Not that much power: Linguistic alignment is influenced more by low-level linguistic features rather than social power

Yang Xu, and **Jeremy Cole** and **David Reitter**
College of Information Sciences and Technology
The Pennsylvania State University
yang.xu@psu.edu and jrcole@psu.edu and reitter@psu.edu

Abstract

Linguistic alignment between dialogue partners has been claimed to be affected by their relative social power. A common finding has been that interlocutors of higher power tend to receive more alignment than those of lower power. However, these studies overlook some low-level linguistic features that can also affect alignment, which casts doubts on these findings. This work characterizes the effect of power on alignment with logistic regression models in two datasets, finding that the effect vanishes or is reversed after controlling for low-level features such as utterance length. Thus, linguistic alignment is explained better by low-level features than by social power. We argue that a wider range of factors, especially cognitive factors, need to be taken into account for future studies on observational data when social factors of language use are in question.

1 Introduction

The effect of social power on language use in conversations has been widely studied. The Communication Accommodation Theory (Giles, 2008) states that the *social power* of speakers influence the extent to which conversation partners accommodate (or align, coordinate) their communicating styles towards them. This theory is supported by findings from qualitative studies on employment interviews (Willemyns et al., 1997), classroom talks (Jones et al., 1999), and the more recent data-driven studies on large online communities and court conversations (Danescu-Niculescu-Mizil et al., 2012; Jones et al., 2014; Noble and Fernández, 2015). In particular, Danescu-

Niculescu-Mizil et al. (2012) uses a probability-based measure of linguistic alignment to demonstrate that people align more towards conversation partners of higher power, i.e., the admin users in Wikipedia talk-page, and the justices in U.S. supreme court conversations, than those of lower power, i.e., the non-admin users and the lawyers.

However, while these results find sound explanations from socio-linguistic theories, they are still somewhat surprising from the perspective of cognitive mechanisms of language production, because the mutual alignment between interlocutors of in natural dialogue can be explained by an automatic and low-level priming process (Pickering and Garrod, 2004). It is known that the strength of alignment is sensitive to low-level linguistic features (e.g., words, syntactic structures etc.), such as temporal clustering properties (Myslín and Levy, 2016), syntactic surprisal measured by prediction error (Jaeger and Snider, 2013), and lexical information density (Xu and Reitter, 2018).

Then why, or under what mechanisms, can alignment be affected by the relatively high-level social perceptions of power as reported? Could it be the case that the effect of power on alignment is actually due to the other low level features in language, such as the ones mentioned above? Is the effect of power still observable, if we control for other factors? How large is the effect? Is it significant enough to be captured by computational measures of alignment? Answering these questions will help clarify the role of social factors in linguistic alignment, and improve our understanding of language production in general.

In this study, we conduct a two-step model analysis. First, we use a basic model that has two predictors, *count* (number of a certain linguistic marker in the preceding utterance) and *power* (power status of the preceding speaker), to predict the occurrence of the same marker in the follow-

ing utterance. Here, the linguistic markers are derived from 11 Linguistic Inquiry and Word Count (LIWC; Pennebaker et al., 2001) categories (e.g., *article*, *adverb*, etc.). With the basic model, the main effect of *count* characterizes the strength of alignment, and the interaction between *count* and *power* characterizes the effect of power on alignment (Section 3). Second, we use an extended model that includes a third predictor, *utterance length* (It is chosen as a typical low-level linguistic feature, discussed in Section 2.3), on top of the basic model. With the extended model, we aim to examine whether the inclusion of *utterance length* will influence the interaction between *count* and *power* (Section 4). Therefore, we can examine the extent to which the effect of power on alignment is confounded by low-level linguistic features.

To clarify, our goal is not to disprove the existence of social accommodation in dialogue. Nonetheless, it is important to distinguish between what is caused by automatic priming-based alignment and what is caused by high-level, intentional accommodation. As we will discuss, these are different processes with different predictions. Throughout this paper we use the term *alignment* to refer to the priming-based process, and accommodation to refer to the intentional process.

2 Related Work

2.1 Social power and linguistic alignment

The social factors of language use have been widely studied. Communication Accommodation Theory (Giles, 2008) posits that individuals adapt their communication styles to increase or decrease the social distance from their interlocutors. One factor that affects the adaptation of linguistic styles is *social power*. Typically, people of lower power converge their linguistic styles to those of higher power; for example, interviewees towards interviewers (Willemyns et al., 1997), or students towards teachers (Jones et al., 1999).

More recently, sensitive quantitative methods have been applied to this line of inquiry. Danescu-Niculescu-Mizil et al. (2012) computed the probability-based linguistic coordination measure among Wikipedia editors and participants of the US supreme court, and they showed that people with low power (e.g., lawyers, non-admins) exhibit greater coordination than people with high power (Justices, admins). Using the same data, Noble and Fernández (2015) found that linguis-

tic coordination is positively correlated with social network centrality, and this effect is even greater than the effect of power status distinction.

The aforementioned studies do not include low-level language features in their analysis and thus overlook the possibility that cognitive mechanisms may be able to more readily explain the data. Importantly, as we will later discuss, these studies use a measurement of alignment that we believe is more appropriately measuring the automatic process, rather than the intentional one.

2.2 Quantifying linguistic alignment

A variety of computational measures of linguistic alignment have been developed. Some quantify the increase in conditional probability of certain elements (words or word types) given that they have appeared earlier (Church, 2000; Danescu-Niculescu-Mizil et al., 2012). Some compute the proportion of repeated lexical entries or syntactic rules between two pieces of text (Fusaroli et al., 2012; Wang et al., 2014; Xu and Reitter, 2015). Some use the coefficients returned by generalized linear models (McCullagh, 1984; Breslow and Clayton, 1993; Lindstrom and Bates, 1990) as an index of alignment (Reitter and Moore, 2014). A large body of the existing computational measures intensively use LIWC (Pennebaker et al., 2001) to construct representations of language users' styles, which can be used to measure alignment with distance-like metrics (Niederhoffer and Pennebaker, 2002; Jones et al., 2014). Many of these approaches do not distinguish between different levels of linguistic analysis and different psycholinguistic processes (phonological, lexical, syntactic, etc.), and neither do we. Alignment is consistently present across these levels and processes, although it is not as clear in naturalistic language as it is in the constrained utterances of experiments, particularly at the syntactic level (Healey et al., 2014). We are concerned with the question of whether alignment is a socially linked, intentional adaptation process, as opposed to addressing any particular cognitive model.

More recently, Doyle et al. (2016) pointed out that most existing measures are difficult to compare, and emphasized the need for a universal measure. The Hierarchical Alignment Model (HAM; Doyle et al., 2016) and Word-Based HAM (WHAM; Doyle and Frank, 2016) use statistical inference techniques, which out-perform other

measures in terms of robustness of capturing linguistic alignment in social media conversations.

In this study, we choose to use generalized linear models to quantify linguistic alignment, avoiding issues with more complex, and less inspectable models. For instance, the commonly used probability based methods and their more advanced variants (HAM and WHAM) lack the flexibility to jointly examine multiple factors (e.g., speaker groups, utterance length etc.) that influence alignment. Another issue is that they do not take into account the number of occurrences of linguistic markers, which is known to affect alignment (see Section 2.3). Conversely, though linear models do not give an accurate per-speaker estimate of alignment (which we do not need for the purpose of this study), they do provide the ability to examine multiple factors that influence alignment by simply including multiple predictors in the model. As should be clear, a generalized linear model also already takes into account baseline usage with a fitted intercept. Given these considerations, we use generalized linear models for quantitative analysis. The formulation of our models is described in Sections 3.2 and 4.1.

2.3 Cognitive constraints on linguistic alignment: why utterance length matters

There are many, at times competing, cognitive explanations of linguistic alignment in both comprehension and production. Jaeger and Snider (2013) explained alignment as a consequence of expectation adaptation, and they found that stronger alignment is associated with syntactic structures that have higher surprisal (roughly speaking, less common). Alignment in language production can also be modeled as a general memory phenomenon (Reitter et al., 2011), which explains a number of known interaction effects. Myslín and Levy (2016) found that sentence comprehension is faster when the same syntactic structure clusters in time in prior experience than when it is evenly spaced in time. Myslín and Levy (2016) cast comprehension priming as the rational expectation for repetition of stimuli. Though this result is not directly related to comprehension-to-production priming, it makes sense to anticipate that production could also be sensitive to the clustering patterns of linguistic elements, because comprehension and production are closely coupled processes (Pickering and Garrod, 2007).

Utterance length, i.e., the number of words in utterance, is a feature that closely relates to both surprisal and clustering properties. Longer utterances tend to have higher syntactic surprisal (Xu and Reitter, 2016a), and it is reasonable to assume they tend to contain more evenly distributed stimuli. Thus, utterance length is a low-level linguistic feature that correlates with many of the causes of alignment. In this way, we use utterance length as a stand-in for low-level linguistic features as a whole when comparing it with social power, a much higher-level feature. Examining alignment (in social science research and elsewhere) therefore calls for controlling sentence length.

3 Experiment 1: Basic model

In Experiment 1, we justify the practice of using generalized linear models to quantify linguistic alignment. We compare two ways of characterizing the occurrence of LIWC-derived markers in a preceding utterance, binary presence and numeric count, to determine which results in better model. We use an interaction term in the model to quantify the effect of the power status of speakers on linguistic alignment, which serves as the basis for the following sections.

3.1 Corpus data

We use two datasets compiled by Danescu-Niculescu-Mizil et al. (2012): Wikipedia talk-page corpus (*Wiki*) and a corpus of United States supreme court conversations (*SC*). Wiki is a collection of conversations from Wikipedia editor's talk Pages[1], which contains 125,292 conversations contributed by 30,732 editors. SC is a collection of conversations from the U.S. Supreme Court Oral Arguments[2], with 51,498 utterances making up 50,389 conversational exchanges, from 204 cases involving 11 Justices and 311 other participants (lawyers or amici curiae).

A conversation consists of a sequence of utterances, $\{u_i\}(i = 1, 2, \ldots, N)$, where N is the total number of utterances in the conversation. Because people take turns to talk in conversation, u_i and u_{i+1} are always from different speakers. Since our interest here is the alignment between different speakers (as opposed to within the same speaker), we use a sliding window of size 2 to go through

---

[1] http://en.wikipedia.org/wiki/Wikipedia:Talk_page_guidelines
[2] http://www.supremecourt.gov/oral_arguments/

the whole conversation, generating a sequence of adjacent utterance pairs, $\{\langle prime_i, target_i \rangle\}(i = 1, 2 \ldots N - 1)$.

Next, we process each utterance u_i by counting the number of occurrences of 14 linguistic markers that are derived from LIWC categories, resulting in 14 counts for each utterance. These 14 linguistic markers are: high frequency adverbs (*adv*), articles (*art*), auxiliary verbs (*auxv*), certainty (*certain*), conjunctions (*conj*), discrepancy (*discrep*), exclusion (*excl*), inclusion (*incl*), impersonal pronouns (*ipron*), negations (*negate*), personal pronouns (*ppron*), prepositions (*prep*), quantifiers (*quant*), and tentativeness (*tentat*). These fourteen markers come from taking the union of the 8 markers used by Danescu-Niculescu-Mizil et al. (2012) and the 11 markers used by Doyle and Frank (2016), which are the main studies we wanted to compare with.

3.2 Statistical models

We formulate alignment as the impact of using certain linguistic elements in the preceding utterance on their chance to appear again in the following utterance. In the language of generalized linear models, we use the occurrence of linguistic markers in *target* as the response variable and the predictor is their occurrence in *prime*. These occurrences can be represented as either a boolean or a count. Thus alignment is characterized by the β coefficient of the predictor, which allows the model to distinguish the prevalence of *Occurrence* or another feature in primed situations as compared to its prior in the corpus. Factors that may influence alignment (e.g., social power) can then be examined by adding a corresponding interaction term to the model.

Our first step, then, is to replicate the previous studies' findings of the effect of social power on alignment. Two models were fitted, predicting the *presence* of the linguistic marker *m* in *target* utterance over its absence. We fit models both corresponding to a binary predictor (C_{presence}) and a count-based one (C_{count}). Both models include a second binary predictor, C_{power}, indicating the power status of the *prime* speaker (*high* vs. *low*), and its interaction with C_{presence} and C_{count}, respectively. Additionally, random intercepts on linguistic marker and *target* speaker are fitted, based on the consideration that individuals might have different levels of alignment towards

different markers. C_{count} is log-transformed to maximize model fit according to Bayesian Information Criterion; this is commensurate with standard psycholinguistic practice and known cumulative priming and memory effects. Equation (1) shows the count-based model. To reiterate, the interaction term $C_{\text{count}} * C_{\text{power}}$ characterizes the effect of power on alignment.

$$
\begin{aligned}
\text{logit}(m) &= \ln \frac{p(m \text{ in } target)}{p(m \text{ not in } target)} \\
&= \beta_0 + \beta_1 C_{\text{count}} + \beta_2 C_{\text{power}} \\
&\quad + \beta_3 \mathbf{C_{\text{count}}} * \mathbf{C_{\text{power}}}
\end{aligned} \tag{1}
$$

3.3 Model coefficients

The main effects of C_{presence} and C_{count} are significant ($p < 0.001$) and positive in both corpora (SC: $\beta_{\text{presence}} = 0.439$, $\beta_{\text{count}} = 0.291$; Wiki: $\beta_{\text{presence}} = 0.440$, $\beta_{\text{count}} = 0.395$), which captures the linguistic alignment from *prime* to *target*. However, there is difference in how alignment is influenced by power between the two corpora: In SC, $C_{\text{count}} * C_{\text{power}}$ is significant ($\beta = 0.078$, $p < 0.001$), but $C_{\text{presence}} * C_{\text{power}}$ is non-significant; In Wiki, on the contrary, $C_{\text{presence}} * C_{\text{power}}$ is marginally significant ($\beta = 0.014$, $p = .055$), but $C_{\text{count}} * C_{\text{power}}$ in not significant. No collinearity is found between C_{count} (or C_{presence}) and C_{power} (Pearson correlation $r < 0.2$).

To explore why using C_{presence} vs. C_{count} results in different significance levels for SC and Wiki, we fit a individual linear model for each linguistic marker, using 14 *disjoint* subsets of each corpus. We present the z scores and significance levels of the two interaction terms are reported in Table 1. First, in SC the interaction term $C_{\text{presence}} * C_{\text{power}}$ is significant for 9 out of 14 markers. In Wiki, $C_{\text{count}} * C_{\text{power}}$ is significant for 5 out of 14 markers. This suggests that the interaction between the occurrence of linguistic markers and the power status of speakers exists within a subset of the linguistic categories, but not across all of them. Thus, we consider this first experiment a replication of past findings of the effect of social power on alignment: social power has a significant effect across certain markers, but its overall effect is neutralized in the full model since some markers at not significant. This analysis also revealed that $C_{\text{count}} * C_{\text{power}}$ is more reliable in capturing this effect, which is what we will use in the following experiment.

Table 1: Summary of the 14 models that fit individual markers on disjoint data subsets. Wald's z-score and significance level (*** for $p < 0.001$, ** for $p < 0.01$, * for $p < 0.05$, and † for $0.05 < p < 0.1$) of the interaction terms ($C_{\text{presence}} * C_{\text{power}}$ or $C_{\text{count}} * C_{\text{power}}$) are reported.

Marker	z score			
	$C_{\text{presence}} * C_{\text{power}}$		$C_{\text{count}} * C_{\text{power}}$	
	SC	Wiki	SC	Wiki
adv	1.19	-0.33	**6.16*****	-0.40
art	**1.99***	0.36	**4.60*****	1.27
auxv	**3.72*****	-0.62	**5.81*****	-0.83
certain	-0.02	**3.19****	**1.94**†	**2.84****
conj	0.54	-0.20	**6.79*****	0.39
discrep	**5.44*****	-0.05	**8.03*****	0.25
excl	-0.53	**1.96***	**2.94****	**2.16***
incl	**2.86****	0.80	**5.24*****	**2.15***
ipron	**6.84*****	**1.70**†	**10.22*****	**1.90**†
negate	**2.83****	**3.14****	**5.49*****	**3.11****
ppron	**2.74****	**-1.86**†	1.29	-1.13
prep	**4.76*****	**2.37***	**6.87*****	-0.19
quant	0.89	1.01	**4.14*****	-0.04
tentat	**3.69*****	0.17	**4.52*****	-0.78

3.4 Visualizing the effect of power

To better understand the interaction term $C_{\text{count}} * C_{\text{power}}$, we divide the data into two groups by whether C_{power} is *high* or *low*, and fit a model on each of the groups. In the models we include only one predictor C_{count} (see Equation (2)). Then we compare the main effects (β_1 coefficients) from the two groups.

$$\text{logit}(m) = \beta_0 + \beta_1 C_{\text{count}} \qquad (2)$$

Unsurprisingly, the main effects of C_{count} are significant for both groups ($p < 0.001$). But more importantly, the β_1 coefficients of the high power group are larger than those of the low power group. For SC, the difference is very salient: $\beta_1^{\text{high}} = 0.416$ ($SE = 0.006$), $\beta_1^{\text{low}} = 0.272$ ($SE = 0.005$). For Wiki, the difference is smaller: $\beta_1^{\text{high}} = 0.424$ ($SE = 0.007$), $\beta_1^{\text{low}} = 0.386$ ($SE = 0.005$). This is in line with the non-significant coefficient of $C_{\text{count}} * C_{\text{power}}$ in Wiki. In fact, the models of Wiki are fitted on a subset of data that contain the 5 (out of 14) markers that have significant coefficients of $C_{\text{count}} * C_{\text{power}}$ in the individual models shown in Table 1 (*certain*, *excl*, *incl*, *ipron*, *negate*), so that the difference in slopes is presented at maximal degree.

In Figure 1 we illustrate the β_{high} and β_{low} coefficients of C_{count} by plotting the predicted probability (the reversed logit transformation of the left-hand side term of Equation (2)) against C_{count} (log-transformed). It is obvious that the slope of β_{high} is larger than that of β_{low} (more salient in SC), indicating the significant interaction between C_{count} and C_{power}.

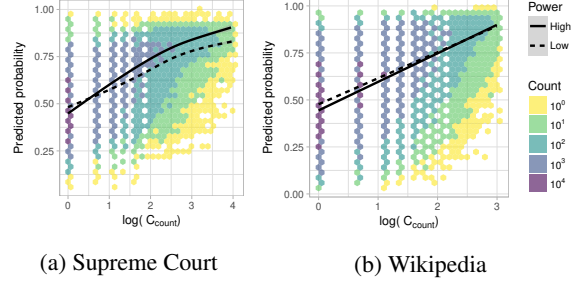(a) Supreme Court   (b) Wikipedia

Figure 1: The predicted probability of marker appearing in *target* (the reverse logit transform of the left hand side of Equation (2)) against the number of markers in *prime*, i.e., C_{count} (log-transformed), grouped by the power of *prime* speaker, i.e., *high* vs. *low*. Divergent slopes indicate significant interactions. Colored hexagons indicate the number of data points within that region.

3.5 Discussion

The occurrence of linguistic markers in *prime* is a strong predictor of whether the same marker will appear again in *target*. The coefficients of C_{count} can be viewed as indicators of the linguistic alignment between interlocutors: larger positive βs indicate stronger alignment, while smaller or even negative βs indicate weaker and reverse alignment, respectively (not found in our data).

Our results confirm the previously reported effect of power on linguistic alignment. The significant β' coefficient of $C_{\text{count}} * C_{\text{power}}$ means that the β of C_{count} is dependent on C_{power}. In other words, the strength of alignment varies significantly depending on different power levels (i.e., *high* vs. *low*) of the *prime* speaker (reflected by the different slopes in Figure 1). However, we need to keep in mind that this affirmative finding is *not* safe, because it based on a simple model that has only one key predictor, C_{power}. According to our hypothesis, the strength of alignment can be influenced by a lot of low-level linguistic features, and we are not sure yet if the effect of power will still be visible after we includes more predictors representing

those features. This will be the next step experiment.

Additionally, the results also suggest that the influence of power on linguistic alignment is better characterized by the more fine-grained cumulative effect of linguistic markers than when it is simply explained by the mere difference between their *absence* or *presence*. Thus, we will discard C_{presence} and proceed with C_{count}.

4 Experiment 2: Extended model

In our first experiment, we replicated the effect of *prime* speakers' power status on the linguistic alignment from *target* speakers, from the significant interaction term $C_{\text{count}} * C_{\text{power}}$. Now, we want to determine if the effect of power remains significant after taking into account utterance length. As discussed, our hypothesis is that alignment (as measured by changes in probability of using LIWC categories) is best explained by low-level linguistic features that would be taken into account by an automatic priming process.

4.1 Statistical models

We add a new predictor to Equation (1), C_{pLen}, which is the number of words in *prime*, resulting in an extended model shown in Equation (3). We are interested to see if β_4 remains significant when the other interaction terms (with corresponding coefficients β_5, β_6 and β_7) are added.

$$
\begin{aligned}
\text{logit}(m) = \ln \frac{p(m \text{ in } target)}{p(m \text{ not in } target)} \\
= \beta_0 + \beta_1 C_{\text{count}} + \beta_2 C_{\text{power}} + \beta_3 C_{\text{pLen}} \\
+ \beta_4 C_{\text{count}} * C_{\text{power}} \\
+ \beta_5 C_{\text{count}} * C_{\text{pLen}} \\
+ \beta_6 C_{\text{power}} * C_{\text{pLen}} \\
+ \beta_7 C_{\text{count}} * C_{\text{power}} * C_{\text{pLen}}
\end{aligned}
\tag{3}
$$

Note that we used the same subset of Wiki as used in Section 3.4 (using the five most significant LIWC categories), so that the strongest effect of $C_{\text{count}} * C_{\text{power}}$ is considered.

4.2 Model coefficients

The coefficients of the full model are in Table 2. Surprisingly, the coefficient of $C_{\text{count}} * C_{\text{power}}$ is significantly *negative* in SC, and non-significant in Wiki (see highlighted rows), which are in contrast to the *positive* coefficients of the same term

Table 2: Summary of the model described in Equation (3): β coefficients, Wald's z-score and significance level (*** for $p < 0.001$, ** for $p < 0.01$, * for $p < 0.05$) for all predictors and interactions.

Corpus	Predictor	β	z
SC	Intercept	0.360	2.40*
	C_{count}	0.213	26.92***
	C_{power}	-0.060	-3.39***
	C_{pLen}	0.080	13.03***
	$\boldsymbol{C_{\text{count}} * C_{\text{power}}}$	**-0.103**	**-9.95*****
	$C_{\text{count}} * C_{\text{pLen}}$	-0.066	-15.35***
	$C_{\text{power}} * C_{\text{pLen}}$	0.231	25.25***
	$C_{\text{count}} * C_{\text{power}} * C_{\text{pLen}}$	0.036	4.79***
Wiki	Intercept	0.330	1.40
	C_{count}	0.149	31.11***
	C_{power}	-0.074	-10.52***
	C_{pLen}	0.179	40.80***
	$\boldsymbol{C_{\text{count}} * C_{\text{power}}}$	**0.001**	**0.14**
	$C_{\text{count}} * C_{\text{pLen}}$	0.022	6.13***
	$C_{\text{power}} * C_{\text{pLen}}$	0.042	5.52***
	$C_{\text{count}} * C_{\text{power}} * C_{\text{pLen}}$	-0.010	-1.61

in Table 1. It indicates that the observed effect of power on alignment depends on the presence of C_{pLen} in the model. No collinearity is found between C_{power} and other predictors: Pearson correlation $r < 0.2$; Variance inflation factor (VIF) is low (< 2.0) (O'brien, 2007).

To further demonstrate how the coefficient of $C_{\text{power}} * C_{\text{count}}$ is dependent on C_{pLen}, we remove $C_{\text{count}} * C_{\text{pLen}}$, $C_{\text{power}} * C_{\text{pLen}}$ and $C_{\text{count}} * C_{\text{power}} * C_{\text{pLen}}$ from Equation (3) stepwisely, and examine $C_{\text{count}} * C_{\text{power}}$ in the corresponding remaining models. z-scores, significance levels, and the Akaike information criterion (AIC) score (Akaike, 1998) of the remainder models are reported in Table 3. In the full model, and when $C_{\text{count}} * C_{\text{power}} * C_{\text{pLen}}$ or $C_{\text{count}} * C_{\text{pLen}}$ is removed from the model, the coefficients of $C_{\text{power}} * C_{\text{count}}$ are significantly negative in SC and non-significant in Wiki. Only when $C_{\text{power}} * C_{\text{pLen}}$ is removed, the coefficients of $C_{\text{count}} * C_{\text{power}}$ become significantly positive (the last two rows in Table 3). However, the models that have negative or non-significant coefficient for $C_{\text{power}} * C_{\text{count}}$ have lower AIC scores than those that have positive coefficient (The full model has the lowest AIC score), which indicates that the former ones have higher quality. Altogether, the stepwise analysis not only indicates that the positive interaction between C_{power} and C_{count} shown in our basic model (Section 3) is unreliable, but also suggests that a negative interaction (SC) or

non-significant interaction is more preferable.

## 4.3 Visualizing interaction effect

To illustrate how the interaction $C_{\text{power}} * C_{\text{count}}$ diminishes after adding C_{pLen} into the extended model, we cluster different ranges of C_{pLen} and determine how the amount of priming changes with C_{count} w.r.t. different combinations of C_{power} and C_{pLen}. This is a common practice to interpret linear models with three-way interactions (Houslay, 2014).

To cluster, we first compute the mean of C_{pLen} (i.e., the average utterance length), M_{pLen}. Then we divide the data by whether C_{pLen} is above or below M_{pLen}. Then we compute the mean of C_{pLen} for the upper and lower parts of data, resulting in M_{pLen}^{L} and M_{pLen}^{S} respectively (*L* for long and *S* for short). Now, we can replace the continuous variable C_{pLen} to a categorical and ordinal one that has two values, $\{M_{\text{pLen}}^{S}, M_{\text{pLen}}^{L}\}$, which represent the length of relatively short and long utterances respectively. Together with the other categorical variable, C_{power}, which has two values, *high* and *low*, we have four combinations: $C_{\text{pLen}} = M_{\text{pLen}}^{S}$ and $C_{\text{power}} = high$ (SH), $C_{\text{pLen}} = M_{\text{pLen}}^{L}$ and $C_{\text{power}} = high$ (LH), $C_{\text{pLen}} = M_{\text{pLen}}^{S}$ and $C_{\text{power}} = low$ (SL), $C_{\text{pLen}} = M_{\text{pLen}}^{L}$ and $C_{\text{power}} = low$ (LL). In Figure 2 we plot the smoothed regression lines of predicted probability against C_{count}, w.r.t. the above four groups of C_{pLen} and C_{power} combinations. Here C_{count} is not log-transformed, because it better demonstrates the trend of the fitted regression lines.

Figure 2 intuitively shows that C_{pLen} is a more determinant predictor than C_{power}. Division by power, i.e., *high* (SH and LH groups) vs. *low* (SL and LL groups), does not result in a salient difference in slopes, as it can be seen that the slopes of *high* (solid) and *low* (dashed) power lines do not differ much from each other within the same *prime* utterance length group (indicated by color). However, division by *prime* utterance length, i.e. *short* (SH and SL) vs. *long* (LH and LL), results in very significant differences in slopes: in Figure 2a, *short* C_{pLen} group (orange) has larger slopes than *long* C_{pLen} group (blue), while in Figure 2b, *short* group has smaller slopes than *long* group.

4.4 Discussion

Adding C_{pLen} to the model has strong impact on the previous conclusion about the effect of power
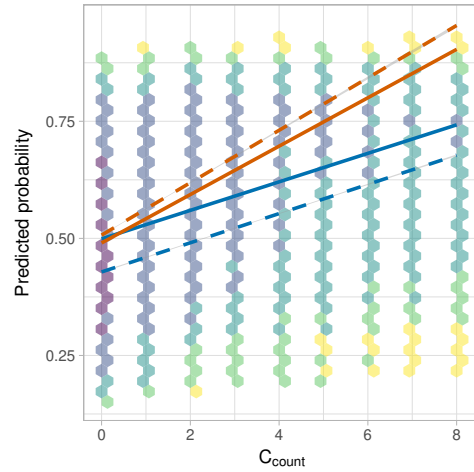
(a) Supreme Court
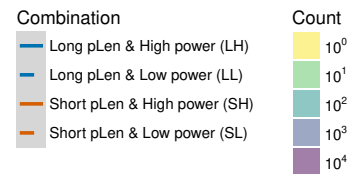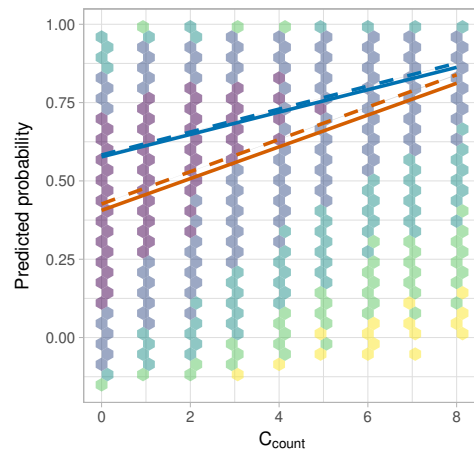
(b) Wikipedia

Figure 2: The predicted probability of marker appearing in *target* against C_{count}, grouped by the four combinations of C_{pLen} (*long* vs. *short*, indicated by color) and C_{power} (*high* vs. *low*, indicated by line type): LH, LL, SH, and SL. Colored hexagon indicates the number of data points.

on alignment. First of all, we find a negative interaction between C_{count} and C_{power} in SC and a non-significant effect in Wiki, which is contrary to the previous findings reported by Danescu-Niculescu-Mizil et al. (2012). Moreover, we doubt the reliability of a positive interaction because the valence of its β varies when other interaction terms (associated with C_{pLen}) are removed or added, and a negative or non-significant interaction is preferred

Table 3: Wald's z-score and significance level (*** for $p < 0.001$) of the $C_{\text{count}} * C_{\text{power}}$ term, and the AIC scores of the remainder models after removing other interaction terms from the full model stepwisely. The full model is described in Equation (3).

Remainder model	SC		Wiki	
	z-score	AIC	z-score	AIC
Full	-9.95***	697588	0.14	890685
Full $- C_{\text{count}} * C_{\text{power}} * C_{\text{pLen}}$	-8.75***	697609	-0.62	890686
Full $- C_{\text{count}} * C_{\text{power}} * C_{\text{pLen}}$ $\quad - C_{\text{count}} * C_{\text{pLen}}$	-5.61***	697838.9	-0.74	890723.5
Full $- C_{\text{count}} * C_{\text{power}} * C_{\text{pLen}}$ $\quad - C_{\text{power}} * C_{\text{pLen}}$	10.90***	698254.7	3.85***	890726.7
Full $- C_{\text{count}} * C_{\text{power}} * C_{\text{pLen}}$ $\quad - C_{\text{count}} * C_{\text{pLen}}$ $\quad - C_{\text{power}} * C_{\text{pLen}}$	15.02***	698461.8	3.72***	890763.8

by a simple model selection criterion.

Second, there is a significant interaction between C_{count} and C_{pLen}, though it is in different directions for the two corpora: negative β in SC and positive β in Wiki. Both observations have some theoretical justification from previous studies. Myslín and Levy (2016)'s work is in favor of the negative β: language comprehension is facilitated by the clustering of linguistic stimuli in time. In our case, the linguistic markers in the utterance of speaker *A* function as stimuli to speaker *B*. A longer utterance means that all the stimuli span wider in time, and thus demonstrate less clustering, which make the stimuli less salient features for speaker *B* to adapt to. This in turn causes speaker *B* to be less likely reuse those stimuli in the near future. Meanwhile, evidence from the line of works on surprisal and syntactic priming supports the positive β. In syntactic alignment, structures with higher surprisal (less common) are associated with stronger alignment (Jaeger and Snider, 2013; Reitter and Moore, 2014). Since surprisal has been found to be closely related with utterance length in dialogue (Genzel and Charniak, 2003; Xu and Reitter, 2016b,a), it is reasonable to expect that longer utterances receive stronger alignment because they contain content of higher surprisal.

The discrepancy between Wiki and SC in terms of the direction of $C_{\text{count}} * C_{\text{pLen}}$ is an interesting phenomenon to explore, because it can tell us something about how the form of dialogue (Wiki consists of online conversations and SC consists of face-to-face ones) affects the underlying cognitive mechanism of language production.

Regardless, our main finding is that low-level linguistic features, such as utterance length, have a strong effect on linguistic alignment. These effects are an important confound to take into account when examining higher-level features, such as social power. In particular, the effect of social power cannot be reliably detected by linear models once introducing utterance length.

Another interesting piece of result is the significant interaction term $C_{\text{power}} * C_{\text{pLen}}$, which implies that the power status of speaker and how long he/she tends to speak are not totally unrelated. Significant but weak correlation are found between C_{power} and C_{pLen} (using Pearson's correlation score): $r = -0.059$ in SC; $r = -0.018$ in Wiki. This correlation may show some kind of a linguistic manifestation of social power, but since it is not directly related to the alignment process, we do not further discuss it in this paper.

In summary of the results, we conjecture that the previously reported effect of power (Danescu-Niculescu-Mizil et al., 2012) is likely to be caused by the correlation between power status and utterance length, though further investigation is needed to confirm this. Moreover, utterance length is just one simple factor, and there are many more other linguistic features that can correlate with social power: e.g., the surprisal based measure of lexi-

cal information etc.

5 Conclusion

To sum up, our findings suggest that the previously reported effect of power on linguistic alignment is not reliable. Instead, we consistently align towards language that shares certain low-level features. We call for the inclusion of a wider range of factors in future studies of social influences on language use, especially low-level but interpretable cognitive factors. Perhaps in most scenarios, alignment is primarily influenced by linguistic features themselves, rather than social power.

We are not denying the existence of accommodation caused by the social distance between interlocutors. However, we want to stress the difference between the priming-induced alignment at lower linguistic levels and the intentional accommodation that is caused by higher-level perception of social power. The latter should be a relatively stable effect that is independent on the low-level linguistic features. In particular, our findings suggest that the probability change of LIWC categories is more likely to be a case of automatic alignment, rather than an intentional accommodation, because it is better explained by lower-level linguistic features (utterance length). Therefore, we suggest that future work on social power and language use should consider other (maybe higher-level) linguistic elements.

Acknowledgement

References

Hirotogu Akaike. 1998. Information theory and an extension of the maximum likelihood principle. In *Selected Papers of Hirotugu Akaike*, pages 199–213. Springer.

Norman E Breslow and David G Clayton. 1993. Approximate inference in generalized linear mixed models. *Journal of the American Statistical Association*, 88(421):9–25.

Kenneth W Church. 2000. Empirical estimates of adaptation: the chance of two Noriega's is closer to $p/2$ than p^2. In *Proceedings of the 18th Conference on Computational Linguistics*, volume 1, pages 180–186, Saarbrücken, Germany.

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, Lyon, France.

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*, pages 526–536, Berlin, Germany.

Gabriel Doyle, Dan Yurovsky, and Michael C Frank. 2016. A robust framework for estimating linguistic alignment in Twitter conversations. In *Proceedings of the 25th International Conference on World Wide Web*, pages 637–648, Montreal, Canada.

Riccardo Fusaroli, Bahador Bahrami, Karsten Olsen, Andreas Roepstorff, Geraint Rees, Chris Frith, and Kristian Tylén. 2012. Coming to terms quantifying the benefits of linguistic coordination. *Psychological Science*, 23(8):931–939.

Dmitriy Genzel and Eugene Charniak. 2003. Variation of entropy and parse trees of sentences as a function of the sentence number. In *Proceedings of the 2003 Conference on Empirical Methods in Natural Language Processing*, pages 65–72. Association for Computational Linguistics.

Howard Giles. 2008. Communication accommodation theory. In L. A. Baxter and D. O. Braithewaite, editors, *Engaging theories in interpersonal communication: Multiple perspectives*, pages 161–173. Sage, Thousand Oaks, CA.

Patrick G. T. Healey, Matthew Purver, and Christine Howes. 2014. Divergence in dialogue. *PLOS ONE*, 9(6):1–6.

Thomas M. Houslay. 2014. Understanding 3-way interactions between continuous variables. https://tomhouslay.com/2014/03/21/understanding-3-way-interactions-between-continuous-variables/.

T Florian Jaeger and Neal E Snider. 2013. Alignment as a consequence of expectation adaptation: Syntactic priming is affected by the prime's prediction error given both prior and recent experience. *Cognition*, 127(1):57–83.

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.

Simon Jones, Rachel Cotterill, Nigel Dewdney, Kate Muir, and Adam Joinson. 2014. Finding zelig in

text: A measure for normalizing linguistic accommodation. In *25th International Conference on Computational Linguistics*, Bath, UK.

Mary J Lindstrom and Douglas M Bates. 1990. Nonlinear mixed effects models for repeated measures data. *Biometrics*, pages 673–687.

Peter McCullagh. 1984. Generalized linear models. *European Journal of Operational Research*, 16(3):285–292.

Mark Myslín and Roger Levy. 2016. Comprehension priming as rational expectation for repetition: Evidence from syntactic processing. *Cognition*, 147:29–56.

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 CMCL 2015*, pages 29–38, Denver, CO.

Robert M O'brien. 2007. A caution regarding rules of thumb for variance inflation factors. *Quality & Quantity*, 41(5):673–690.

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

Martin J Pickering and Simon Garrod. 2004. Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, 27(02):169–190.

Martin J Pickering and Simon Garrod. 2007. Do people use language production to make predictions during comprehension? *Trends in cognitive sciences*, 11(3):105–110.

David Reitter, Frank Keller, and Johanna D Moore. 2011. A computational cognitive model of syntactic priming. *Cognitive Science*, 35(4):587–637.

David Reitter and Johanna D Moore. 2014. Alignment and task success in spoken dialogue. *Journal of Memory and Language*, 76:29–46.

Yafei Wang, David Reitter, and John Yen. 2014. Linguistic adaptation in conversation threads: Analyzing alignment in online health communities. In *Proceedings of Cognitive Modeling and Computational Linguistics. Workshop at the Annual Meeting of the Association for Computational Linguistics*.

Michael Willemyns, Cynthia Gallois, Victor Callan, and J Pittam. 1997. Accent accommodation in the employment interview. *Journal of Language and Social Psychology*, 15(1):3–22.

Yang Xu and David Reitter. 2015. An evaluation and comparison of linguistic alignment measures. In *Proceedings of Cognitive Modeling and Computational Linguistics (CMCL)*, pages 58–67, Denver, DO.

Yang Xu and David Reitter. 2016a. Convergence of syntactic complexity in conversation. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 443–448, Berlin, Germany.

Yang Xu and David Reitter. 2016b. Entropy converges between dialogue participants: Explanations from an information-theoretic perspective. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 537–546, Berlin, Germany.

Yang Xu and David Reitter. 2018. Information density converges in dialogue: Towards an information-theoretic model. *Cognition*, 170:147–163.