Investigating the Components of Linguistic Accommodation

Anonymous ACL submission

Abstract

In conversation, people tend to "accommodate" or "align" to their partners, adopting similar styles of speech. This accommodation is robust, appearing across many features, and its degree correlates with important social factors such as power and likability, even for non-human accommodators. We focus on "linguistic alignment", the amount that one's word use is influenced by others. While there are a variety of computational measures for assessing this alignment, they have not delved into the different components of linguistic alignment. We propose an extension to one of these measures to account for alignment in messages of various lengths, and to separate out the influence of words and word categories within alignment behaviors. We compare alignment behaviors on Twitter with those of the telephone conversations in Switchboard, and show substantial differences between the datasets. Furthermore, we find that different discourse acts also have different alignment behavior in Switchboard.

1 Introduction

In conversation, people tend to adapt to each other across a broad range of behaviors, collectively known as "communication accommodation" (Giles et al., 1991). Linguistic alignment, using similar words to one's interlocutor, is one prominent form of accommodation. It is a robust form of accommodation, found in many settings, including computer-mediated and web-based conversation (Giles, Pennebaker, DNM). It is also interesting because the strength of this alignment varies with important sociological factors, such as power, network centrality, and likability, and may be a useful way of inferring these factors in situations where they are unobserved.

While substantial work has suggested that linguistic alignment is important, there are a number of open questions about the nature of this alignment. One is the issue of different components of linguistic accommodation. While linguistic alignment has been studied in multiple ways, the most prominent strand has focused on the level of word categories, looking at how interlocutors change their frequency of using, for instance, pronouns or quantitative words. But alignment can occur at many levels, including individual words (**?**), of syntactic structure (**?**), and message length (**?**). There are conflicting expectations about how different types of linguistic alignment should interact, ranging from the Interactive Alignment Model's expectation that different types of alignment should be approximately uniform (Pickering and Garrod, 2004) to arguments that, for instance, lexical and syntactic alignment may push in different directions (Healey et al). We propose a method that uses the same framework to look at three different levels of alignment: word, category, and category-not-word alignments, defined in Section **??**.

A second major question is how the desire to align, which repeats preceding words and concepts, is balanced against a need to move the conversation forward by introducing new words and concepts. We propose that linguistic alignment is modulated by the discourse act being performed, with certain acts showing greater or lesser alignment depending on their purpose within the conversation. In particular, we examine the influence of back-channels, short utterances communicating agreement, understanding, or misunderstanding on alignment.

To address these questions, we propose a new version of an existing measure of linguistic alignment, increasing its robustness to messages of different lengths. We also extend the measure to consider multiple levels of alignment. We perform two experiments. First, we look at the differences in estimated alignment at the different levels, and show that alignment varies substantially by representation level. Second, we estimate alignment on different subsets of the Switchboard corpus of telephone conversations, finding that alignment is highly dependent what discourse actions are being undertaken.

2 Previous Work

2.1 Why does alignment matter?

Linguistic alignment, like other kinds of accommodation, can be a critical part of achieving social goals. Performance in a variety of cooperative decision-making tasks has been positively related to the participants' linguistic convergence (Fusaroli et al., 2012; Kacewicz et al., 2013). Match-matching in speed dating as well as stability in established relationships have been linked to increased alignment (Ireland et al., 2011). Alignment can also improve perceived persuasiveness, encouraging listeners to follow good health practices (Kline and Ceropski, 1984) or to leave larger tips (van Baaren et al., 2003).

Alignment is also important as an indicator of implicit sociological variables. Less powerful conversants generally accommodate more to powerful conversants. Prominent examples include interviews and jury trials (Willemyns et al., 1997; Gnisci, 2005; Danescu-Niculescu-Mizil et al., 2012). A similar effect is found for network structure: speakers align more to more network-central speakers (Noble and Fernández, 2015). Additionally, factors such as gender, likability, respect, and attraction all interact with the magnitude of accommodation (Bilous and Krauss, 1988; Natale, 1975). Such differences in accommodation can also be indicative of changes to the power dynamic: In U.S. Supreme Court transcripts, (Guo et al., 2015) showed that depending on the accommodation dimension, justices – who are more powerful by any intuitive assessment – may nevertheless accommodate more to lawyers, perhaps because the lawyers have the local power to answer justices' questions.

2.2 What's the nature of alignment?

The preceding results show that linguistic alignment is a robust phenomenon, but the specific nature of this alignment is something of a black box. One major question is how different structural levels of language align. The Interactive Alignment Model (Pickering and Garrod, 2004) makes the theoretical claim that alignment should be similar across all structural levels as conversants increasingly share their representation. However, (**?**) find in phone conversations that speakers syntactically diverge from their interlocutors once lexical alignment is accounted for. Because most work on alignment has been done either on categories of words or aggregating across the lexicon, we do not have a good sense of whether there are systematic differences between alignment at different levels.

Futhermore, positive alignment is treated as an inherently good thing, but there is clearly a limit to its goodness, as alignment is inherently backward-looking, while the general goal of a conversation is to exchange information that is not already known by both parties, and inherently forward-looking goal. There are suggestive results that alignment based on function words, which can stay constant even as the topic changes, is more appropriate than alignment based on all words for predicting performance in a decision-making task (Fusaroli et al., 2012), and some recent work on accommodation has limited itself to "non-topical" word categories (Danescu-Niculescu-Mizil et al., 2011; Doyle et al., 2016).

3 Measures of alignment

Multiple metrics for linguistic alignment have been used in previous work, which fall into two basic categories: distributional and conditional. The distributional methods, such as Linguistic Style Matching (LSM) (Niederhoffer and Pennebaker, 2002; Ireland et al., 2011) and the Zelig Quotient (Jones et al., 2014), calculate the similarity between the rate of usage of words or word categories between the conversation participants. Conditional metrics, such as Local Linguistic Alignment (LLA) (Fusaroli et al., 2012; Wang et al., 2014) and the metric used by (Danescu-Niculescu-Mizil et al., 2011), look at how a message conditions its reply, with convergence indicated by elevated word use in the reply when that word was in the preceding message. This paper extends a recent conditional metric, the Hierarchical Alignment Model (Doyle et al., 2016).

Distributional methods [overview of LSM]

By-message conditional methods By-message conditional methods rest on a binary view of utterances. If HAM is being used to assess alignment on pronouns, for instance, Bob aligns to Alice in the HAM model if his replies are more likely to contain a pronoun when in response to a message from Alice that contains a pronoun. An example of positive HAM alignment is shown in the table below:

	Bob's reply	
Alice's message	has pronoun	no pronoun
has pronoun	8	2
no pronoun	5	5

Here, Alice sends 10 messages that contain at least one pronoun, and 8 of Bob's replies contain at least one pronoun. But Alice also sends 10 messages that don't contain any pronouns, and only 5 of Bob's replies to these contain pronouns. Alignment is this increased likelihood of a pronoun-containing reply to a pronoun-containing message.

Different models quantify this alignment slightly differently. (Danescu-Niculescu-Mizil et al., 2011) propose a subtractive conditional probability model, where alignment is the difference between the likelihood of a pronoun-containing reply to a pronoun-containing message and the probability of a pronoun-containing reply to any message:

$$align_{SCP} = p(B|A) - p(B) \tag{1}$$

(Doyle et al., 2016) show that this measure can be affected by the overall frequency of the category being aligned on. They propose the Hierarchical Alignment Model (HAM), which quantifies this alignment as a linear effect on the log-odds of a reply containing the relevant marker (e.g., a pronoun), similar to a linear predictor in a logistic regression:[1]

$$align_{HAM} \approx logit^{-1}(p(B|A)) - logit^{-1}(p(B|\neg A)) \tag{2}$$

[1] Because the HAM defines a Bayesian inference structure for alignment, the inferred alignment value depends on the prior and number of message observed, so unlike the other measures, this equality is only approximate.

Both of these binarized conditional methods, though, depend on the assumption that all messages have similar, and small, numbers of words. The probability that a message contains at least one of any marker of interest is of course dependent on the message's length, so if messages vary substantially in their length, these alignment values can be at least noisy, if not biased. They are also not robust as messages increase in length, since the likelihood that a message contains any marker approaches 1 as message length increases.

By-word conditional methods A solution to this is simply to shift from binarized data to count data. Instead of modeling whether or not a reply contains a marker of interest, we can model that marker's frequency within replies to messages that do or do not contain the marker. Local Linguistic Alignment (LLA) (Fusaroli et al., 2012; Wang et al., 2014) and the lexical similarity (LS) measure of (**?**), are two examples that estimate the proportion of the preceding message that appear in the reply. LLA is shown below, but LS is similar:

$$align_{LLA} = \frac{\sum_{w_i \in M_b} \delta(w_i \in M_a)}{length(M_a) length(M_b)} \tag{3}$$

In LLA, alignment is the number of word tokens (w_i) that appear in both the message (M_a) and the reply (M_b) are divided by the product of the total number of word tokens in the message and reply.

These measures have an aspect of conditionality, as they only count words that appear in both the message and the reply, but they lack baselining, one of (Doyle et al., 2016)'s desired features in an alignment measure. They also have inherent length dependence, as the maximum alignment estimate is only possible when the reply is shorter than the message.[2]

We need, therefore, an alignment measure that has the benefits of the existing by-message conditional models (HAM outperformed SCP, LLA, and LSM in simulations in (Doyle et al., 2016)) while gaining the length-robustness of a by-word conditional method. A simple change to the HAM framework satisfies this goal.

4 The Word-Based Hierarchical Alignment Model (WHAM)

The Hierarchical Alignment Model (HAM) was proposed by (Doyle et al., 2016) as a way to

[2] Proof in Supplemental?

increase the robustness of conditional alignment metrics to sparse data and rare words. HAM conceptualizes alignment as the increase in the likelihood that a reply will contain a "marker"—a word or word category—given that the preceding message contained it compared to when the preceding message did not contain it, as discussed above.

The HAM framework makes a substantial simplification: it treats each message as a binary variable, either containing or not containing the word of interest. This approach follows that of (Danescu-Niculescu-Mizil et al., 2011), who found alignment affects despite the simplification, and may be appropriate for data such as the Twitter dataset used in that paper, which has a 140-character limit and thus ensures all messages will be short, rarely containing more than one instance of a word. However, in looking at more natural dialogues, such as phone conversations, length may vary substantially between messages. As such, we need a framework that is able to handle different message lengths.

We propose the Word-Based Hierarchical Alignment Model (WHAM) to address this. Like HAM, this framework assumes that replies are shaped by whether or not they are in reply to a message that contained the marker of interest. But where the HAM framework only looks at whether or not the reply contains the marker, the WHAM framework looks at the marker token frequencies within the reply. For each marker, each reply is treated as a sample of independent draws from a binomial distribution. The binomial probability is dependent on whether the reply is responding to a message that does or does not contain the marker.

[Graphical model and a re-vamp of the model outline paragraph from WWW to come.]

Generative Model
Alice says x, Bob wants to reply.
If $m \notin x$, then $p(B_m \vert A_m) = \mu^{base}(m)$
If $m \in x$, then $p(B_m \vert \neg A_m) = \mu^{align}(m)$
$\mu^{base}(m) = logit^{-1}(\eta^{base}(m))$
$\mu^{align}(m) = logit^{-1}(\eta^{base}(m) + \eta^{align}(m))$

5 Experiment 1

5.1 Levels of alignment

We want to examine the interaction of two levels of alignment. The first is lexical-level alignment, looking at how the presence of a word in a message increases its likelihood of appearing in the reply. The second is category-level alignment, looking at how the presence of a word in a message increases the likelihood of it or any other member of that category appearing in the reply.

Message	$\emptyset$	*he*	*she*
$\emptyset$	25	25	25
he	20	50	10
she	20	10	50

Table 1: A case where lexical alignment surpasses categorical alignment due to non-independence between the words.

These two levels interact in a non-obvious way, as shown in Table 1. Intuitively, category alignment should be higher than lexical alignment, since each instance of lexical agreement is also an instance of category agreement. However, interactions between lexical items within a category can cause the lexical alignment to overshoot the category alignment. Table 1 illustrates this with a theoretical distribution over the pronouns *he* and *she*; one use of the pronoun *he* makes another use more likely (*Did he like the movie? Yeah, he loved it.*) while also reducing the likelihood of *he*, since the topic of conversation is now a female, and vice versa for *she*. For both *he* and *she*, the lexical alignment is approximately $logit^{-1}(p(B|A) - p(B|\neg A)) = logit^{-1}(\frac{50}{80} - \frac{25}{75}) \approx 1.2$, but categorical alignment is approximately $logit^{-1}(\frac{120}{160} - \frac{50}{75}) \approx 0.4$.

Lexical alignment can differ from categorical alignment for a variety of reasons, so we consider three quantities: the lexical and categorical alignments, but also the "category-not-word" (CNW) alignment: the increased likelihood of seeing another member of the category other than the target word. Separating these three factors is complex, and the CNW alignment is not simply the categorical alignment minus the lexical alignment.

[Define CNW here]

References

FR Bilous and RM Krauss. 1988. Dominance and accommodation in the conversational behaviours of same-and mixed-gender dyads. *Language & Communication.*

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

World Wide Web - WWW '11, page 745, New York, New York, USA. ACM Press.

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 - WWW '12*, page 699.

Gabriel Doyle, Dan Yurovsky, and Michael C. Frank. 2016. A robust framework for estimating linguistic alignment in Twitter conversations. In *WWW 2016*.

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

H Giles, N Coupland, and J Coupland. 1991. 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.

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

Fangjian Guo, Charles Blundell, Hanna Wallach, Katherine Heller, and U C L Gatsby Unit. 2015. The Bayesian Echo Chamber: Modeling Social Influence via Linguistic Accommodation. In *Proceedings of the Eighteenth International Conference on Artificial Intelligence and Statistics*, pages 315–323.

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

S. Jones, R. Cotterill, N. Dewdney, K. Muir, and A. Joinson. 2014. Finding Zelig in text: A measure for normalising linguistic accommodation. In *Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics*, pages 455–465.

Ewa Kacewicz, James W Pennebaker, Matthew Davis, Moongee Jeon, and C Arthur. 2013. Pronoun use reflects standings in social hierarchies. *Journal of Language and Social Psychology*, 33(2):125–143.

SL Kline and JM Ceropski. 1984. Person-centered communication in medical practice. In J T Wood and G M Phillips, editors, *Human Decision-Making*, pages 120–141. SIU Press, Carbondale.

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

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

B Noble and R Fernández. 2015. Centre Stage: How Social Network Position Shapes Linguistic Coordination. In *Proceedings of the Workshop on Cognitive Modeling and Computational Linguistics*.

Martin J Pickering and Simon Garrod. 2004. Toward a mechanistic psychology of dialogue. *Behavioral and brain sciences*, 27(2):169–190.

R B van Baaren, R W Holland, B Steenaert, and Ad van Knippenberg. 2003. Mimicry for money: Behavioral consequences of imitation. *Journal of Experimental Social Psychology*, 39(4):393–398.

Yafei Wang, David Reitter, and John Yen. 2014. Linguistic Adaptation in Conversation Threads: Analyzing Alignment in Online Health Communities. In *Proceedings of 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.

A Supplemental Material

(Temporary collection of material that I'm considering adding as supplemental material – separate from the paper, doesn't count against 8-page limit. [references also don't count against limit])

Message	{∅}	{X}	{Y}	{X,Y}
{∅}	160	20	10	10
{X}	140	30	20	10
{Y}	170	10	15	5

Table 2: Another case where (mean lexical) ¿ categorical alignment.

[Proof that LLA and Healey's measure of lexical similarity can depend on message/reply length.]