Alignment at Work: Accommodation and Enculturation in Corporate Communication*

Gabriel Doyle
Department of Psychology
Stanford University
Stanford, CA 94305
gdoyle@stanford.edu

Amir Goldberg
Graduate School of Business
Stanford University
Stanford, CA 94305
amirgo@stanford.edu

Sameer B. Srivastava
Haas School of Business
UC Berkeley
Berkeley, CA 94720
srivastava@haas.berkeley.edu

Michael C. Frank
Department of Psychology
Stanford University
Stanford, CA 94305
mcfrank@stanford.edu

ABSTRACT

Cultural fit is an elusive construct that is often assumed to play an important role in individual, group, and organizational success. Most existing approaches to measuring culture are static, making it difficult to understand how cultural dynamics relate to success. By contrast, we develop a directed, dynamic measure of "linguistic alignment," which estimates the extent to which one person's word use is influenced by another's, and use this measure to trace employees' enculturation trajectories across a large, multi-year corpus of corporate emails.

We show that the employees' changes in language use, especially pronouns, are consistent with changes in their status; new hires' word usage moves toward the organization's average behavior as they assimilate, suggesting that directed linguistic alignment is an appropriate measure of cultural fit. This interpretation is further supported by the finding that predictive classifiers trained on the first six months of an employee's communication can predict eventual employment outcomes (i.e., leaving the company voluntarily or involuntarily) at levels above chance. Finally, we use the employment outcome classifiers to better understand the nature of alignment, compare the predictive quality of the traditional, lexical formulation of linguistic alignment against a novel semantic alignment scheme, and find that the semantic formulation is even more informative.

1. INTRODUCTION

Entering a new group can be difficult. Whether transitioning to a new school, showing up for the first day at a new job, or making one's first post to an established online community, the newcomer faces the daunting task of trying to understand the group's established culture and how to fit into it. In most cases, increased cultural fit is beneficial. Within organizational contexts, in particular, fitting in culturally has been shown to yield positive career outcomes such as faster time-to-promotion, higher performance ratings, and reduced risk of being fired [21, 16].

In the present research, we apply concepts and methods from the psychology of language to more fully illuminate the dynamics of enculturation. In particular, we develop a directed, dynamic measure of linguistic alignment that reveals how newcomers linguistically accommodate their colleagues as they assimilate into a group, and how they linguistically diverge from colleagues prior to their departure from the group. Importantly, our approach distinguishes between an individual's (1) *base rate* of word use and (2) *linguistic alignment* with interlocutors, and estimates changes in these two parameters, as they relate to different linguistic categories, during the individual's initial entry and terminal period prior to exit. The distinction between baseline and alignment allows to differentiate between internalization of a linguistic norm and accommodation to a norm when instantiated by an interlocutor during interaction.

We demonstrate the utility of this approach using personnel records and a corpus of internal email communications from a mid-sized technology company over a seven-year period. We show that the changes in language use, especially pronouns, are consistent with successful assimilation into a group in an employee's first six months and separation from the group in their last six months, suggesting that directed linguistic alignment is a useful indicator of cultural fit. This interpretation is further supported by results based on a classifier trained on only the first six months of an employee's communications, which can predict eventual employment outcomes at levels above chance. In particular, the rate of change in an employee's baseline and alignment in these first months is critical in determining whether an employee remains employed, exits voluntarily, or exits involuntarily. Lastly, we show that the joint investigation of linguistic alignment and cultural fit can improve our understanding of alignment as a form of communication accommodation, by testing the efficacy of lexically-based and

*This paper is under review and may change in the final version. Please contact the authors before citing.

Submitted to WWW 2017
ACM ISBN xxxx. . . .
DOI: xxx/xxx

semantically-based alignment estimates, and find predictive benefits to the semantically-based alignment estimates.

2. CULTURAL FIT AND ALIGNMENT

2.1 Cultural Fit

Cultural fit has both cognitive and behavioral components. Cognitive cultural fit refers to the extent to which a person shares assumptions, beliefs, and values with other members of a group, whereas behavioral culture fit represents the degree to which a person adheres to the normative expectations of the group [4, 26, 28]. These two forms of fit are interrelated—people draw inferences about their cognitive cultural fit based on observations of the alignment between their own and others' behavior. In other words, behavioral cultural fit provides a window into cognitive cultural fit. Yet the two forms of fit can also be decoupled—for example, when people strategically adjust their self-presentations to fit group expectations even when they do not share the group's beliefs and assumptions [15, 27]

With the exception of ethnographic research on group and organizational cultures (e.g. [19]), prior work has tended to emphasize cognitive cultural fit based on self-report measures that are inherently static and that implicitly assume a high degree of correspondence between the cognitive and behavioral dimensions of fitting in (e.g. [21]). Ethnographic research can reveal the behavioral manifestations of cultural fit but is difficult to scale to all members of large social groups or organizations as a whole and can be costly or simply infeasible to undertake for extended periods of time.

Recent studies have tried to address these inherent limitations of self-reports and ethnographic research by using language-based measures of cultural fit. Building on the insight that linguistic accommodation can serve as an indicator of the social distance between individuals [7], Srivastava et al. [28] developed an interactional language use measure of cultural fit that was based on the similarity or dissimilarity in linguistic style between individuals and their colleagues in an organization. This time-varying measure highlights an important facet of behavioral fit—the extent to which a person is linguistically compatible with her group—and reveals distinct trajectories of enculturation for employees who experience different career outcomes. Goldberg et al. [16] extend this approach to examine how the consequences of cultural fit for career success depend upon a person's position within the organization's internal network structure.

Although this recent work has shed new light on the dynamics and consequences of cultural fit, critical features of the enculturation process remain obscured given the details of the analyses used in previous studies. First, cultural fit in these studies has not incorporated directionality information. In particular, [16] defined fit in terms of the Jensen-Shannon divergence between the word distributions in incoming and outgoing email messages in a given month. Because this measure is symmetric, it provides an overall indicator of how a person's level of cultural fit changes over time, but it cannot reveal the direction of change. It does not, for example distinguish if a person's cultural fit has increased from one time period to the next because she has moved closer to her group, because her group has moved closer to her, or because of a combination of these shifts.

Second, previous research measured cultural fit as distributional changes in relative usage frequencies over a wide range of categories, rather than focusing on changes within the individual categories. In [16], messages were binned using the 64 categories in the standard Linguistic Inquiry and Word Count (LIWC) lexicon [24]. Fit was then computed over the LIWC distributions of incoming and outgoing messages, without considering how alignment with respect to particular linguistic categories—for example, pronouns—might or might not matter for successful enculturation into an organization.

2.2 Linguistic Alignment

Linguistic alignment is the tendency to use the same or similar words to one's conversational partner. Alignment is an instance of a pervasive human behavior: *communication accommodation* [13], the tendency of two interacting people to nonconsciously adopt similar behaviors. Evidence of accommodation appears in many different behavioral dimensions, including gestures, postures, speech rate, self-disclosure, and language/dialect choice (see [13] for a review).

Accommodation more generally is widespread and socially important. People who accommodate their interlocutors more are rated by their interlocutors as more intelligible, attractive, and cooperative [10, 18, 29]. These internal perceptions are accompanied by practical effects: High accommodation requests are more likely to be fulfilled, and pairs who accommodate more in how they express uncertainty perform better in lab-based tasks [2, 11].

Variability in accommodation and alignment.

While accommodation is widespread, individuals can also vary their levels of accommodation in informative ways for discourse-strategic purposes. More powerful people are accommodated more strongly in a variety of settings, including trials [14], online forums [6], and Twitter [9]. Speakers may also, within a conversation, increase their accommodation to signal a sense of group camaraderie, or decrease it to assert their senses of self. For example, Welsh English speakers increased their use of the Welsh accent and language in response to an English speaker who dismissed it [1]. Accommodation behavior also varies from person to person, and previous work has found that factors such as age, introversion, and desire for social approval relate to this variation [3, 20].

We focus on linguistic alignment in this paper for two reasons. First, linguistic alignment is well-established as a type of communication accommodation, and has been shown to correlate well with social outcomes in previous work (e.g., dating [18], joint decision making [11]). Second, in contrast to accent, prosody, or posture, linguistic alignment is one of the few aspects of communication accommodation that is retained in nearly all types of computer-mediated communication. While the reader of an email thread doesn't know the posture or the speech rate of the sender, they do see what words they chose to use. This feature of linguistic alignment allows us to investigate its effects in relative isolation, leading both to insights about alignment itself as well as controlling for outside factors when understanding its relationship to cultural fit.

Estimating alignment.

Linguistic alignment has been studied using a wide variety of quantitative measures [5, 9, 11, 18, 31]. Here we adopt the Word-Based Hierarchical Alignment Model (WHAM; [8]), to estimate the levels of linguistic alignment in the corpus. This measure has been shown to be robust to sparse, short messages such as those found in emails and microblog texts.

WHAM is a conditional measure of alignment, meaning that it separately estimates baseline usage (how often a speaker uses a word of their own accord) from alignment (the level of adaptation of word use to an interlocutor), distinguishing two sources of speaker similarity: homophily versus adaptation. This separation of baseline and alignment is an important part of understanding enculturation, similar to the cognitive versus behavioral components of cultural fit. WHAM also represents a directed measure of alignment, in that it estimates a *replier's* adaptation to the other conversational participant, independent of the participant's adaptation to the replier. As such, we can use WHAM alignment estimates to focus on how much an individual adapts to the prevailing culture.

2.3 Connecting Alignment and Cultural Fit

Why use alignment as a measure of cultural fit? Linguistic alignment reflects an individual's affinity to their interlocutor, and has been shown to correlate with better performance on group tasks. Coupled with previous work on linguistic similarity as an indicator of cultural fit [28], this pattern of findings suggests that linguistic alignment should reflect an individual's attachment to an organization's culture. Moreover, alignment differs from two common approaches for measuring cultural fit, participant observation [19, 30] and self-report surveys [21, 17], which lack fine time-resolution and may be affected by reporting biases. Using linguistic alignment as a measure of cultural fit also provides several practical advantages. First, the ubiquitous use of email provides an opportunity to produce finer-scale estimates than other methods. In addition, while some general aspects of tone in writing can be self-conscious, the particular behaviors being captured in linguistic alignment are thought to be non-conscious [23] and not directly related to the primary task of communication.

Furthermore, a joint investigation of linguistic alignment and cultural fit may be beneficial in refining our understanding of how alignment behaviors should be interpreted more generally. While linguistic alignment has proven itself an effective methodology in a variety of fields, there is still much to learn about it. What drives alignment, for example? Is it a primarily low-level cognitive behavior (e.g., hearing someone else say a word activates it in one's own mind, increasing its probability of re-use) or a higher-level discourse-strategic behavior motivated by goals like establishing rapport? Are the accommodative aspects of linguistic alignment primarily realized at the lexical level (i.e., repeating the same or similar words) or the semantic level (i.e., referring to the same objects, even if using different words for them)? We take advantage of the high resolution data with real-world organizational outcomes in Analysis 3 to inform our view of linguistic alignment in general, finding that lexical and semantic alignment are both predictive of employment outcomes, but semantic alignment is more predictive.

3. DATA: CORPORATE EMAIL CORPUS

Our data encompass the complete corpus of electronic messages, including metadata and content, exchanged among full-time employees at a mid-sized technology company between 2009 to 2014 [28]. All data were stored on secure research servers to protect employee privacy and company confidentiality, and all messages exchanged with anyone outside the company or with any of the company's attorneys were deleted. All identifying information about employees was removed, and each email was summarized into a count of the number of occurrences of 11 different word categories within it. These categories are a subset of the Linguistic Information and Word Count system [24], which is commonly used in linguistic alignment research. The categories were chosen because they are likely to be indicative of one's standing within a group.[1]

We divided email chains into message-reply pairs in order to investigate how messages shape their replies. To ensure that we were investigating pairs where the reply was likely to be related to the preceding message and its sender, we removed all emails with more than one sender or recipient (including CC/BCC), identical sender and recipient, or where the sender or recipient was an automatic notification system, mailing list, or any other mailbox that was not specific to a single employee. We also excluded emails with no body text or more than 500 words in the body text, and message-reply pairs with more than a week's latency between them. Finally, because we are investigating enculturation dynamics over the first six months of employment, we excluded replies sent by an employee who was at the company for less than six months. This resulted in a collection of 407,779 message-reply pairs with 485 replying employees. Crucially for our analyses, we also have access to which calendar month each employee joined and left the company, and whether they left voluntarily or involuntarily. Of the 485, 66 left voluntarily, 90 left involuntarily, and 329 remained employed at the end of the observation period.

4. MODEL: THE WHAM FRAMEWORK

To assess alignment, we use the Word-Based Hierarchical Alignment Model (WHAM) framework [8]. The core principle of WHAM is that alignment is realized as a change in the frequency of using a word in a reply when the word was used in the preceding message. Alignment is generally an increase in word usage, as people tend to re-use words and word categories that their interlocutors used. For instance, a reply to the message *What **will** we discuss at tomorrow's meeting?*, is likely to have more instances of future tense than a reply to the message *What **did** we discuss at yesterday's meeting?*.[2] Under this definition, alignment is defined relative to the *baseline* frequency, the frequency of the word in a reply when the preceding message did **not** contain the word.

Enculturation in this framework can be realized as a shift in alignment or baseline levels. For instance, a new employee

[1]Six pronoun categories (first singular (*I*), first plural (*we*), second (*you*), third singular personal (*he, she*), third singular impersonal (*it, this*), and third plural (*they*)) and five time/certainty categories (past tense, present tense, future tense, certainty, and tentativity).

[2]What accounts for this alignment effect? The psychological sources of alignment are still unknown but we believe there is substantial evidence for at least some discourse-strategic alignment effects [8]; we remain agnostic about whether there are other lower-level mechanisms at work as well, potentially including psycholinguistic priming [25] or topic-induced correlations in word usage.

may try to fit in by closely following the same linguistic patterns as the more established employees until they feel comfortable in the company's culture, in which case the employee's early trajectory may show a decrease in alignment and a shift in the baseline toward the company's average. Persisting in high alignment without a concomitant shift toward the company's baseline may reflect a failure to enculturate.

WHAM is a hierarchical generative modeling framework. By estimating alignment values over a hierarchy of different grouping variables, it is able to use information from related observations (e.g., multiple repliers with similar demographics) to improve its robustness on sparse observations [9]. A chain of normal distributions over a hierarchy generates two key parameters: η^{base}, the log-odds of a given word category c when the preceding message did not contain c, and η^{align}, the increase in the log-odds of c when the preceding message did contain c.

4.1 A dynamic extension

Our focus in this work is to track changes in alignment (and the baseline) over time, so we introduce a by-month change term to WHAM, and build a piecewise linear model of alignment and baseline frequency over the course of an employee's tenure. This modified version of WHAM is similar to the alignment model proposed by [32] to account for alignment changes between parents and children as the children age (though their model had only a single linear segment and ours has two). Each employee's tenure is broken into two or three segments: their first six months after being hired, their last six months before leaving (if they leave), and the rest of their tenure. The linear segments for their alignment are fit as an intercept term η^{align}, based at their first month (for the initial period) or their last month (for the final period), and per-month slopes α. Baseline segments are fit similarly, with parameters η^{base} and β.[3] In our present work, we are concerned primarily with changes in cultural fit as an employee transitions into or out of the group, so we treat all observations outside the first/last six months as a stable point estimate, constraining α and β to be zero. This simplification avoids the issue of different employees having different length middle periods.

To visualize the results, we create "sawhorse" plots, with an example in Figure 1. The model estimates three point parameters: the value in an employee's first month ($\eta_{align,start}$); the value in an employee's last month ($\eta_{align,end}$, if applicable); and the steady-state estimate in the rest of the employee's tenure ($\eta_{align,mid}$). It also estimates two slope parameters: the change in alignment per month for the first six and last six months of employment (α_{start}, α_{end}). The same five parameters are estimated for the baseline rate of category use. These parameters are not constrained to create a continuous function, as previous work has shown that cultural fit continues to change throughout an employee's tenure [28], so alignment after six months is not necessarily representative of average behavior at the company.

4.2 Qualitative intuitions

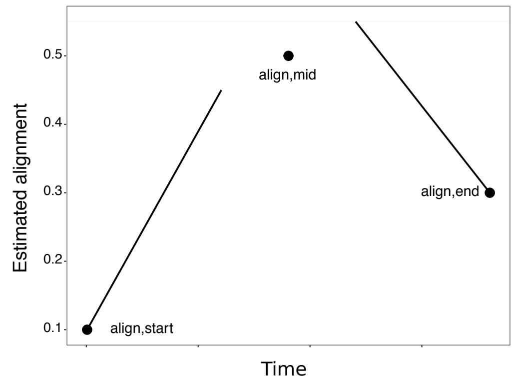

Figure 1: Sample sawhorse plot. The three point parameters (first month, last month, and middle average) and two by-month slope (entrance, exit) parameters are estimated by WHAM for each word category and group.

One important benefit of this extension to the WHAM framework is that it allows separate estimation of trajectories for both alignment behaviors and baseline rates of word usage. These baselines are estimated as the frequency of using a word category when the preceding message did not contain that category. The comparison of baseline and alignment measures gives us a sense of whether the employee is using the word more on their own, as might be expected if they are learning and adapting to the speech style of the company as a whole, or only changing how they use it in response to others' use.

Examined qualitatively, there are at least two interesting cases to be distinguished in terms of their patterns of baseline and alignment change. Imagine an employee who enculturates by first having an elevated alignment level, repeating the words she hears so as to appear to fit in, but then also independently adopting the way her colleagues speak, resulting in a change in baseline use. Such an employee could be thought of as a "believer" who adopts the corporate culture. Contrast this case with an employee who enculturates by adapting her use of a particular word or way of speaking to her interlocutor without independently changing her baseline use. This employee is more of a "chameleon", generating the appearance of cultural fit without internalizing it. The believer is distinguished by a baseline shift toward the average, regardless of alignment behavior; the chameleon by stable alignment or an upward shift, regardless of baseline behavior. We interpret alignment behavior as behavioral cultural fit, and baseline shift as indication of cognitive fit. These cases of course are not mutually exclusive. We assume that successfully assimilated individuals exhibit both believer and chameleon behaviors, reinforcing their cultural fit.

4.3 Model structure

The graphical model for our instantiation of WHAM is shown in Figure 2. For each word category c, WHAM's generative model represents each reply as a series of token-by-token independent draws from a binomial distribution. The binomial probability μ is dependent on whether the

[3]Informal investigations into the change in baseline usage over time showed roughly linear changes over the first/last six months, but we note that our linearity assumption may mask interesting variation in the shape of enculturation across different employees that may be worth further investigation.

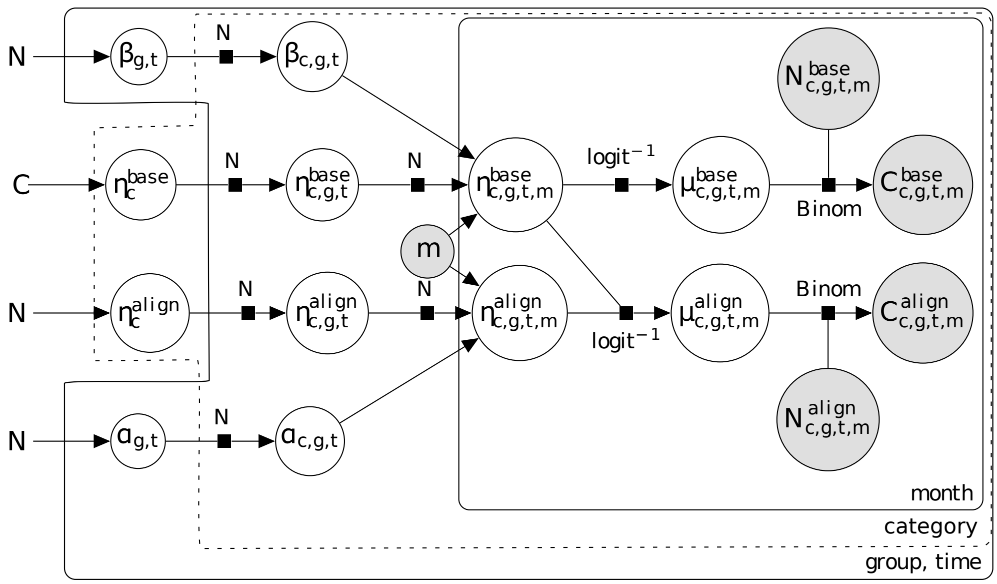

Figure 2: The Word-Based Hierarchical Alignment Model (WHAM). Hierarchical chains of normal distributions capture relationships between word categories, individuals, outcome groups, and time, and generate linear predictors η, which are converted into probabilities μ for binomial draws of the words in replies.

preceding message did (μ^{align}) or did not (μ^{base}) contain a word from category c, and the inferred alignment value is the difference between these probabilities in log-odds space (η^{align}).

The specific values of these variables depend on three hierarchical features: the word category c, the group g that a given employee falls into, and the time period t (a piece of the piece-wise linear function: beginning, middle, or end). Note that the hierarchical ordering is different for the η chains and the α/β chains; c is above g and t for the η chains, but below them for the α/β chains. This is because we expect the static (η) values for a given word category to be relatively consistent across different groups and at different times, but we expect the values to be independent across the different word categories. Conversely, we expect that the enculturation trajectories across word categories (α/β) will be similar, while the trajectories may vary substantially across different groups and different times. Lastly, the month m in which a reply is written (measured from the start of the time period t) has a linear effect on the η value, as described below.

To estimate alignment, we first divide the replies up by group, time period, and month. We separate the replies into two sets based on whether the preceding message contained the category c (the "alignment" set) or not (the "baseline" set). All replies within a set are then aggregated in a single bag-of-words representation, with category token counts $C^{align}_{c,g,t,m}$ and $C^{base}_{c,g,t,m}$, and total token counts $N^{base}_{c,g,t,m}$ and $N^{base}_{c,g,t,m}$ comprising the observed variables on the far right of the model. Moving from right to left, these counts are assumed to come from binomial draws with probability $\mu^{align}_{c,g,t,m}$ or $\mu^{base}_{c,g,t,m}$. The μ values are then in turn generated from η values in log-odds space by an inverse-logit transform, similar to linear predictors in logistic regression.

The η^{base} variables are representations of the baseline frequency of a marker in log-odds space, and μ^{base} is simply a conversion of η^{base} to probability space, the equivalent of an intercept term in a logistic regression. η^{align} is an additive value, with $\mu^{align} = \text{logit}^{-1}(\eta^{base} + \eta^{align})$, the equivalent of a binary feature coefficient in a logistic regression. The specific month's η variables are calculated as a linear function: $\eta^{align}_{c,g,t,m} = \eta^{align}_{c,g,t} + m\alpha_{c,g,t}$, and similarly with β for the baseline.

The remainder of the model is a hierarchy of normal distributions that integrate social structure into the analysis. In the present work, we have three levels in the hierarchy: category, group, and time period. In Analysis 1, employees are grouped by their employment outcome (stay, leave voluntarily, leave involuntarily); in Analyses 2 & 3, where we predict the employment outcomes, each group is a single employee. The normal distributions that connect these levels have identical standard deviations $\sigma^2 = .25$.[4] The hierarchies are headed by a normal distribution centered at 0, except for the η_{base} hierarchy, which has a $Cauchy(0, 2.5)$ distribution.[5]

Length has been shown to have a substantial effect on alignment estimates; the WHAM model was developed in part to reduce this effect. As different employees had substantially different reply length distributions, we further accounted for length by dividing all replies into five quintile length bins, and treated each bin as separate observations

[4]The deviation is not a theoretically motivated choice, and was chosen as a good empirical balance between reasonable parameter convergence (improved by smaller σ^2) and good model log-probability (improved by larger σ^2).

[5]As η^{base} is the log-odds of each word in a reply being a part of the category c, it is expected to be substantially negative. For example, second person pronouns (*you*), are around 2% of the words in replies, approximately -4 in log-odds space. We follow [12]'s recommendation of the Cauchy prior as appropriate for parameter estimation in logistic regression.

for each employee. This design choice adds an additional control factor, but results were qualitatively quite similar without it. All of our analyses are based parameter estimates based on RStan fits of WHAM with 500 iterations over four chains.

5. ANALYSIS 1: ALIGNMENT AND BASELINE RATES OVER TIME

We begin our analyses by examining the behavior of a few categories that are likely to indicate incorporation into the company's culture. Specifically, these are the use of first- and second-person pronouns (e.g., *I, we, you*), under the assumption that increases in *we* usage will occur as the employee is integrated into the group, while *I* and *you* usage will decrease. We also examine employees' use of expressions of certainty or tentativity, which may be triggered by the transitions around their start date and end date.

5.1 Experiment design

We divided employees' tenures by calendar month, then divided this data into an employee's first six months, their last six months (if an employee left the company by the last month of data), and the middle of their tenure. The emails from an employee's first and last six months were treated as twelve separate observations, and all from the rest of their tenure aggregated. Only employees with a twelve month or longer tenure were included, so that the first and last months did not overlap.

We fit four WHAM models in this experiment. Two aggregated all employees, regardless of employment outcome, and two separated them by outcome. For each of these structures, one model was fit using only pronouns and one using only expressions of certainty, tentativity, and verb tense, in case the two sets had sufficiently different alignment behaviors that combining them into the same hierarchy would be inappropriate.

5.2 Results

To facilitate interpretation, we start by examining the behavior of all employees (irrespective of outcome). Figure 3 focuses on five especially interpretable markers. We will only discuss the first six months in this initial examination, as the behaviors over the last six months should be quite different depending on why the employee leaves the company. For baselines on pronouns, we see that levels of *I* and *you* use do drop in the first six months, with *we* usage increasing over the same period, confirming the expected result that incorporating into the group is accompanied by more inclusive pronoun usage. As for alignment, all three pronouns are fairly stable through the first six months. Alignment on *I* and *you* is lower than *we* and most other pronouns, due at least in part to the fact that the referents of *I* and *you* are different between the two emailers, which is discussed in Analysis 3.

Certainty and tentativity markers both decrease in baseline use over the first six months; this change is more likely due to a shift in discussion topics than a shift in confidence, given that the two markers pattern together rather than moving in opposite directions. Alignment estimates are noisier, but it appears that alignment on certainty decreases while alignment on tentativity increases. Speculatively, this pattern could reflect increased comfort by the employees in expressing doubts as they enculturate.

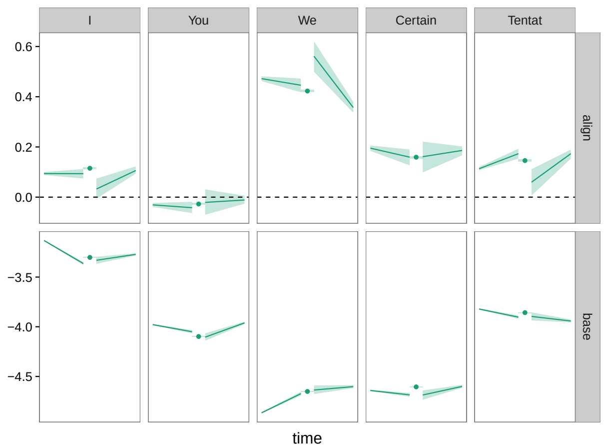

Figure 3: Sawhorse plots showing dynamics of alignment behavior aggregated across all employees. Vertical axis shows log-odds for baseline and alignment. Top row shows estimated alignment, positive for all categories but *you*-pronouns. Bottom row shows baseline dynamics, with employees shifting toward the average usage as they enculturate. The shaded region is one standard deviation over parameter samples.

We next turn to splitting by outcome. Figures 4 and 5 show outcome-specific trajectories, with green lines showing involuntary leavers (i.e., those who are fired or experience downsizing), blue showing voluntary leavers, and orange showing employees who remained at the company in the final month of the data. The use of *I* and *you* is similar to the aggregates in Figure 3, regardless of group. There is an interesting difference in the last six months of *I* usage, where involuntary leavers align more on *I* but retain a stable baseline while voluntary leavers retain a stable alignment but use *I* more overall, which is consistent with group separation.

Certainty and tentativity show large by-outcome differences. Involuntary leavers use certainty expressions less overall, throughout their tenure, while people who stay at the company have the highest levels of certainty use. Furthermore, while both leavers align more on certainty throughout their first six months, they follow different paths in their last six months, with involuntary leavers aligning more and voluntary leavers aligning less. This likely reflects the involuntary leavers' lack of control over their destiny, which the voluntary leavers retain. Stayers also consistently use more tentative expressions than leavers, another sign of comfort within the company. There is another large alignment crossover in tentativity during the last six months, with involuntary leavers greatly increasing their alignment on tentativity, while voluntary leavers remain stable.

The most compelling result we see here, though, is the changes in *we* usage by different groups of employees. Employees who eventually leave the company involuntarily show the behavior of "chameleons" over the first six months, increasing their alignment while decreasing their baseline use (though they return to more similar levels as other employees later in their tenure). Employees who stay at the company, as well as those who later leave voluntarily, show the

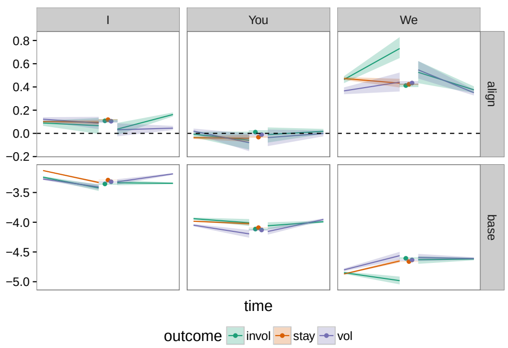

Figure 4: Sawhorse plots for pronouns, split by employment outcome. Mid-tenure points are jittered for improved readability.

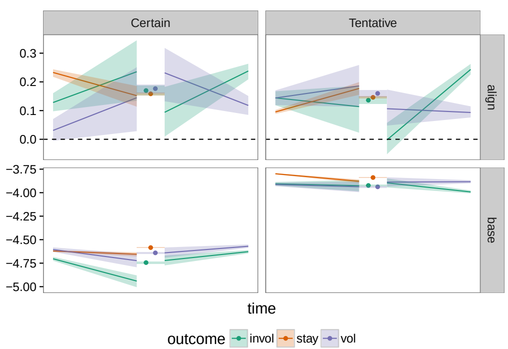

Figure 5: Sawhorse plots for certainty andtentativity, split by employment outcome.

true believer pattern, increasing their baseline usage to the company average, as well as adapting their alignment levels to the mean. This finding suggests that how quickly the employees acquire culturally-standard language use predicts their eventual employment outcome, even if they eventually end up near the average. We test this hypothesis in the second experiment by predicting outcomes from only the first six months of email.

6. ANALYSIS 2: PREDICTING OUTCOMES THROUGH ALIGNMENT

The goal of this second analysis was to test the hypothesis that there are meaningful differences in employees' alignment behaviors, especially during initial enculturation. To that end, we examine the first six months of communications and attempt to predict whether the employee will leave the company, a sign of poor cultural fit. We find that, even with a simple prediction model, alignment behaviors are in fact predictive of employment outcome.

6.1 Design

We fit the WHAM model to the first six months of email correspondence for all employees who had at least six months of email. The model estimated the initial level of baseline use ($\eta_{base,0}$) and alignment ($\eta_{align,0}$) for each employee, as well as the slope (α, β) for baseline and alignment over those

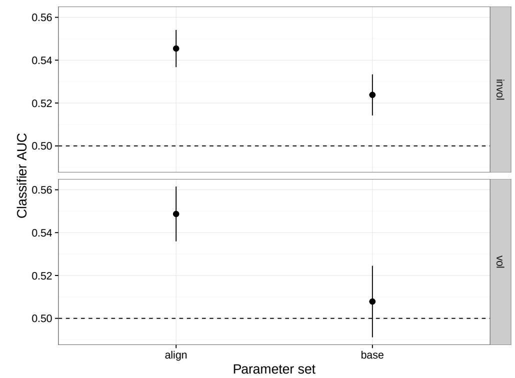

Figure 6: AUC (area under the curve) values for 10 runs of 10-fold cross-validated logistic classifiers. Predictability is consistently above chance for the alignment-based classifier; error bars indicate 95% confidence intervals.

first six months, over the 11 word categories. This procedure yielded 44 parameter estimates for each employee.

To create a predictive model of employment outcomes, we next created two logistic regression models, predicting whether an employee would leave the company. We fit separate models for leaving voluntarily or involuntarily, since these outcomes have substantially different motivations. In particular, as our results show, early alignment behaviors provide little information to identify employees who will leave voluntarily, consistent with [28]'s findings that voluntary leavers deviate from stayers primarily late in their tenure. We fit separate classifiers using the alignment parameters and the baseline parameters to investigate their relative informativity.

For each model, we report "area under the curve" (AUC). This value is estimated from the receiver operating characteristic (ROC) curve, which plots the true positive rate against the false positive rate over different classification thresholds. Changing thresholds shifts the classifier from favoring positive to negative predictions, so the AUC represents the overall performance as the preference for false positives or false negatives changes. An AUC of 0.5 represents chance performance, where any positive prediction is equally likely to be correct or incorrect. AUCs above 0.5 indicate that the classifier is effectively capturing predictive power. We use balanced, stratified cross-validation to reduce AUC misestimation due to unbalanced outcome frequencies and high noise [22].

6.2 Results

Figure 6 shows the mean and standard error over 10 runs of 10-fold balanced logistic classifiers with stratified cross-validation. The alignment-based classifiers (left) are both above chance at predicting that an employee will leave the company, whether involuntarily or voluntarily. The baseline-based classifiers are above chance for involuntary leavers, but are non-predictive for voluntary leavers. This finding is consistent with the idea that voluntary leavers resemble stayers (who form the bulk of the employees) until late in their tenure when their cultural fit declines.

We fit a model using both alignment and baseline parameters as well, but this model suffered from overfitting,

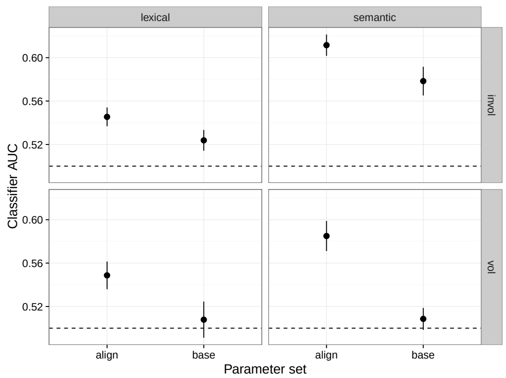

Figure 7: AUC values for 10 runs of 10-fold cross-validated logistic classifiers. Both lexical and semantic alignment parameters lead to above chance classifier performance, but semantic alignment outperforms lexical alignment at predicting both voluntary and involuntary departures.

yielding an AUC value between the alignment and baseline parameters. This finding suggests that where alignment and baseline behaviors are both predictive, they do not provide sufficiently different predictive power to overcome overfitting and the curse of dimensionality. It is likely that a more sophisticated classification model would help in overcoming these challenges; our goal here was not to achieve maximal classification performance but simply to create a minimal estimate of the information content in the alignment estimates.

7. ANALYSIS 3: TYPES OF ALIGNMENT

Our final analysis examines the assumption, made in previous work, that alignment should track word category use. There are strong psycholinguistic reasons to expect word repetition to correlate with amount of attention or affinity to an interlocutor (e.g., the Interactive Alignment Model [25]). But, for many of the words used in alignment analyses, the referent of the word may change depending on which member of the conversation is saying the word. For instance, linguistic alignment has generally treated the message-reply pair *I'm going/I am too* as aligning, even though *I* refers to two different people; by comparison, *I'm going/You are?* shows semantic alignment, since the same person is referred to in both sentences. This referential "semantic" alignment may even be more predictive of effective communication by tracking how well the people involved in a conversation match in what they are talking about. For this reason, we compare the predictability of employment outcome using the lexical alignment values of the previous analyses against the predictability using "semantic" alignment values, which attempt to determine whether the speakers are more likely to refer to the same objects.

To test this hypothesis, we make a small change to the alignment calculations. Lexical alignment as defined above is based on the conditional probability of the replier using a word category c given that the preceding message used that same category c. For semantic alignment, we examine the conditional probability of the replier using a word category c_j given that the preceding message used the category c_i, where c_i and c_j are likely to be referentially linked. We also consider cases where c_i is likely to transition to c_j throughout the course of the conversation, such as present tense verbs turning into past tense as the event being described recedes into the past. The pairs of categories that are likely to be referentially or evolutionarily linked are: (*you, I*); (*we, I*); (*you, we*); (past, present); (present, future); and (certainty, tentativity). We include both directions of these pairs, so this provides approximately the same number of predictor variables for both situations to maximize comparability (12 for the semantic alignments, 11 for the lexical). This modification does not change the structure of the WHAM model, but rather changes the compilation of the C and N counts by changing which replies go into the baseline or alignment pathways.

7.1 Results

Figure 7 plots the differences in predictive model performance using lexical versus semantic alignment parameters. We find that the semantic parameters provide more accurate classification than the lexical both for voluntarily and involuntarily-leaving employees. This suggests that while previous work looking at lexical alignment successfully captures social structure, semantic alignment may reflect a deeper and more accurate representation of the social structure. It is unclear if this behavior holds in less formal situations or with weaker organizational structure and shared goals, but these results suggest that semantic alignment should be investigated alongside lexical alignment in conversational studies.

8. CONCLUSIONS

This paper described an effort to use directed linguistic alignment—the tendency of speakers to be influenced by their interlocutors' word use—as a measure of cultural fit within an organization. We adapted a hierarchical alignment model from previous work to estimate alignment within corporate email communications, focusing on changes in language during employees' entry to and exit from the company. Our results showed substantial changes in the use of pronouns, with pronoun patterns varying by employees' outcomes within the company: those who stayed tended to use "we" more and more over their first six months, for example, while those who left involuntarily tended to repeat "we" when their conversation partner used it, but not to introduce it into conversation as often. Quantitatively, rates of usage and alignment in the first six months of employment carried information about whether employees left involuntarily, pointing towards early markers of employment outcomes in the fit of employees with the broader corporate culture. Future research should extend and further validate these observations, which are correlational in nature and hence subject to test via intervention. Finally, we saw ways in which the application of alignment to cultural fit might help to refine ideas about alignment itself: preliminary analysis suggested that semantic, rather than lexical, alignment was more predictive of employment outcomes. More broadly, these results suggest ways that quantitative methods can be used to make precise application of concepts like "cultural fit" at scale.

9. REFERENCES

[1] R. Y. Bourhis and H. Giles. The language of intergroup distinctiveness. In H. Giles, editor, *Language, Ethnicity, and Intergroup Relations*, pages 119–135. Academic Press, London, 1977.

[2] D. B. Buller and R. K. Aune. The effects of vocalics and nonverbal sensitivity on compliance: A speech accommodation theory explanation. *Human Communication Research*, 14:301–32, 1988.

[3] B. R. Burleson. Age, social-cognitive development, and the use of comforting strategies. *Communication Monographs*, 51:140–53, 1984.

[4] J. A. Chatman, J. T. Polzer, S. G. Barsade, and M. A. Neale. Being different yet feeling similar: The influence of demographic composition and organizational culture on work processes and outcomes. *Administrative Science Quarterly*, 43(4):749–780, 1998.

[5] C. Danescu-Niculescu-Mizil and L. Lee. Chameleons in imagined conversations: a new approach to understanding coordination of linguistic style in dialogs. In *Proceedings of the Workshop on Cognitive Modeling and Computational Linguistics*, pages 76–87. Association for Computational Linguistics, 2011.

[6] C. Danescu-Niculescu-Mizil, L. Lee, B. Pang, and J. Kleinberg. Echoes of power: Language effects and power differences in social interaction. In *Proceedings of the 21st international conference on World Wide Web*, page 699, 2012.

[7] C. Danescu-Niculescu-Mizil, R. West, D. Jurafsky, J. Leskovec, and C. Potts. No country for old members: user lifecycle and linguistic change in online communities. In *Proceedings of the 22nd International Conference on World Wide Web*, pages 307–318, 2013.

[8] G. Doyle and M. C. Frank. Investigating the sources of linguistic alignment in conversation. In *Proceedings of ACL*, 2016.

[9] G. Doyle, D. Yurovsky, and M. C. Frank. A robust framework for estimating linguistic alignment in Twitter conversations. In *Proceeedings of WWW*, 2016.

[10] R. E. Feldman. Response to compatriots and foriegners who seek assistance. *Journal of Personality and Social Psychology*, 10:202–14, 1968.

[11] R. Fusaroli, B. Bahrami, K. Olsen, A. Roepstorff, G. Rees, C. Frith, and K. Tylén. Coming to Terms: Quantifying the Benefits of Linguistic Coordination. *Psychological Science*, 23(8):931–939, 2012.

[12] A. Gelman, A. Jakulin, M. G. Pittau, and Y.-S. Su. A weakly informative default prior distribution for logistic and other regression models. *The Annals of Applied Statistics*, 2008.

[13] H. Giles, N. Coupland, and J. Coupland. Accommodation theory: Communication, context, and consequences. In H. Giles, J. Coupland, and N. Coupland, editors, *Contexts of accommodation: Developments in applied sociolinguistics*. Cambridge University Press, Cambridge, 1991.

[14] A. Gnisci. Sequential strategies of accommodation: A new method in courtroom. *British Journal of Social Psychology*, 44(4):621–643, 2005.

[15] E. Goffman. *The presentation of self in everyday life*. Anchor Books, New York, 1959.

[16] A. Goldberg, S. B. Srivastava, V. G. Manian, and C. Potts. Fitting in or standing out? The tradeoffs of structural and cultural embeddedness. *American Sociological Review*, forthcoming.

[17] G. Hofstede, G. J. Hofstede, and M. Minkov. *Cultures and Organizations: Software of the Mind*. McGraw-Hill, New York, 3rd edition, 2010.

[18] M. E. Ireland, R. B. Slatcher, P. W. Eastwick, L. E. Scissors, E. J. Finkel, and J. W. Pennebaker. Language style matching predicts relationship initiation and stability. *Psychological Science*, 22:39–44, 2011.

[19] G. Kunda. *Engineering culture: Control and commitment in a high-tech corporation*. Temple University Press, Philadelphia PA, 2006.

[20] M. Natale. Convergence of mean vocal intensity in dyadic communication as a function of social desirability. *Journal of Personality and Social Psychology*, 32(5):790–804, 1975.

[21] C. A. O'Reilly, J. Chatman, and D. F. Caldwell. People and organizational culture: a profile comparison approach to assessing person-organization fit. *Academy of Management Journal*, 34(3):487–516, 1991.

[22] B. J. Parker, S. Günter, and J. Bedo. Stratification bias in low signal microarray studies. *BMC bioinformatics*, 8(1):1, 2007.

[23] J. Pennebaker and L. King. Linguistic styles: language use as an individual difference. *Journal of Personality and Social Psychology*, 77(6):1296–1312, 1999.

[24] J. W. Pennebaker, C. K. Chung, M. Ireland, A. Gonzalez, and R. J. Booth. The development and psychometric properties of liwc2007. Technical report, LIWC.net, 2007.

[25] M. J. Pickering and S. Garrod. Toward a mechanistic psychology of dialogue. *Behavioral and brain sciences*, 27(2):169–190, 2004.

[26] E. H. Schein. *Organizational Culture and Leadership*. Jossey-Bass, San Francisco, 4 edition, 2010.

[27] M. Snyder. Self-monitoring of expressive behavior. *Journal of personality and social psychology*, 30(4):526, 1974.

[28] S. B. Srivastava, A. Goldberg, V. G. Manian, and C. Potts. Enculturation trajectories: Language, cultural adaptation, and individual outcomes in organizations. *Management Science*, forthcoming.

[29] H. C. Triandis. Cognitive similarity and communication in a dyad. *Human Relations*, 13:175–183, 1960.

[30] J. Van Maanen. The smile factory. In P. J. Frost, L. F. Moore, M. R. Louis, C. C. Lundberg, and J. Martin, editors, *Reframing Organizational Culture*, pages 58–76. SAGE Publications, Inc, 1991.

[31] Y. Xu and D. Reitter. An Evaluation and Comparison of Linguistic Alignment Measures. In *Proceedings of the Workshop on Cognitive Modeling and Computational Linguistics*, 2015.

[32] D. Yurovsky, G. Doyle, and M. C. Frank. Linguistic input is tuned to children's developmental level. In *Proceedings of the 38th Annual Meeting of the Cognitive Science Society*, 2016.