



UvA-DARE (Digital Academic Repository)

Centre Stage: How Social Network Position Shapes Linguistic Coordination

Noble, B.; Fernández, R.

Publication date

2015

Document Version

Final published version

Published in

NAACL HLT 2015 : The Sixth Workshop on Cognitive Modeling and Computational Linguistics (CMCL)

License

Article 25fa Dutch Copyright Act

[Link to publication](#)

Citation for published version (APA):

Noble, B., & Fernández, R. (2015). Centre Stage: How Social Network Position Shapes Linguistic Coordination. In T. O'Donnell, & M. van Schijndel (Eds.), *NAACL HLT 2015 : The Sixth Workshop on Cognitive Modeling and Computational Linguistics (CMCL): proceedings of the workshop : June 4, 2015, Denver, Colorado, USA* (pp. 29-38). The Association for Computational Linguistics. <http://www.aclweb.org/anthology/W15-1104.pdf>

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

Centre Stage: How Social Network Position Shapes Linguistic Coordination

Bill Noble and Raquel Fernández

Institute for Logic, Language and Computation
University of Amsterdam

winobes@gmail.com, raquel.fernandez@uva.nl

Abstract

In conversation, speakers tend to echo the linguistic style of the person they are interacting with. This paper contributes to a body of work that addresses how this linguistic style coordination is affected by the social context in which the interaction occurs. In particular, we investigate the effect that an agent’s social network centrality has on the coordination exhibited in replies to their utterances. We find that linguistic coordination is positively correlated with social network centrality and that this effect is greater than previous results showing a similar connection between status-based power and linguistic coordination. We conjecture that the social value of coordination may reside in the wish to conform to the linguistic norms of a community.

1 Introduction

In communicative contexts, there is more to language use than the individual processing of representations. When two or more interlocutors take part in a conversation, they engage in a *joint* activity — a type of social interaction that requires an intricate level of interpersonal coordination. This often leads to interlocutors converging on similar patterns of language use, including phonetic production (Kim et al., 2011; Babel, 2012), lexical choice (Brennan, 1996), use of function words (Niederhoffer and Pennebaker, 2002), and of syntactic constructions (Pickering and Ferreira, 2008).

It is a matter of debate what mechanisms give rise to the observed convergences and whether different factors may simultaneously and complementarily be

at play. For instance, the “collaborative” approach led by Clark (1996) considers that adaptation or entrainment is mainly motivated by the communicative need to reach mutual understanding, which leads to speakers reasoning about their common ground (Clark and Murphy, 1982; Brennan and Clark, 1996; Brennan and Hanna, 2009). In contrast, Pickering and Garrod (2004) have argued that stimulus-response priming is the key mechanism underlying alignment of representations in conversation (Braniigan et al., 1995; Pickering and Garrod, 2006; Reitter and Moore, 2014). Yet, within social psychology researchers have emphasised the role of social processes and goals as triggers of linguistic “imitation” (Shepard et al., 2001; Giles, 2008; Babel, 2012).

In this paper, we contribute to this latter line of research by investigating the effects of social factors on linguistic coordination. In particular, we exploit notions of *network centrality* from Social Network Analysis (Wasserman, 1994) to study the extent to which an individual’s position in a social network is connected to differences in linguistic adaptation observed within a community. We use online discussions amongst Wikipedia¹ editors as our case study, leveraging a corpus compiled by Danescu-Niculescu-Mizil et al. (2012), who found that the social status of editors (whether they have the role of administrator) is relevant to explain the patterns of linguistic coordination observed in this online community. In the present study we show that network centrality factors (which are largely implicit) also bear on linguistic coordination in this scenario, largely independently of explicit social status.

¹<https://www.wikipedia.org>

The paper proceeds as follows: In Sections 2 and 3, we introduce key notions related to socially-driven linguistic coordination and social network analysis, respectively, and review related work in these areas. In Section 4, we put forward our working hypotheses. The experimental setup and the particular measures we use to test these hypotheses are described in Section 5. In Section 6, we present our results in detail. Finally, we conclude in Section 7 with a discussion of the implications of these results.

2 Socially-Driven Linguistic Coordination

The influence of social factors on how we linguistically communicate with each other has been studied, amongst others, by sociologists working within the framework of Conversation Analysis (Atkinson and Drew, 1979; Heritage, 2005) and by social psychologists employing more quantitative approaches such as Communication Accommodation Theory (Giles, 2008). CAT claims that linguistic adaptation is motivated by individuals' aim to be socially accepted and to negotiate the social distance that separates them from their interlocutors (increasing, maintaining, or decreasing it). Such adaptation can take place at different levels: pitch, vocabulary, gestures, etc. (Giles et al., 1991). An approach in this direction has been put forward by Pennebaker and colleagues, who focus on *style matching*, in particular the matching of function words such as pronouns, articles, and quantifiers (Chung and Pennebaker, 2007). For instance, it has been shown that function word matching in speed dating conversations predicts relationship initiation and stability (Ireland et al., 2011) and that it is indicative of relative social status between interlocutors (Niederhoffer and Pennebaker, 2002). An advantage of focusing on function words is that their choice is genuinely *stylistic*, i.e. not directly related to the topic of the conversation and thus largely domain independent.

Linguistic style matching has been exploited by Danescu-Niculescu-Mizil et al. (2012) to investigate power differences in two domains: the community of Wikipedia editors and the U.S. Supreme Court. The authors found that power differences between interlocutors bear upon the degree to which a speaker echoes the linguistic style of the addressee to whom they are responding. In the Wikipedia

domain (which is the most relevant for our own study), it was observed that speakers tend to linguistically coordinate more with editors who have the role of administrator. Adminship confers a certain authority since these editors can block user accounts and protect or delete Wikipedia pages.² Therefore admins have a higher social status than other editors (non-admins), which endows them with *status-based power*.

We examine the same corpus of textual conversational exchanges amongst Wikipedia editors, but in addition to status-based power, we consider linguistic style coordination in relation to an individual's *position in a social network structure*. In particular, we investigate the effect that a speaker's network *centrality* has on how much other individuals coordinate with her linguistically.

3 Importance in Social Network Structure

A social network is a graph model of a community where nodes represent individuals (of some kind) and edges represent links between those individuals. Edges may be weighted to capture the strength of certain links or directed to represent asymmetrical relationships. Given a social network, one may extract information about how *important* an individual is in the community. Importance is of course a slippery notion. Which individuals are important depends on what it means to be important in a particular community and how these criteria are encoded in the network model. Network centrality is a family of measures that attempt to capture importance by assigning a numerical value to each node based on its position in the network structure. We shall consider two measures of network centrality, which will be defined in Section 5.3.

High centrality might be seen as a source of power (be it status-based or of some other kind), but this is not always so. A case in point are exchange networks: An exchange network is one where social relations involve the exchange of valued commodities, be they physical, like goods and services, or less tangible, like affection or information. In such networks, power often increases with access to non-central individuals who have less choice in partners

²<https://en.wikipedia.org/wiki/Wikipedia:Administrators>

for exchange. In such situations it may present a power advantage *not* to be centrally located (Cook et al., 1983). Regardless of whether importance corresponds to high or low centrality values in a particular social community, network centrality certainly does not confer any institutionalised title or explicit authority. For this reason we consider centrality as related to *implicit social power*, to be considered in parallel with status-based power.

The ease with which social network structure can be extracted from online communities has made network analysis of such communities a very active method of research in different fields concerned with social interaction at large, not necessarily with language (Chakrabarti, 2003; Guha et al., 2004; Leskovec et al., 2010). In sociolinguistics, however, the effect of network structure on language had been observed long before the prevalence of the Internet. Sociolinguists have examined the relationship between social network features and aspects of linguistic variation and change to draw conclusions on how the linguistic behaviour of individuals reflects their membership in small-scale social clusters (Milroy, 1987; Milroy and Milroy, 1997). For example, Eckert (1988) considers the effects of the social network of suburban Detroit area adolescents on their susceptibility to phonological innovations. She argues that linguistic change can be better explained by features of the social network than by unstructured demographic data and that alignment of linguistic styles is an important factor in maintaining acceptance in a rapidly developing social structure.

Thanks to the ubiquity of the Internet, it is now possible to apply social network analysis techniques to address sociolinguistic concerns to much larger amounts of linguistic data than ever before. Here we exploit this opportunity, in particular the availability of interactional data from a rich online community such as the Wikipedia editors. Although Wikipedia has been used as a testbed to study the connection between structural network properties and factors such as contentious topics (Laniado et al., 2011), quality improvement of Wikipedia articles (Kittur and Kraut, 2008) or editing activity (Crandall et al., 2008), to the best of our knowledge this is the first study that investigates the links between social network features of this community and linguistic style coordination amongst its members.

4 Hypotheses

We investigate the following three hypotheses that relate linguistic style coordination to the concept of social network centrality:

- H1. *Speakers coordinate more towards individuals that occupy more central social positions.*
- H2. *Individuals in more central social network positions tend to possess status-based power.*
- H3. *The effect hypothesized in H1 holds independently of any effect observed in relation to H2.*

While we are primarily interested in the relationship between linguistic coordination and network centrality, we do so in a context where something is known about the status-based power born by individuals: We have access to the adminship role of editors and editors themselves are aware of the admin status of other users.³ We expect to find that social network centrality correlates with status-based power, i.e., that individuals at the centre stage of the community are more likely to be admins (H2). With this in mind, we would like to separate the effect of status-based power on coordination from that of implicit centrality-based power (H3). Given that linguistic style coordination correlates with status-based power (Danescu-Niculescu-Mizil et al., 2012), it is not enough simply to show that it also correlates with network centrality (H1) since such a result may be wholly explained by status-based power as a confounding factor.

5 Experimental Setup

In this section we describe the corpus used in our experiments and define the measures of linguistic coordination and network centrality that we compute from the data.

5.1 Data

The Wikipedia Talk Page Conversations Corpus consists of a collection of exchanges from Wikipedia editors' *user talk pages*. A talk page, as opposed to a Wikipedia *article*, is a page for discussion between

³There is a symbol identifying admins as such on their Wikipedia user page, although it is worth noting that no such identifying marks are visible where discussions take place.

editors that is not part of the content of Wikipedia. User talk pages, which are not associated with a particular article, tend to feature more community-oriented discussions.⁴

The corpus contains information on 26,397 users, including whether or not the editor is a Wikipedia administrator (an *admin*). There are 1,825 admins. Each utterance (or *post*) is annotated with metadata including the username of the editor who made the post and which previous contribution it is a reply to (if any). Of 391,294 total posts in the corpus, we consider a subset of 342,800 that were made by users whose admin status is known (i.e., by one of the 26,397 for whom we have metadata). The corpus was collected in August 2011 and made available by Danescu-Niculescu-Mizil et al. (2012).⁵

We derived a weighted network structure from the corpus as follows: A node was created for each individual. Undirected, weighted edges were formed between editors based on the number of direct replies between them. An edge with weight w between user a and user b indicates that w is the total number of times user a replied to a post of user b or *visa versa*. The edges in this model are intended to represent the degree to which two users *know* each other (as members of the Wikipedia editors' community). The talk page is the locus of an editor's involvement in Wikipedia as a community. The rationale for our edge definition is that a 's reply to b 's contribution is directed at b as a member of the community. The more that a and b reply to one another, the better connected they are in the network.

The resulting network was pruned to its largest connected component such that there is a path between every pair of nodes. Pruning eliminated 575 users from 556 different disconnected components (mostly singletons). The final network consists of 25,822 nodes and 103,992 edges with an average weight of 3.3.

5.2 Linguistic Coordination Measures

We want to measure how much participants align their language with that of the interlocutors to whom they are immediately replying. Following Danescu-

⁴http://en.wikipedia.org/wiki/Wikipedia:Talk_page_guidelines

⁵http://www.cs.cornell.edu/~cristian/Echoes_of_power.html

Niculescu-Mizil et al. (2012), we use the presence of a word in a particular category of *function words* as linguistic style markers. We consider the same eight categories of functional markers as these authors: quantifiers, personal pronouns, impersonal pronouns, articles, auxiliary verbs, conjunctions, prepositions, and adverbs. However, while Danescu-Niculescu-Mizil and colleagues use the markers provided by the commercial text analysis software LIWC (Pennebaker et al., 2007),⁶ we compiled our own list of markers for each of these eight categories using freely available frequency lists of part-of-speech classes from the British National Corpus (Burnard, 2000).⁷ We took the most common words for each relevant POS (manually filtering out any content words) to match the length of the LIWC category lists reported in the software documentation.⁸

Linguistic style coordination quantifies the degree to which an agent b immediately echoes the linguistic style of agent a . We use the linguistic style coordination measure $C^m(b, a)$ defined by Danescu-Niculescu-Mizil, et al. (2012), where the coordination of b (the *speaker*) towards a (the *target*) with respect to a marker m encodes how much a 's use of m increases the probability that b will use that marker in her reply to a , relative to the overall frequency of m in b 's replies to a :⁹

$$C^m(b, a) = P(\mathcal{E}_{u_b}^m | \mathcal{E}_{u_a}^m) - P(\mathcal{E}_{u_b}^m)$$

where \mathcal{E}_u^m indicates that utterance u exhibits a marker m and (u_a, u_b) belongs to the set of pairs of utterances made by b in response to a . If none of the utterances of a that b responds to exhibit m , then $C^m(b, a)$ is undefined.

We introduce a variant of this measure that considers the coordination of a group of speakers B towards an individual a by making the same calculation across all utterance pairs (u_a, u_B) where some member of B is replying to individual a . When B

⁶<http://www.liwc.net>

⁷<http://ucrel.lancs.ac.uk/bncfreq/flists.html>

⁸The relevant LIWC documentation can be found at <http://www.liwc.net/descriptiontable1.php>. Our lists of markers are freely available upon request.

⁹This ensures that $C^m(b, a)$ captures the influence of a 's use of marker m on b 's immediate reply. It may be that b uses m more *in general* when speaking to a , but such an effect would have no bearing on $C^m(b, a)$.

is the set of all members of a social network who have addressed a , we refer to $C^m(B, a)$ as a 's *coordination received*. This is the main measure we will employ in the analyses reported in Section 6.

It is sometimes desirable to have a single score that combines a 's coordination received across all markers. This is made complicated by the fact that coordination may be undefined for one or more markers. Since we generally consider coordination received by a in the context of a 's membership in some group A (e.g., admins), there are several different aggregation schemes available. Again, we follow Danescu-Niculescu-Mizil et al. (2012) in the naming and definition of these schemes, but apply them to *coordination received*.

Aggregate 1 Take a simple average across markers, but only for those users a in A for whom $C^m(B, a)$ is defined for all markers. Otherwise the aggregate is undefined for a .

Aggregate 2 Wherever $C^m(B, a)$ is undefined, substitute with the average of $C^m(B, a')$ across those a' in A for whom it is defined.

Aggregate 3 Whenever $C^m(B, a)$ is undefined, substitute with the average of $C^{m'}(B, a)$ across those m' for which it is defined.

When taking an average aggregate coordination received over a group of users A , aggregate 1 takes into account only those users for whom all measures are defined. Aggregates 2 and 3 take into account all users for whom at least one measure is defined, but exhibit slightly different smoothing assumptions. Aggregate 2 assumes that people in B would have behaved towards a with regard to m as they did towards the rest of A , whereas aggregate 3 assumes that members of B would have behaved towards a with regard to m as they did with regard to the other markers.

5.3 Network Centrality Measures

As mentioned in Section 3, a network centrality measure assigns a numerical value to each individual in the network based on features of their position in the graph. This value is intended to represent the importance of that individual in the social network.

What kind of importance it captures depends on exactly how centrality is calculated. Here we consider two well-known measures of network centrality.

Eigenvector centrality (Bonacich, 1987) tries to capture the notion that your importance in a network depends on the importance of your closest contacts. Let $M(n)$ be the *neighborhood* of n ; that is, the nodes in N that share an edge with n . Then the Eigenvector centrality of n^* is defined by

$$EC(n^*) = \frac{1}{\lambda} \sum_{n \in M(n^*)} EC(n)$$

where λ is a constant, the *eigenvalue*. There may be multiple values of λ for which the Eigenvector centrality is defined, but taking the largest value provides a consistent measure across the network.

Betweenness centrality (Freeman, 1977) measures how much a node contributes to the overall connectivity of the network. Nodes who lie on more *shortest paths* between pairs of other nodes have higher Betweenness centrality. Specifically it looks at all of the shortest paths between each pair of nodes, and counts how many of them contain the node in question. Letting $\text{Path}(m, n)$ stand for the set of shortest paths between m and n , the Betweenness centrality of n^* is defined by:

$$BC(n^*) = \sum_{n \neq m \in N} \frac{|\{\sigma \in \text{Path}(m, n) | n^* \in \sigma\}|}{|\text{Path}(m, n)|}$$

Both Eigenvector and Betweenness centrality measures have generalizations for weighted networks which we use here. For Betweenness, path length is calculated using the inverse weight of edges (so that paths along edges with higher weights are considered shorter). For Eigenvector centrality, the notion of “neighbour” is adjusted so that adjacent nodes connected with a higher weight count for more. We use the implementation of these algorithms available in the Python library NetworkX (Hagberg et al., 2008).¹⁰

6 Analyses and Results

To investigate our hypotheses regarding the impact of social network position on linguistic coordination, we computed scores for all the measures in-

¹⁰<https://networkx.github.io>

	Agg. 3	Eigenvector	Betweenness	Users
admins	1.85 (6.44)	1.01 (15.9)	1.66 (9.44)	1825
non-admins	0.48 (6.85)	0.16 (4.72)	0.14 (2.67)	23,997
Total	0.60 (6.83)	0.22 (6.22)	0.25 (3.62)	25,822

Table 1: Descriptive statistics of the computed measures: Averages (and standard deviations) for Coordination received according to Aggregate 3 (scaled by 100), Eigenvector and Betweenness centrality (scaled by 1000) for admin and non-admin users.

roduced in the previous section (coordination received, Eigenvector and Betweenness centrality) for each individual in the social network. Table 1 provides some basic descriptive statistics.

6.1 Coordination and Status-Based Power

We start by replicating the relevant result by Danescu-Niculescu-Mizil et al. (2012), according to which Wikipedia editors coordinate more towards admins than non-admins. We compare the average linguistic coordination received by each of these two social groups for each functional marker as well as for each of the aggregate measures. The results are shown in Figure 1. As can be seen, in all cases individuals coordinate significantly more towards admin addressees than towards non-admins (independent Welch two sample t-test, $p < 0.001$). We are thus able to reproduce this basic result despite the fact that we use a self-compiled list of functional markers instead of the full power of the LIWC tool (Pennebaker et al., 2007). The results we obtain are in fact stronger, since we observe high levels of significance for all aggregate measures and markers while Danescu-Niculescu-Mizil and colleagues obtained significant results only for the aggregate measures and for conjunctions, indefinite pronouns, adverbs and articles. We point out however that regardless of the high significance values across the board, the size of the effect is larger for the aggregate measures (average Cohen’s $d = 0.2$) than for any of the individual markers (for which the effect size is in fact very small: on average $d < 0.15$).¹¹

6.2 Coordination and Centrality

We now turn to investigate each of the hypotheses formulated in Section 4. Our data provide evidence

¹¹Danescu-Niculescu-Mizil et al. (2012) do not report effect size and hence a more detailed comparison is not possible.

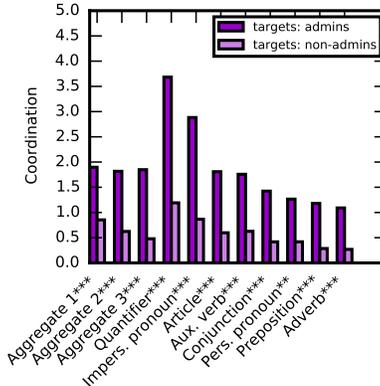


Figure 1: Linguistic style coordination towards admins/non-admins. Note on all figures: Coordination scores are reported as percentages for clarity (i.e., multiplied by 100). Measures are marked for significance by independent t-test as follows: * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$.

in support of H1 (*speakers coordinate more towards individuals that occupy more central social positions*). As shown in Table 2, we find significant (albeit weak) positive correlations in all cases between the level of coordination received by an individual a ($C^m(B, a)$) and a ’s position in the social network, as quantified by Eigenvector and Betweenness centrality scores.¹² Betweenness centrality correlates slightly better with coordination received.

	Agg. 1	Agg. 2	Agg. 3
Eigenvector	0.1911	0.1175	0.1878
Betweenness	0.2028	0.1491	0.2113

Table 2: Spearman’s rank correlation ρ between network centrality measures and linguistic coordination received ($p < 0.001$ for all values).

We consider an individual *highly central* with respect to some centrality measure if this individual’s centrality score is higher than one standard deviation above the mean score. Given the large amount of variation in our dataset (see Table 1), this is a very selective criterion: Out of 25,822 editors, only 119 (~0.5%) are highly Eigenvector-central and 239 (~0.9%) are highly Betweenness-central. We observe that speakers coordinate more with individu-

¹²This is also the case for the individual markers, which are not shown in Table 2 for conciseness.

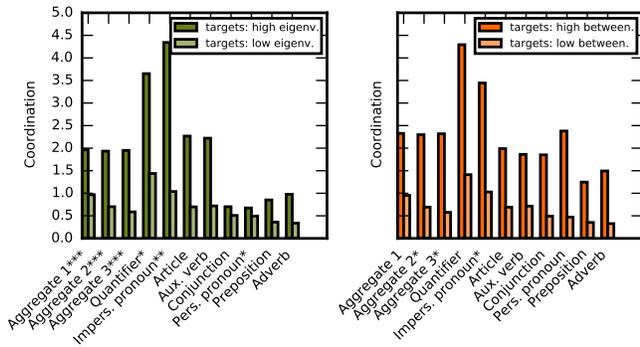


Figure 2: Linguistic style coordination towards users with high/low Eigenvector and Betweenness centrality.

als that are highly central, significantly so for most of the aggregate measures, pronouns and quantifiers (independent Welch two sample t-test; see Figure 2). The average effect size for the aggregate measures is $d = 0.32$ for Eigenvector centrality and $d = 0.35$ for Betweenness centrality.

Regarding H2 (*individuals in more central social network positions tend to possess status-based power*), in our dataset only 29% of editors who are highly Eigenvector-central are also admins. The percentage goes up to 53% in the case of highly Betweenness-central editors. Editors with admin status make up around 45% of those individuals who are highly central according to at least one measure (319 editors) and around 49% of those who are highly central according to both measures (39 editors). To further investigate the connection between status-based power as represented by adminship and network centrality, we examined the mean centrality scores of admin versus non-admin editors. We find that on average admins are more centrally positioned in the network than non-admins. An independent Welch t-test confirms that this is significant for both Eigenvector centrality ($p < 0.5$, Cohen’s $d = 0.07$) and Betweenness centrality ($p < 0.001$, Cohen’s $d = 0.22$), although the effect is practically nonexistent for the former centrality measure. Thus, H2 is only relatively supported, with Betweenness centrality exhibiting a closer connection with adminship than Eigenvector centrality.

Finally, to assess H3 (*the effect hypothesized in H1 holds independently of any effect observed in re-*

lation to H2), we compared the average coordination received scores for admins and non-admins within each class of highly central users. Our aim was to check whether users coordinate more towards editors who are both administrators and highly central in the social network. As shown in Figure 3 (left), no effects of adminship were found (for all markers and aggregate measures, $p > 0.05$ in an independent Welch t-test), which provides evidence supporting the hypothesis. Analogous calculations were made for users who are not highly central (Figure 3, right). Amongst these, significant differences were found (albeit with small effect sizes: $0.16 < d < 0.2$ across the board for both centrality measures), with admins receiving more linguistic coordination for all markers and aggregate measures.¹³

7 Discussion and Conclusions

In this paper we have put forward the hypothesis that speakers coordinate more with targets who occupy a more central position in the social context in which the communication takes place. We have provided evidence for this claim by measuring linguistic style coordination in the Wikipedia talk pages corpus.

We confirmed the result by Danescu-Niculescu-Mizil et al. (2012) that correlates coordination with explicit status-based power represented by adminship, and went on to show that there is a further positive relationship between how much linguistic style coordination an editor receives and her network centrality. Furthermore, while editors coordinate more with admins in general, we found that adminship has no significant effect on how much coordination *highly central* editors receive. We may conclude that in certain situations, considerations of network centrality trump explicit status-based power in determining how much a speaker immediately aligns with the person to whom she is responding.

This conclusion provides evidence for the claim of Communication Accommodation Theory according to which linguistic adaptation is motivated by the desire for social acceptance (Giles, 2008). Exactly how aligning with highly central members of

¹³We chose this analysis over a regression analysis because the data violates the normality assumption and is affected by very severe heteroscedasticity, i.e., the variance in the error is not constant, with very large residuals when the centrality measures are low and much smaller ones as they increase.

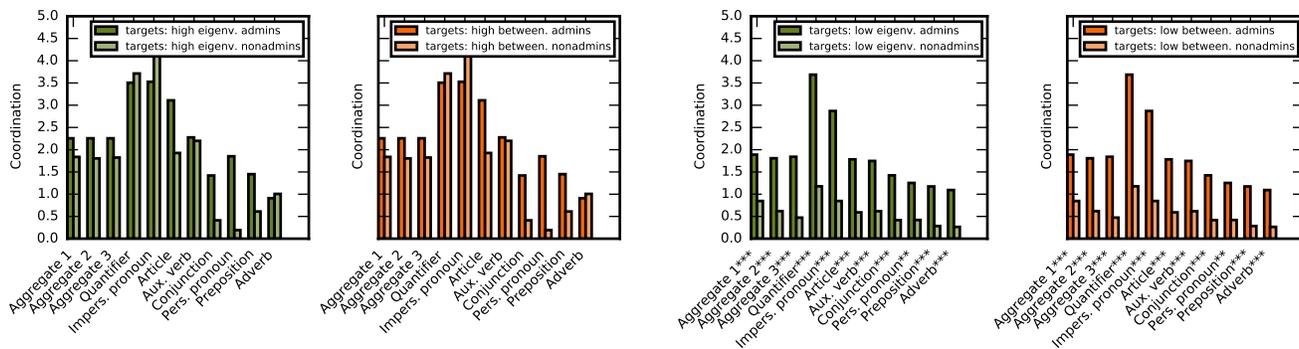


Figure 3: Linguistic style coordination towards admins/non-admins among users with high (left) and low (right) Eigenvector/Betweenness centrality.

the community achieves this goal is open to some interpretation, though it is likely that there is more than one mechanism at play. We consider two possibilities.

First, aligning with highly central community members can be seen as an instance of aligning to power, since network centrality (especially Eigenvector centrality) is often used as a proxy for implicit social power. Aligning with highly central members helps to achieve social acceptance since those with implicit social power have more power to grant it. This interpretation is supported by our results: Just as coordination more closely follows network centrality than it does adminship, it is natural to assume that the power to confer social acceptance more closely follows implicit social power than it does any official title.

The second possible interpretation has more expressly linguistic motivations. It has been observed that those in central social positions make more utterances characteristic of the group they are central to (Eckert, 1988; Kooti et al., 2012). Since learning the linguistic practices of a community is important to social acceptance, it is beneficial to coordinate with highly central members of the community as a way of picking up those linguistic practices. In other words, coordination towards a highly central individual may have a social goal that does not have anything to do with that individual in particular, but rather with adapting to the linguistic norms of the community at large.

Although Eigenvector centrality is most closely associated with implicit power (Bonacich, 1987), we

found that in the Wikipedia corpus it is actually Betweenness centrality that correlates somewhat more strongly with coordination. This may have something to say about the relationship between the two mechanisms mentioned above. While it may still be true that people align to power for the immediate social benefit, the greater effect of Betweenness centrality suggests that the social benefit of coordination may be mediated by something more than the power of the individual who is being responded to. One possible explanation is that community members who are more vital to the connectivity of the social network (i.e., those with high Betweenness centrality) tend to conform better with the linguistic norms of the community as a whole (rather than, for example, to the norms of some clique or subgroup). Assuming that some of the social motivations for alignment are expressly linguistic, this would explain why Betweenness centrality correlates better with coordination received than the centrality measure more typically associated with social power. More research is needed to determine whether (a) community members with high Betweenness centrality better represent the linguistic norms of the community and (b) immediate linguistic coordination is the mechanism by which those norms are propagated.

The results of this paper are suggestive in both those regards, while providing good evidence that, at the very least, network centrality is a more important factor in linguistic coordination than is formal status-based power.

References

- John Maxwell Atkinson and Paul Drew. 1979. *Order in court: The organisation of verbal interaction in judicial settings*. Macmillan London.
- Molly Babel. 2012. Evidence for phonetic and social selectivity in spontaneous phonetic imitation. *Journal of Phonetics*, 40(1):177–189.
- Phillip Bonacich. 1987. Power and centrality: A family of measures. *American journal of sociology*, pages 1170–1182.
- Holly P. Branigan, Martin J. Pickering, Simon P. Liv-ersedge, Andrew J. Stewart, and Thomas P. Urbach. 1995. Syntactic priming: Investigating the mental representation of language. *Journal of Psycholinguistic Research*, 24(6):489–506.
- Susan E. Brennan and Herbert H. Clark. 1996. Conceptual Pacts and Lexical Choice in Conversation. *Journal of Experimental Psychology*, 22(6):1482–1493.
- Susan E. Brennan and Joy E. Hanna. 2009. Partner-specific adaptation in dialog. *Topics in Cognitive Science*, 1(2):274–291.
- Susan E. Brennan. 1996. Lexical entrainment in spontaneous dialog. In *Proc. of the International Symposium on Spoken Dialogue*, pages 41–44.
- Lou Burnard. 2000. *Reference Guide for the British National Corpus (World Edition)*. Oxford University Computing Services.
- Soumen Chakrabarti. 2003. *Mining the Web: Discovering knowledge from hypertext data*. Morgan Kaufmann.
- Cindy Chung and James W. Pennebaker. 2007. The psychological functions of function words. *Social Communication*, pages 343–359.
- Herbert H. Clark and Gregory L. Murphy. 1982. Audience design in meaning and reference. *Advances in psychology*, 9:287–299.
- Herbert H. Clark. 1996. *Using language*. CUP.
- Karen S. Cook, Richard M. Emerson, Mary R. Gillmore, and Toshio Yamagishi. 1983. The distribution of power in exchange networks: Theory and experimental results. *American journal of sociology*, pages 275–305.
- David Crandall, Dan Cosley, Daniel Huttenlocher, Jon Kleinberg, and Siddharth Suri. 2008. Feedback effects between similarity and social influence in online communities. In *Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 160–168.
- Cristian Danescu-Niculescu-Mizil, Lillian Lee, Bo Pang, and Jon Kleinberg. 2012. Echoes of power: Language effects and power differences in social interaction. In *Proceedings of the 21st international conference on World Wide Web*, pages 699–708. ACM.
- Penelope Eckert. 1988. Adolescent social structure and the spread of linguistic change. *Language in Society*, 17(2):183–207.
- Linton C. Freeman. 1977. A set of measures of centrality based on betweenness. *Sociometry*, pages 35–41.
- Howard Giles, Justine Coupland, and Nikolas Coupland. 1991. Accommodation theory: Communication, context, and consequence. *Contexts of accommodation: Developments in applied sociolinguistics*, 1.
- Howard Giles. 2008. Communication accommodation theory. In Leslie A. Baxter and Dawn O. Braithewaite, editors, *Engaging theories in interpersonal communication: Multiple perspectives*, pages 161–173. Sage Publications, Inc.
- Ramanathan Guha, Ravi Kumar, Prabhakar Raghavan, and Andrew Tomkins. 2004. Propagation of trust and distrust. In *Proceedings of the 13th international conference on World Wide Web*, pages 403–412.
- Aric A. Hagberg, Daniel A. Schult, and Pieter J. Swart. 2008. Exploring network structure, dynamics, and function using NetworkX. In *Proceedings of the 7th Python in Science Conference (SciPy2008)*, pages 11–15, Pasadena, CA USA, August.
- John Heritage. 2005. Conversation analysis and institutional talk. In *Handbook of language and social interaction*, pages 103–147. Erlbaum.
- Molly E. Ireland, Richard B. Slatcher, Paul W. Eastwick, Lauren E. Scissors, Eli J. Finkel, and James W. Pennebaker. 2011. Language style matching predicts relationship initiation and stability. *Psychological Science*, 22(1):39–44.
- Midam Kim, William S. Horton, and Ann R. Bradlow. 2011. Phonetic convergence in spontaneous conversations as a function of interlocutor language distance. *Laboratory Phonology*, 2(1):125–156.
- Aniket Kittur and Robert E. Kraut. 2008. Harnessing the wisdom of crowds in wikipedia: quality through coordination. In *Proceedings of the 2008 ACM conference on Computer supported cooperative work*, pages 37–46.
- Farshad Kooti, Haeryun Yang, Meeyoung Cha, Krishna P. Gummadi, and Winter A. Mason. 2012. The emergence of conventions in online social networks. In *ICWSM*.
- David Laniado, Riccardo Tasso, Yana Volkovich, and Andreas Kaltenbrunner. 2011. When the Wikipedians talk: Network and tree structure of Wikipedia discussion pages. In *ICWSM*.
- Jure Leskovec, Daniel Huttenlocher, and Jon Kleinberg. 2010. Predicting positive and negative links in online social networks. In *Proceedings of the 19th international conference on World Wide Web*, pages 641–650.

- James Milroy and Lesley Milroy. 1997. Network structure and linguistic change. In Nikolas Coupland and Adam Jaworski, editors, *Sociolinguistics: a reader*. St. Martin's Press.
- Lesley Milroy. 1987. *Language and Social Networks*. Oxford: Blackwell, 2nd edition.
- Kate G Niederhoffer and James W Pennebaker. 2002. Linguistic style matching in social interaction. *Journal of Language and Social Psychology*, 21(4):337–360.
- James W. Pennebaker, Martha E. Francis, and Roger J. Booth. 2007. Linguistic Inquiry and Word Count (LIWC): A computerized text analysis program. Technical report, LIWC.net, Austin, Texas.
- Martin J. Pickering and Victor S. Ferreira. 2008. Structural priming: a critical review. *Psychological Bulletin*, 134(3):427–459.
- Martin J. Pickering and Simon Garrod. 2004. Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, 27(02):169–190.
- Martin J. Pickering and Simon Garrod. 2006. Alignment as the basis for successful communication. *Research on Language and Computation*, 4(2-3):203–228.
- David Reitter and Johanna D. Moore. 2014. Alignment and task success in spoken dialogue. *Journal of Memory and Language*, 76:29–46.
- Carolyn A. Shepard, Howard Giles, and Beth A. Le Poire. 2001. Communication accommodation theory. In W. P. Robinson and H. Giles, editors, *The new Handbook of Language and Social Psychology*, pages 33–56. John Wiley & Sons Ltd.
- Stanley Wasserman. 1994. *Social network analysis: Methods and applications*, volume 8. Cambridge University Press.