social correlates of linguistic style accommodation. *Human Communication Research*, 42(3):462–484.

Kate Muir, Adam Joinson, Rachel Cotterill, and Nigel Dewdney. 2017. Linguistic style accommodation shapes impression formation and rapport in computer-mediated communication. *Journal of Language and Social Psychology*, 36(5):525–548.

Sik Hung Ng and James J Bradac. 1993. *Power in language: Verbal communication and social influence*. Sage Publications, Inc.

Kate G Niederhoffer and James W Pennebaker. 2002. Linguistic style matching in social interaction. *Journal of Language and Social Psychology*, 21(4):337–360.

Bill Noble and Raquel Fernández. 2015. Centre stage: How social network position shapes linguistic coordination. In *Proceedings of the 6th workshop on cognitive modeling and computational linguistics*, pages 29–38.

Ellie Pavlick and Joel Tetreault. 2016. An empirical analysis of formality in online communication. *Transactions of the Association for Computational Linguistics*, 4:61–74.

James W Pennebaker and Cindy K Chung. 2007. Expressive writing, emotional upheavals, and health. In *Foundations of health psychology*. Oxford University Press.

James W Pennebaker, Martha E Francis, and Roger J Booth. 2001. Linguistic inquiry and word count: Liwc 2001. *Mahway: Lawrence Erlbaum Associates*, 71(2001):2001.

Sudha Rao and Joel Tetreault. 2018. Dear sir or madam, may I introduce the GYAFC dataset: Corpus, benchmarks and metrics for formality style transfer. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long Papers)*, pages 129–140, New Orleans, Louisiana. Association for Computational Linguistics.

Daniel M Romero, Roderick I Swaab, Brian Uzzi, and Adam D Galinsky. 2015. Mimicry is presidential: Linguistic style matching in presidential debates and improved polling numbers. *Personality and Social Psychology Bulletin*, 41(10):1311–1319.

RL Street. 1991. Accommodation in medical consultations. *Contexts of accommodation: Developments in applied sociolinguistics*, pages 131–156.

Yla R Tausczik and James W Pennebaker. 2010. The psychological meaning of words: Liwc and computerized text analysis methods. *Journal of language and social psychology*, 29(1):24–54.

Paul J Taylor and Sally Thomas. 2008. Linguistic style matching and negotiation outcome. *Negotiation and conflict management research*, 1(3):263–281.

Michael Willemyns, Cynthia Gallois, Victor J Callan, and Jeffery Pittam. 1997. Accent accommodation in the job interview: Impact of interviewer accent and gender. *Journal of Language and Social Psychology*, 16(1):3–22.

Walt Wolfram. 1974. *Sociolinguistic Aspects of Assimilation: Puerto Rican English in New York City*. Center for Applied Linguistics.

Yang Xu, Jeremy Cole, and David Reitter. 2018. Not that much power: Linguistic alignment is influenced more by low-level linguistic features rather than social power. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 601–610.