

# Annotations for Power Relations on Email Threads

Vinodkumar Prabhakaran\*, Huzaifa Neralwala†, Owen Rambow‡, Mona Diab‡

\* Computer Science Department, Columbia University, New York, NY 10027, USA.

† LinkedIn Corp., 2029 Stierlin Ct., Mountain View, CA 94043, USA.

‡ Center for Computational Learning Systems, Columbia University, New York, NY 10027, USA.

vinod@cs.columbia.edu, hneralwala@linkedin.com, rambow@ccls.columbia.edu, mdiab@ccls.columbia.edu

## Abstract

Social relations like power and influence are difficult concepts to define, but are easily recognizable when expressed. In this paper, we describe a multi-layer annotation scheme for social power relations that are recognizable from online written interactions. We introduce a typology of four types of power relations between dialog participants: hierarchical power, situational power, influence and control of communication. We also present a corpus of Enron emails comprising of 122 threaded conversations, manually annotated with instances of these power relations between participants. Our annotations also capture attempts at exercise of power or influence and whether those attempts were successful or not. In addition, we also capture utterance level annotations for overt display of power. We describe the annotation definitions using two example email threads from our corpus illustrating each type of power relation. We also present detailed instructions given to the annotators and provide various statistics on annotations in the corpus.

**Keywords:** Social Power, Email Corpus, Annotations

## 1. Introduction

In this paper, we propose a typology of social power in the context of online written interactions and we present an annotated email corpus capturing these types of power relations between participants. Recently, various studies have tried to infer social power relations from emails using the Enron email corpus (Diesner and Carley, 2005; Creamer et al., 2009; Bramsen et al., 2011; Peterson et al., 2011). These studies relied upon Enron’s organizational hierarchy as the reference for actual power relations. This limits such approaches’ applicability to genres for which a pre-set hierarchy is absent. Also, when people interact with one another, even within an organizational setting, there are various other forms of power that come into play, which may or may not align with their hierarchical roles. We aim to capture these other types of power (as well as power attributed by hierarchical position) that are recognizable from the nature of the interaction. Specifically, we capture four types of power: hierarchical power, situational power, influence and control of communication. We derive our typology from a careful analysis of our data. It also draws on various theories on power in sociology literature.

Most definitions of power in the sociology literature (e.g., (Dahl, 1957; Emerson, 1962; Bierstedt, 1950)) include “an element indicating that power is the capability of one social actor to overcome resistance in achieving a desired objective or result” (Pfeffer, 1981). Influence closely resembles power, although some consider it as one of the means by which power is used (Handy, 1985). The five bases of power proposed by French and Raven (1959) (Coercive, Reward, Legitimate, Referent, and Expert) and its extensions are widely used in sociology to study power. We find these definitions and typologies helpful as general background, but not specific enough for a data-oriented study of how they are expressed in written dialogs such as emails. Wartenberg (1990) makes the distinction between

two notions of power: power-over and power-to. Power-over refers to hierarchical relationships between interactants, while power-to refers to the ability an interactant possesses and uses. Our notions of hierarchical power and situational power roughly correspond to Wartenberg’s notions of power-over and power-to, respectively. Both can be considered special cases of French & Raven’s notion of legitimate power. Our notion of Influence captures the forms of expert power and referent power introduced by French & Raven. Our notion for control of communication is based on the concept of conversational control introduced by Ng and Bradac (1993). We also capture the dynamic aspect of power through annotations for exercise of power and influence. Our utterance level annotations for overt display of power capture instances of “restriction of an interactant’s action-environment”, a key element to identify exercise of power in interactions (Wartenberg, 1990).

## 2. Corpus

We used the email corpus presented in (Hu et al., 2009) which was annotated with Dialog Functional Units (DFU). A DFU is a contiguous subset of an email which has a coherent communicative intention. The corpus contains 122 email threads with a total of 360 messages with each message segmented into a series of DFUs. This set of email threads is chosen from a version of the Enron email corpus with some missing messages restored from other emails in which they were quoted (Yeh and Harnly, 2006). Most emails are concerned with exchanging information, scheduling meetings, and solving problems, but there are also purely social emails.

Table 1 presents some statistics on participants and messages in the corpus.<sup>12</sup>

<sup>1</sup>Some messages did not have sender information. They were excluded from the *messages per active participant* statistic

<sup>2</sup>We define *active participant* of a given thread as someone who have sent at least one email message in the thread.

Statistic	Count / Mean (SD)
Number of email threads	122
Participants per thread	8.47(13.82)
Active participants per thread	1.81(0.73)
Messages per thread	2.95(2.24)
Messages per active participant	1.45(1.01)
Number of word tokens	20,740

Table 1: Corpus statistics

### 3. Annotations

In this section, we define annotations for different types of power. The annotations were performed by a single annotator. Table 2 and Table 3 show two example threads chosen from the annotated corpus. The first column shows the email thread and second column shows corresponding power annotations.

#### 3.1. Intention of Communication

We annotate the purpose or content type of the thread along with a very short description of the topic of the thread. The purpose could be one or more of the following.

1. Knowledge-Acquisition: The thread purpose is mainly to convey or exchange information.
2. Argumentation: The thread purpose is mainly to argue or explore the pros and cons of a position or claim.
3. External-event-planning: The thread purpose is to plan one or more events that will take place outside of the email exchange, such as a meeting, or performance of a task.
4. Collaboration-on-information-product: The thread purpose is to collaborate on a document or information, and the work will be done within the email communication channel.
5. Social: The main purpose of the thread is simply being social.

For the example thread in Table 2, our annotator identified the intention of the thread to be knowledge acquisition and that of the thread in Table 3 to be argumentation.

#### 3.2. Hierarchical Power

We annotate (person<sub>1</sub>, person<sub>2</sub>, reason) tuples such that person<sub>1</sub> appears to be above person<sub>2</sub> in the organizational hierarchy, based on the communication in the given thread. (The annotators do not have access to independent information about the organizational hierarchy.) Possible clues to hierarchical power include (by way of example): 1) a characteristic of a part of a message such as being an approval, or being a direct order; 2) a person's behavior such as asking for approval; 3) a person's authority to make the final decision. The hierarchical relation could also be explicitly or implicitly stated in one of the messages.

For the thread in Table 2, our annotator found *Rick Dietz* to have hierarchical power over *Alma Carrillo* since he mentions that *Alma* knows how to get in touch with him, suggesting that she is his assistant or secretary. For the thread

in Table 3, our annotator found no instances of hierarchical power relations.

#### 3.3. Situational Power

We annotate (person<sub>1</sub>, person<sub>2</sub>, reason) tuples such that based on the communication in the current thread, person<sub>1</sub> has power (authority to direct / approve other people's actions) in the current situation or while a particular task is being performed. Situational power is independent of organizational hierarchy: person<sub>1</sub> with situational power may or may not be above person<sub>2</sub> in the organizational hierarchy (or there may be no organizational hierarchy at all). The annotation for situational power is performed independently of the annotation for hierarchical power.

For the thread in Table 2, *Shelley Corman* is found to have situational power over *Rick Dietz* based on the fact that *Shelley* has a task (contracts) that she needs to perform and that she would normally involve *Rick* in. While in this case *Rick* is not actually working with her on the task (because of his golf game), it is his obligation to make sure someone else (presumably from his organization) can help *Shelley*. For the thread in Table 2, our annotator assessed no situational power relations.

#### 3.4. Attempt at Exercising Hierarchical or Situational Power

We annotate (utterances, person<sub>1</sub>, person<sub>2</sub>, type, successful?, reason) tuples where utterances denotes a contiguous sequence of utterances which indicate an attempt by some participant to exercise power. Person<sub>1</sub> denotes the participant associated with the utterance sequence. Person<sub>2</sub> denotes the participant being addressed. Type denotes the type of power i.e. Hierarchical or Situational. Successful? is one of {Y, N, NoReply} depending on whether the attempt was successful, not successful or not responded to. The reason is a list of observations that support the judgment.

Our annotator did not find any instances of attempts at exercising power in either example threads.

#### 3.5. Annotation for Influence

We annotate (person<sub>1</sub>, person<sub>2</sub>, reason) tuples where person<sub>1</sub> has influence and person<sub>2</sub> (if identifiable) is the person (or set of persons) on which person<sub>1</sub> has influence. The reason is the supporting evidence, ideally referring to specific parts of the thread where person<sub>1</sub> seem to have influence over person<sub>2</sub>.

We define a person to have influence if he or she:<sup>3</sup>

1. has credibility in the group;
2. persists in attempting to convince others, even if some disagreement occurs;

<sup>3</sup>We adopt this definition from the IARPA Socio-Cultural Content in Language (SCIL) program, where many of the researchers participating in the SCIL program contributed to the scope and refinement of the definition of a person with influence.

Example email thread	Annotations
<p>&lt;from name="Rick Dietz" &gt; &lt;to name="Shelley Corman" &gt; &lt;cc name="Alma Carrillo" &gt; ----- M1.1. Shelley, M1.2. I will be out of the office tomorrow morning participating in a charity golf tournament sponsored by EDS at South Shore Harbor. M1.3. Mark Giglotti, Jeannie Licciardo, Don Stacy and I are playing together. M1.4. I know the timing may be bad because of the financial project we have been preparing but the charitable contribution to play was quite generous and I do not want to back out of my commitment to the other team members. M1.5. HOWEVER, IF YOU NEED ME FOR ANY REASON, PLEASE PAGE ME AT 1(800) 609-6967 OR CALL MY CELL PHONE AT (713) 569-4140. M1.6. ALMA WILL ALSO KNOW HOW TO GET A HOLD OF ME. M1.7. I will only be 45 minutes away and will be able to come straight into the office, if needed. M1.8. I will check in during the day with Linda Trevino, as she will be sitting in for me at your staff meeting. M1.9. Rick</p> <p>&lt;from name="Shelley Corman" &gt; &lt;to name="Rick Dietz" &gt; ----- M2.1. Rick M2.2. I have a number of contracts that the bankers want early tomorrow. M2.3. I am assuming Elizabeth will be in to help me?</p> <p>&lt;from name="Rick Dietz" &gt; &lt;to name="Shelley Corman" &gt; ----- M3.1. Elizabeth will be in. M3.2. Also, Linda Trevino will be in the office. M3.3. All contracts are in Envision as well as in the file room on 39. M3.4. If you need me to be there, just let me know.</p>	<p><b>Annotation for Overt Show of Power:</b> N/A</p> <p><b>Annotation for Intention:</b> Knowledge-Acquisition: Rick out of the office to play golf for charity, and locations of contracts for bankers</p> <p><b>Annotation for Control:</b> Rick Dietz: Rick initiates the thread and tells Shelley Corman about his absence from work and where the contracts she needs are.</p> <p><b>Hierarchical Power Annotation:</b> Person.1: Rick Dietz Person.2: Alma Carrillo Reason: In M1.6, Rick tells Shelley Corman that Alma knows how to get in touch with him, suggesting that Alma is his assistant or secretary.</p> <p><b>Situational Power Annotation:</b> Person.1: Shelley Corman Person.2: Rick Dietz Reason: Rick is intimating Shelley about his absence and ensuring work is not affected while he is away. His messages have an apologetic tone, while Shelley's message are short and precise.</p> <p><b>Attempt at Exercising Power</b> <b>Annotation:</b> N/A</p> <p><b>Influence Annotation:</b> N/A</p> <p><b>Attempt at Influence Annotation:</b> N/A</p>

Table 2: Example thread and annotations

Example email thread	Annotations
<p>&lt;from name="Sara Shackleton" &gt; &lt;to name="Mark Taylor" &gt; ----- M1.1. Martin sent this message to London and Michael advised that Sullivan &amp; Cromwell be retained. M1.2. With respect to interest by the CFTC in the proposed transaction, would you agree that the same law firm advise on that issue as well? M1.3. Makes sense to me. M1.4. Also, what is Energydesk.com Limited? M1.5. Sara</p> <p>&lt;from name="Sara Shackleton" &gt; &lt;to name="Mark Taylor" &gt; ----- M2.1. Per my voice mail. M2.2. Let me know what you think. M2.3. SS</p> <p>&lt;from name="Mark Taylor" &gt; &lt;to name="Sara Shackleton" &gt; ----- M3.1. I think S&amp;C is fine - they are helping us with CFTC issues related to online trading M3.2. and Energy Desk.com seems somewhat related.</p> <p>&lt;from name="Sara Shackleton" &gt; &lt;to name="Martin Rosell" &gt; ----- M4.1. Martin: M4.2. Sorry for the log jam but I always thought that the law firm was the best idea. M4.3. Call if you need assistance. M4.4. Sara</p>	<p><b>Annotation for Overt Show of Power:</b> N/A</p> <p><b>Annotation for Intention:</b> Argumentation: belief that Sullivan &amp; Cromwell should be retained</p> <p><b>Annotation for Control:</b> Sara Shackleton: Sara initiates the thread and asks Mark Taylor for his opinion on retaining Sullivan &amp; Cromwell.</p> <p><b>Hierarchical Power Annotation:</b> N/A</p> <p><b>Situational Power Annotation:</b> N/A</p> <p><b>Attempt at Exercising Power Annotation:</b> N/A</p> <p><b>Influence Annotation:</b> N/A Person_1: Mark Taylor Person_2: Sara Shackleton Reason: In message 1, Sara asks Mark for his advice on using a law firm for a second time, and after he expresses his view in message 3, she concurs in message 4, to Martin Rosell.</p> <p><b>Attempt at Influence Annotation:</b> DFUs: M3.1 Person_1: Mark Taylor Person_2: Sara Shackleton Successful?: Y Reason: After Sara asks for Mark's advice about using a law firm, he tells her his opinion in message 3, and in message 4, she agrees with him in her message to Martin Rosell.</p>

Table 3: Example thread and annotations

3. introduces topics/ideas that others pick up on or support;
4. and is a group participant but not necessarily active in the discussion(s) where others support/credit him or her.

In addition, the influencer's ideas or language may be adopted by others and others may explicitly recognize influencer's authority.

Our annotator did not identify any influence relations in the

thread in Table 2. For the thread in Table 3, our annotator found *Mark Taylor* to have influence over *Sara Shackleton*, since *Sara* asks *Mark* for his advice, and follows it.

### 3.6. Attempt at Influence

We identify (utterances, person\_1, person\_2, successful?, reason), where utterances denotes a contiguous sequence of utterances which include an attempt by some participant to influence other participants. The relevant notion of "influence" is the definition given for person with influence in the previous section. Person\_1 denotes the participant

associated with the utterance sequence. *Person\_2* denotes the participant being addressed. *Successful?* is one of {Y, N, NoReply} depending on whether the attempt was successful, not successful or not responded to. The reason is a list of observations that support the judgment.

Our annotator did not identify any attempts at influence in the thread in Table 2. In the example thread in Table 3, our annotator found *Mark Taylor* to have attempted to influence *Sara Shackleton* in the DFU M3.1 and was successful.

### 3.7. Controller of Communication

We also identify participants who actively attempt to achieve the intended goals of the communication. These are people who ask questions, request others to take action, etc. and not people who simply respond to questions or perform actions when directed to do so. There could be multiple such participants in a given thread.

Our annotator found *Rick Dietz* and *Sara Shackleton* to be in control for the threads in Table 2 and Table 3, respectively.

### 3.8. Overt Display of Power Annotation

In addition to thread level power annotations, we also annotate each utterance to specify whether there was an Overt display of power (*ODP*) or not. An utterer can choose linguistic forms in his or her utterance to signal that he or she is imposing constraints on the addressee’s choice of how to respond, which go beyond those defined by the standard set of dialog acts. For example, if the boss’s email is “Please come to my office right now”, and the addressee declines, he is clearly not adhering to the constraints the boss has signaled, though he is adhering to the general constraints of cooperative dialog by responding to the request for action. We are interested in these additional constraints imposed on utterances through choices in linguistic form. We define an utterance to have *ODP* if it is interpreted as creating additional constraints on the response beyond those imposed by the general dialog act. An *ODP* can be an order, command, question or even a declarative sentence. The presence of an *ODP* does not presuppose that the utterer actually possess social power: the utterer could be attempting to gain power.

Table 4 presents some sample utterances chosen from our corpus. An utterance with *ODP* could be an explicit order or command (*s3*, *s8*) or an implicit one (*s2*, *s5*). It could be a simple sentence (*s3*) or a complex one (*s1*). It could be an imperative (*s3*), an interrogative (*s5*) or even a declarative (*s2*) sentence. But not all imperatives (*s4*) or interrogatives (*s6*, *s7*) are *ODPs*. *s5*, *s6* and *s7* are all syntactically structured as a question. However, *s5*’s discourse function within an email is to request/order to work on “that” which makes it an instance of *ODP*, while *s6* is merely an inquiry and *s7* is a rhetorical question.

For a further discussion of the annotation of *ODP*, see (Prabhakaran et al., 2012), where we further define the notion of *ODP*, give inter-annotator agreement numbers, and

ID	Sample utterance
s1	If there is any movement of these people between groups can you please keep me in the loop.
s2	I need the answer ASAP, as ....
s3	Please give me your views ASAP.
s4*	Enjoy the rest of your week!
s5	Would you work on that?
s6*	... would you agree that the same law firm advise on that issue as well?
s7*	can you BELIEVE this bloody election?
s8	ok call me on my cell later.

Table 4: Sample utterances from the corpus  
\* next to ID denotes an utterance without an *ODP*

present initial work on building an automatic classifier for *ODP*.

## 4. Additional Annotation Instructions

Apart from definitions for annotations described in Section 3., we gave our annotators the below instructions.

- Consider only people present in at least one of from/to/cc fields and specify the person names in annotations as they appear in the from/to/cc fields.
- If either person\_1 or person\_2 of an annotation seems to be all persons in To/CC/both field(s) of a message, represent them as To(n), CC(n) or ToCC(n) respectively, where n is the message identifier.
- Annotations for situational power, hierarchical power and influence should be made independent to each other.
- Annotation for an attempt at power/influence between two participants should be done independent of whether they were judged to have that relation.
- Annotations should be made based solely on the communication within the thread, disregarding annotator’s world knowledge about any participant or knowledge about participant relations from previously annotated threads.

## 5. Annotation Statistics

This section presents some statistics on the annotated corpus of 122 threads.

Annotations	Total Count
Hierarchical power	26
Hierarchical power attempt	8
Situational power	390
Situational power attempt	282
Influence	16
Influence attempt	77
Control of communication	128
Overt display of power	86

Table 5: Annotation statistics

Table 5 presents counts of each type of annotations in the corpus. Six threads were annotated to have two persons in control while all others have exactly one person in control (no threads have no one in control). Out of 1,734 utterances, only 86 (5.2%) were judged to have an overt display of power. Table 6 presents the percentage breakdown on success/failure/no response for attempts at exercising each type of power or influence.

Relation	Count	Y	N	No Reply
Hierarchical	8	50.0	0.0	50.0
Situational	282	21.3	11.7	67.0
Influence	77	18.2	9.1	72.7

Table 6: Power/Influence attempt and success statistics

Table 7 gives the counts of each type of intention annotation on the 122 email threads. There were 24 threads which were annotated to have two intentions.

Intention	Count
External-event-planning	30
Collaboration-on-information-product	21
Knowledge-Acquisition	70
Social	5
Argumentation	20

Table 7: Intention statistics

## 6. Conclusion

We have presented an email corpus annotated with four types of social power relations between participants — hierarchical power, situational power, influence and control of communication.<sup>4</sup> Annotations also capture attempts at exercise of power and influence, as well as utterances with overt display of power. As future work, we plan to extend these annotations to other genres (discussion forums, blogs), domains (political, scientific) and languages (for example, Arabic).

## 7. Acknowledgments

This work is supported, in part, by the Johns Hopkins Human Language Technology Center of Excellence. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the sponsor. While working on this project, Huzaifa Neralwala was affiliated with the Center for Computational Learning Systems at Columbia University. We thank several anonymous reviewers for their constructive feedback.

## 8. References

Robert Bierstedt. 1950. An Analysis of Social Power. *American Sociological Review*.  
 Philip Bramsen, Martha Escobar-Molano, Ami Patel, and Rafael Alonso. 2011. Extracting social power relationships from natural language. In *ACL*, pages 773–782. The Association for Computer Linguistics.

Germán Creamer, Ryan Rowe, Shlomo Hershkop, and Salvatore J. Stolfo. 2009. Advances in web mining and web usage analysis. chapter Segmentation and Automated Social Hierarchy Detection through Email Network Analysis, pages 40–58. Springer-Verlag, Berlin, Heidelberg.

R. A. Dahl. 1957. The concept of power. *Syst. Res.*, 2(3):201–215.

Jana Diesner and Kathleen M. Carley. 2005. Exploration of communication networks from the enron email corpus. In *In Proc. of Workshop on Link Analysis, Counterterrorism and Security, SIAM International Conference on Data Mining 2005*, pages 21–23.

Richard M. Emerson. 1962. Power-Dependence Relations. *American Sociological Review*, 27(1):31–41.

John R. French and Bertram Raven. 1959. The Bases of Social Power. In Dorwin Cartwright, editor, *Studies in Social Power*, pages 150–167+. University of Michigan Press.

C.B. Handy. 1985. *Understanding Organisations*. Institute of Purchasing & Supply.

Jun Hu, Rebecca Passonneau, and Owen Rambow. 2009. Contrasting the interaction structure of an email and a telephone corpus: A machine learning approach to annotation of dialogue function units. In *Proceedings of the SIGDIAL 2009 Conference*, London, UK, September. Association for Computational Linguistics.

Sik Hung Ng and James J. Bradac. 1993. *Power in language : verbal communication and social influence / Sik Hung Ng, James J. Bradac*. Sage Publications, Newbury Park .

Kelly Peterson, Matt Hohensee, and Fei Xia. 2011. Email formality in the workplace: A case study on the enron corpus. In *Proceedings of the Workshop on Language in Social Media (LSM 2011)*, pages 86–95, Portland, Oregon, June. Association for Computational Linguistics.

Jeffrey Pfeffer. 1981. *Power in organizations*. Pitman, Marshfield, MA.

Vinodkumar Prabhakaran, Owen Rambow, and Mona Diab. 2012. Predicting overt display of power in written dialogs. In *Human Language Technologies: The 2012 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, Montreal, Canada, June. Association for Computational Linguistics.

T.E. Wartenberg. 1990. *The forms of power: from domination to transformation*. Temple University Press.

Jen-yuan Yeh and Aaron Harnly. 2006. Email thread reassembly using similarity matching. In *In Proc. of CEAS*.

<sup>4</sup>The annotation manual and the annotated corpus can be obtained at <http://www.cs.columbia.edu/~vinod/powerann/>