How Emotional Support and Informational Support Relate to Linguistic Alignment

Yafei Wang, David Reitter, and John Yen

College of Information Sciences and Technology,
The Pennsylvania State University,
University Park, PA 16802
`yxw184@psu.edu, reitter@psu.edu, jyen@ist.psu.edu,`

Abstract. Linguistic alignment in text-based communication means that people tend to adjust their language use to one another both in terms of word choice and sentence structure. Previous studies about linguistic alignment suggested that these two forms of adaptation are correlated with each other, and that they build up to alignment at a higher representational level, such as pragmatic alignment for support functions. Two types of social support have been identified as important for online health communities (OHCs): emotional and informational support between support seekers and support providers. Do the two lower-level alignment measures (lexical and syntactic) relate to these two types of social support in the same way or, are they different? Our hypothesis was that they are similar, due to their correlation relationship. However, we found that, based on an analysis of a 10-year online forum for cancer survivors, the lower-level alignment measures have distinct relationships to the two higher-level support functions. In this paper, we describe this finding and its implications regarding potential refinement of the Interactive Alignment Model.

1 Introduction

For people living with a chronically illness like cancer, online "peer-to-peer" healthcare, such as exchanging information and social support, is an essential pathway for patients getting social support, helping each other, and increasing the quality of life [12]. A study from Pew Research Center [13] shows that 34% internet users have read others' health-related experience online, and 18% internet users have tried to find other patients with similar experiences online. A primary way of "peer-to-peer" healthcare interactions is using social network websites, such as Facebook groups, online health communities, etc. Online health communities (OHCs) serve prominent function of exchanging social support.

Generally, there are four types of social support: emotional support, informational support, tangible support and appraisal support [21]. However, the social support people seeking and providing in web-based conversation is either informational or emotional support [31]. This is because patients with chronic disease can benefit from not only information provided by their peers, such as

side effects of chemo therapy, but also emotional support, like understanding and caring. Moreover, anecdotal information and first-hand accounts from peers can be useful to patients in terms of reassurance, when patients do not weigh it against the evidence-based, scientific medical information obtained from professionals. Previous studies showed that receiving appropriate social support in thread-based conversations is correlated with members' commitment of online health support groups [31]. Similarly, [28] reported that members have higher satisfaction of online dialogues when they got the social support they need.

Another line of work in this paper is communication accommodation. Communication accommodation [15] is a communication phenomenon showing that people tend to adapt their gestures [4], speech [6], and language use [16], to accommodate other participants while interaction. Among these accommodation phenomena, linguistic alignment [24] is one of the useful models of accommodation, analyzing accommodation on language usage, with respect to words [16], syntax [6], and more, although general alignment may extend beyond language use. In *web-based communication*, linguistic alignment is the primary form of accommodation which is found under various settings [9, 22] as well as in OHCs [29].

Although linguistic alignment is well-studied, the reason of linguistic alignment is still arguable. This phenomenon may be subject to more or less explicit, conscious control, while many studies suggest an automated, implicit process [20, 26]. Lexical alignment, i.e., adaptation to a conversation partner's choice of words, is arguably more under a speaker's conscious control. [17] suggested that syntactic alignment is due to lexical repetition. Thus, under that finding, we expect consistency, and very little difference in correlation of syntactic and lexical priming with secondary variables.

We choose the higher-level representation of a conversation, such as the outcome of a given conversation as a secondary variable in this paper. Theoretically, Interactive Alignment Model [24] suggests that linguistic alignment at lexical and syntactic levels build up to alignment at a higher representation level, such as pragmatic alignment [30] and mutual understanding, at the conversational level. Therefore, understanding how linguistic alignment at different levels interacts with social support is important. Also, since emotional support and informational support are two important types of support in dialogues from OHCs, it is essential to understand and predict linguistic phenomena such as alignment.

Hence, the research question of this paper is whether different types of discussions influence the linguistic alignment behavior of the support provider? To be specific, are there differences regarding the relationship between lexical alignment and syntactic alignment with regard to emotional support and informational support threaded discussions? Our goal of research is to uncover the relationship between lexical alignment and syntactic alignment in two types of posts in OHCs: emotional support post and informational support post.

This paper is organized as follows. We firstly review previous studies on linguistic alignment and its mechanism. Then, we introduce social support measurements and linguistic measurements including lexical alignment, syntactic

alignment, and the corpus we used in this paper. Further, we conduct an experiment analyzing how linguistic alignment correlates to social support types at post level. We end by comparing the experiment results to analyze different correlations between social support and linguistic alignment measurements.

The contributions of this paper include: (1) providing a deeper understanding about the different relationship between lexical and syntactic alignment and emotional and informational support based dialogues in online health communities, (2) motivating further studies to investigate the explanations or casual factors of these different relationships, and (3) making theoretical progress on psycholinguistic models regarding linguistic alignment.

2   Related Work

While linguistic alignment is a well-studied linguistic phenomenon, its mechanism is still debatable among researchers. Previous studies [6, 27] considered that linguistic alignment happened at syntactic level is due to the internal mechanism of conversation, which happens unconsciously. In other words, people do not consciously adapt the sentence structure they used during conversation. While, many studies argued that alignment is due to the purpose of the dialogues, such as building common ground [8], a higher level strategy [10], or an artificial task that requires such mutual understanding [27].

Also, linguistic alignment is a useful predictor driving the higher level representation, such as the outcome of a given conversation. For example, linguistic convergence helps conversation participants win decision-making games [14], and even the outcome of speed dating [18].

Furthermore, linguistic alignment is a meaningful signal of revealing sociological relationship among conversation participants. [9], and [11] used linguistic alignment measures to identify power and power dynamics on a conversational dataset from Twitter. Generally, conversational participants with a lower power tend to align with participants with higher power. This phenomenon is also identified in Wikipedia editors' discussion groups [23] and general online communities [19].

3   Automated Classification of Social Support Types

We discuss two main types of social support in OHCs, emotional support and informational support [31]. Specifically, *Emotional support* is a type of social support that an individual provides a support seeker the provision of "understanding, empathy, encouragement and concern" [2]. In contrast, *informational support* provides support seeker "facts, advice, referral, teaching, personal experiences, and information" [2].

Table 1 shows an anonymized threaded interaction between support seeker and support provider. The message from support provider, i.e. the message in the second row, includes both emotional support, such as encouragement, and informational support, like personal experience. In OHCs, patients often responds

Table 1: An anonymized threaded interaction between support seeker and support provider. The sentences in gray and black in the second row are providing emotional and informational support, respectively. The support provider tends to use similar words and syntax as the support seeker.

A:	"Newly diagnosed at age XX, went to BS and they want to remove the cells., ... Everything I've read says it is very slow growing and surgery is not always necessary. She said I'm stage 1.,I'm petrified of any surgery. ... My mind is overwhelmed and just won't stop!,I know I'm lucky it was caught so early and I'm lucky it's only stage 1 but ....., Thank you in advance."
B:	"You are lucky to have caught this early. You are probably in shock right now so take a few days to think. I'm with your family. Better to get a few cells removed that wait until it becomes a tumor! Breast surgery is so routine these days. My sister got similar surgery and felt fine the same day. Good luck!"

to support requests in the forum by disclosing his/her personal experiences together with showing understanding and giving encouragement. These two types of social support are often intertwined.

Therefore, we distinguish these two types of social support by leveraging an automated classifier at the sentence level developed by [5]. As reported in [5], the sentence classifier was built on 1,066 hand-tagged sentences selected from Cancer Survivor's Network, which is the same corpus in this paper; and the initial agreement between two taggers was 89%. The features of building the sentence classifier include combined words, part-of-speech, subjective words, cancer-related words, linguistic patterns about emotional support and informational support and etc [5]. The macro-averaged precision, recall and F-1 score of that model reached 0.841, 0.842, and 0.840, respectively. Based on the result of this sentence-level classification, we further determine the support type of a post – if the number of emotional support sentences is larger than that of informational support sentences, the post is an *emotional support post*; otherwise, it is considered an *informational support post*. Following [5], the amount of one type of support, emotional support and informational support, in a reply post is quantified as: $Index_{Type} = SentNum_{Type}/SentNum_{Classified}$.

4 Linguistic Alignment Measures

There are multiple metrics, such as indiscriminate local linguistic alignment (LLA) [14], subtractive conditional probability (SCP) [9], linguistic style matching (LSM) [22], word-based hierarchical alignment model [10], and so on, quantifying linguistic alignment phenomenon. We choose LLA, which was used to measure linguistic alignment in online communities [29, 32], measures linguistic alignment among messages which computes lexical and syntactic alignment in a similar way.

Indiscriminate Local Linguistic Alignment, implemented in [29], measures the linguistic alignment at lexical and syntactic levels. Generally, LLA at the lexical

level measures the normalized word repetition in both prime and target posts in the same conversation (in the same thread). For example, it is computed as the number of words which occur in both the target post (in the second row of Table 1) and the prime post (in the first row of Table 1) normalized by the numbers of words in prime and target posts. Formally, *Lexical Indiscriminate Local Linguistic Alignment (LILLA)* is calculated as:

$$LILLA(\text{target}, \text{prime}) = \frac{\sum_{word_i \in \text{target}} 1_{\text{prime}}(word_i)}{length(\text{prime}) \times length(\text{target})} \qquad (1)$$

where length($post$) is the number of words in the post *post*, $1_{prime}(word_i)$ is an indicator function that the outcome is 1 if $word_i$ is in the prime message.

Similarly, *Syntactic Indiscriminate Local Linguistic Alignment (SILLA)* measures the percentage of syntactic rule repetition which appears in both prime and target posts in the same conversation. For each sentence in a post, we annotated it as a collection of syntactic rules using phrase structure trees generated by Standford CoreNLP Parser [1]. Then, we compute normalized syntactic rule repetition for SILLA.

5 Corpus

We use a collection of online threads from Cancer Survivor's Network (CSN) (`csn.cancer.org`), with more than 166,000 registered users and 41 sub-communities [25]. In CSN, cancer patients and cancer survivors with the same disease and under similar situations are in the same sub-communities. Members in the same sub-community often exchange their personal feelings and experiences of being under similar difficult personal circumstances and the associated emotional burden. Thus, most conversations happened in CSN are support-oriented conversations, which are either seeking and offering emotional support, informational support, or a mixture of social support. We use all the threads taken place from the two largest sub-forums in CSN, Breast cancer and Colorectal cancer, between June 2000 to October 2010, as our two corpora.

Table 2: The number of Emotional and Informational Support Posts in two sub-communities

	Breast	Colorectal
Emotional Support Posts	111,495	93,355
Informational Support Posts	20,610	16,099

A conversational *thread* is an initial post (normally seeking social support) followed by a sequence of replies (normally offering emotional or informational

---

[1] `http://stanfordnlp.github.io/CoreNLP/`

support) in temporal order. Formally, a conversational thread is shown as $<P_0, P_1, \cdots, P_i, \cdots, P_n>$, where P_0 denotes *initial post*, and the author of initial post is called *thread initiator*. The post *distance* indicates how much information has been discussed between two posts. Thus, given a post pair, P_i and P_j, the *distance* is calculated as $j - i$. Following [5], each sentence in a reply post is classified as providing either emotional or informational support. In this paper, we only consider either emotional support post or informational support post. The distribution of these two types of posts is shown in Table 2.

6 Alignment and Support Type at Post Level

Armed with linguistic alignment measures and social support classifier, we then evaluates the correlation between linguistic alignment and social support types at the post level. We look for repetition between posts from the initial author and later posts, which are classified as either emotional or informational support. We examine whether lexical and syntactic alignment between support seekers and support providers in the post pairs can be used to predict support type. Do lexical and syntactic alignment act similarly or differently?

6.1 Methods

To examine this question, we fit a generalized mixed effects linear regression model with a binomial kernel to predict the emotional support index of a target post. The covariates of the predicting model include lexical alignment (i.e., LILLA), syntactic alignment (i.e., SILLA) in *logit* space, post distance, and interaction terms. Because different types of social support could be influenced by various topics and authors, we also include these variables grouped by *ThreadID* as random effects.

The generalized mixed effects linear regression model is estimated with the lme4 R package [3]. We then use step-wise Akaike Information Criterion (AIC) [7], a measure of the quality of the current logistic regression model, to select the best model without overfitting. Overall, the full model with all the features has the best performance. Table 3 presents the main effects of covariates, and Table 4 presents the main effects and interaction terms.

6.2 Results

Initially, we focus on the effect of lexical alignment. The regression model (Table 3) shows that the lexical adaptation (LILLA) between support providers and support seekers is a reliable indicator of emotional support in both forums. In other words, messages with more emotional support tend to repeat more words in support seekers' posts. We note that this result may also be interpreted by properties of emotional support in both sub-forums. Emotional support presents understanding and empathy, including similar words from support seekers' posts. Also, the model (Table 3) shows that emotional support index generally increases

Table 3: Predicting Emotional Support Index in posts using linguistic alignment between posts from initial author and other replies. This model only includes main effects.

	Breast Cancer Sub-Forum			Colorectal Cancer Sub-Forum		
Predictor	β	**SE**	**p**	β	**SE**	**p**
Intercept	1.593	0.065	0.000	1.971	0.072	0.000
Lexical Alignment	0.127	0.010	0.000	0.188	0.011	0.000
Syntactic Alignment	-0.168	0.013	0.000	-0.180	0.014	0.000
Distance	0.002	0.000	0.000	0.002	0.002	0.001

Table 4: Predicting Emotional Support Index in posts using linguistic alignment between posts from initial author and other replies. The model includes main effects and interactions.

	Breast Cancer Sub-Forum			Colorectal Cancer Sub-Forum		
Predictor	β	**SE**	**p**	β	**SE**	**p**
Intercept	1.357	0.076	0.000	1.849	0.085	0.000
Lexical Alignment	0.080	0.011	0.000	0.171	0.013	0.000
Syntactic Alignment	-0.160	0.016	0.000	-0.185	0.016	0.000
Distance	0.021	0.003	0.000	0.015	0.004	0.145
Lexical Alignment × Distance	0.003	0.000	0.000	0.001	0.001	0.012
Syntactic Alignment × Distance	-0.0003	0.000	0.676	-0.0004	0.001	0.598

with post distance. I.e., conversations tend to shift to emotional support. Another noteworthy result is the effect of syntactic alignment between messages from support seekers and providers. Less syntactic rule repetition occurs when emotional support is given.

Table 4 adds the interaction terms between the post distances and the measure of lexical and syntactic alignment between posts (Lexical Alignment × Distance and Syntactic Alignment × Distance). Compared to the models in Table 3, the effects and directions of predictors are similar. The effects of lexical alignment and post distance in both datasets show that lexical alignment is also predictive for emotional support as distance increases. However, the effect size is small. Furthermore, the effects of syntactic alignment and post distance is not reliable in either dataset. According to the models, syntactic alignment differs from lexical alignment. We do not see a positive correlation between lexical and syntactic alignment in either dataset.

7 Discussion

In this paper, we showed that two types of support relate to the two lower-level linguistic alignment measurements in different ways in two forums. When

peers support one another at the emotional level, they tend to align more with the support seeker at the lexical level than those dialogues that are primarily providing information support. This may be due to active adaptation, or it may be a consequence of the available lexicon associated with the language register (emotional vs. informational): a smaller set of possible words might imply more overlap. This does not necessarily mean that people tailor their message to an audience (in fact, they show very little adaptation in that respect: [29]). However, it is possible that lexical choice is influenced by a desire to demonstrate empathy when the intent is to provide emotional support.

Importantly, however, this principle does not extend to adaptation in syntactic structure. Support providers align with support seekers at the syntactic level when providing informational as opposed to emotional support. This pattern of effects can potentially be explained by a model of communication that suggests different control over the social-level message at the lexical level versus than at the level of sentence structure. Syntactic adaptation is increased in the informational-task situation, which complements previous corpus-based comparisons that showed increased syntactic priming in task-oriented dialogue compared to non-consequential chat [27]. It also complements recent findings of reduced (perceived) phonetic alignment in the speech of conversation partners that are subjected to increased cognitive load [1].

We believe the findings reported by this paper is important because no previous study has suggested that different types of social support dialogue may be associated with different relationship to the lower-level linguistic alignment phenomena. It not only opens the door of studying potential relationships between the higher level functions of dialogue and lower-level alignments, but also suggests potential directions in which existing alignment theories can be enriched through further studies. For example, can we separate the effect of lower-level linguistic alignments to previous replies on a threaded discussion from the effect of linguistic alignment to the thread initiator (i.e., the support seeker)? If so, is the relationship between support types and lower-level alignment measures still hold?

8 Conclusion and Future Work

Social support in online health communities, especially in the forms of emotional support and informational support, benefits chronically ill patients. This support can come from peers, as our dataset demonstrates, and probably benefits support seekers as well as support providers. That is why we think this study is important in improving health care and well-being of patients.

Our analysis models how users, especially support provider, provide social support by adapting to each other. The pattern of adaptation is interesting, even from a theoretical perspective. We observed reliable lexical adaptation, and relatively smaller syntactic adaptation. Adaptation differs with type of support, with emotional support attracting greater lexical and less syntactic adaptation, and informational support being correlated with the opposite.

This study makes potential theoretical progress revealing the relationship between different levels of linguistic adaptation. The result implies that lexical and syntactic adaptation have different adaptation levels and are influenced in different ways by the higher-level support function of the dialogue. This difference may be due to different cognitive bases between these two types of adaptation. These results motivate further research and experimentation regarding the mechanism of linguistic adaptation in web-based conversations.

9 Acknowledgements

References

1. Abel, J., Babel, M.: Cognitive load reduces perceived linguistic convergence between dyads. Language and Speech 0(0), 0023830916665652 (2016)
2. Bambina, A.: Online social support: the interplay of social networks and computer-mediated communication. Cambria Press (2007)
3. Bates, D., Maechler, M., Bolker, B., Walker, S.: lme4: Linear mixed-effects models using Eigen and S4 (2014), r package version 1.1-7
4. Bergmann, K., Kopp, S.: Gestural alignment in natural dialogue. In: Proceedings of the 34th Annual Conference of the Cognitive Science Society. pp. 1326–1331 (2012)
5. Biyani, P., Caragea, C., Mitra, P., Yen, J.: Identifying emotional and informational support in online health communities. In: Proceedings of COLING. pp. 827–836. Dublin, Ireland (2014)
6. Bock, J.K.: Syntactic persistence in language production. Cognitive Psychology 18(3), 355–387 (1986)
7. Burnham, K.P., Anderson, D.R.: Model selection and multimodel inference: a practical information-theoretic approach. Springer (2002)
8. Clark, H.H., Brennan, S.E.: Grounding in communication. In: Resnick, L.B., Levine, J.M. (eds.) Perspectives on socially shared cognition, pp. 127–149 (1991)
9. Danescu-Niculescu-Mizil, C., Gamon, M., Dumais, S.: Mark my words!: linguistic style accommodation in social media. In: Proceedings of the 20th International Conference on the World Wide Web. pp. 745–754. ACM (2011)
10. Doyle, G., Frank, M.C.: Investigating the sources of linguistic alignment in conversation. In: Proceedings of ACL. pp. 526–536 (2016)
11. Doyle, G., Yurovsky, D., Frank, M.C.: A robust framework for estimating linguistic alignment in twitter conversations. In: Proceedings of the 25th International Conference on World Wide Web. pp. 637–648. ACM (2016)
12. Fox, S.: Peer-to-peer healthcare (2011), `http://www.pewinternet.org/2011/02/28/peer-to-peer-health-care-2/`
13. Fox, S.: The social life of health information (2011), `http://www.pewinternet.org/2011/05/12/the-social-life-of-health-information-2011/`

14. Fusaroli, R., Bahrami, B., Olsen, K., Roepstorff, A., Rees, G., Frith, C., Tylén, K.: Coming to terms quantifying the benefits of linguistic coordination. Psychological Science 23(8), 931–939 (2012)

15. Giles, H., Coupland, N., Coupland, J.: Accommodation theory: Communication, context, and consequences. In: Contexts of accommodation: Developments in applied sociolinguistics, pp. 1–68 (1991)

16. Gries, S.T.: Syntactic priming: A corpus-based approach. Journal of Psycholinguistic Research 34(4), 365–399 (2005)

17. Healey, P.G., Purver, M., Howes, C.: Divergence in dialogue. PloS one 9(6), e98598 (2014)

18. Ireland, M.E., Slatcher, R.B., Eastwick, P.W., Scissors, L.E., Finkel, E.J., Pennebaker, J.W.: Language style matching predicts relationship initiation and stability. Psychological science 22(1), 39–44 (2011)

19. Jones, S., Cotterill, R., Dewdney, N., Muir, K., Joinson, A.: Finding zelig in text: A measure for normalising linguistic accommodation. In: Proceedings of COLING. pp. 455–465. University of Bath (2014)

20. Kaschak, M.P., Kutta, T.J., Jones, J.L.: Structural priming as implicit learning: Cumulative priming effects and individual differences. Psychonomic Bulletin & Review 18(6), 1133–1139 (2011)

21. Malecki, C.K., Demaray, M.K.: What type of support do they need? investigating student adjustment as related to emotional, informational, appraisal, and instrumental support. School Psychology Quarterly 18(3), 231–252 (2003)

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

23. Noble, B., Fernández, R.: Centre stage: How social network position shapes linguistic coordination. Proceedings of CMCL pp. 29–38 (2015)

24. Pickering, M.J., Garrod, S.: The interactive-alignment model: Developments and refinements. Behavioral and Brain Sciences 27(02), 212–225 (2004)

25. Portier, K., Greer, G.E., Rokach, L., Ofek, N., Wang, Y., Biyani, P., Yu, M., Banerjee, S., Zhao, K., Mitra, P., Yen, J.: Understanding topics and sentiment in an online cancer survivor community. JNCI Monographs 2013(47), 195–198 (2013)

26. Reitter, D., Keller, F., Moore, J.D.: A computational cognitive model of syntactic priming. Cognitive Science 35(4), 587–637 (2011)

27. Reitter, D., Moore, J.D.: Alignment and task success in spoken dialogue. Journal of Memory and Language 76, 29–46 (2014)

28. Vlahovic, T.A., Wang, Y.C., Kraut, R.E., Levine, J.M.: Support matching and satisfaction in an online breast cancer support community. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. pp. 1625–1634. ACM (2014)

29. Wang, Y., Reitter, D., Yen, J.: Linguistic adaptation in conversation threads: Analyzing alignment in online health communities. Proceedings of CMCL pp. 55–62 (2014)

30. Wang, Y., Yen, J., Reitter, D.: Pragmatic alignment on social support type in health forum conversations. Proceedings of CMCL pp. 9–18 (2015)

31. Wang, Y.C., Kraut, R., Levine, J.M.: To stay or leave?: The relationship of emotional and informational support to commitment in online health support groups. In: Proceedings of CSCW. pp. 833–842. ACM (2012)

32. Xu, Y., Reitter, D.: An evaluation and comparison of linguistic alignment measures. Proceedings of CMCL pp. 58–67 (2015)