Identifying Stylometric Correlates Of Social Power

By:

Rachel Emma Cotterill

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

The University of Sheffield
Faculty of Engineering
Department of Computer Science

February 2016

Contents

Index of Figures

Index of Tables

Abstract

This thesis takes a stylometric approach to the measurement of social power, particularly hierarchical power in an organisational setting. Following the social constructionist view of identity, we infer that construction of identity is an ongoing process incorporating the full scope of human behaviour, including linguistic behaviour. We test the primary hypothesis that stylistic choice in language is indicative of power relations, and that a stylometric signal can be extracted from natural language to enable prediction of relationship status. Additionally, we consider the effect of individual variation versus interpersonal variation, and the effects of aggregating predictions to boost the predictive strength of the model. Three different datasets are used to validate the proposed approach across three different genres: email, spoken conversation, and online chat. We also present a vector space approach to modelling linguistic style accommodation, and undertake a preliminary examination of the correlation between linguistic accommodation and social power.

Publications Arising

Cotterill, R. (2011) Question classification for email. *Proceedings of the Ninth International Conference on Computational Semantics*, 330-334. Association for Computational Linguistics.

Cotterill, R. (2013) Using Stylistic Features for Social Power Modelling. *Computación y Sistemas*, 17(2), 219-227.

Jones, S., Cotterill, R., Dewdney, N., Muir, K. and Joinson, A. (2014) Finding Zelig in text: A measure for normalising linguistic accommodation. *Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics*, Dublin, Ireland, August 2014.

Muir, K., Cotterill, R., Joinson, A. and Dewdney, N. (2015) Power and personality in linguistic style accommodation. *British Psychological Society Developmental and Social Section Annual Conference*, Manchester, UK, 9 - 11 September 2015.

Cotterill, R., Muir, K., Joinson, A. and Dewdney, N. (2015) Identifying linguistic correlates of power. *International Journal of Computational Linguistics and Applications*, 6(1).

Muir, K., Joinson, A., Cotterill, R. and Dewdney, N. (2016) Characterising the linguistic chameleon: Personal and social correlates of linguistic style accommodation. *Human Communication Research*, 42(3).

Muir, K., Joinson, A., Cotterill, R. and Dewdney, N. (2016) When Communication Accommodation Backfires: Interpersonal Effects of Social Power and Linguistic Style Accommodation in Computer-Mediated-Communication. *Proceedings of the 66th Annual Conference of the International Communication Association*.

Muir, K., Joinson, A., Cotterill, R. and Dewdney, N. (2016) English Speed Networking Conversational Transcripts LDC2016T16. *Linguistic Data Consortium.*

1. Introduction

Social identity theory posits that 'identity', rather than being a set of static features possessed by an individual, is a social construct that is constantly negotiated in the course of interactions with others (Tajfel & Turner 1979). Similarly, various theories have been proposed to account for politeness and formality in language (e.g. Goffman 1959; Grice 1975; Brown & Levinson 1987), but theorists generally agree that in order to be polite it is necessary to communicate in a manner appropriate to your audience, and that this changes according to circumstance (Goffman 1974).

If an individual is constantly adjusting their communicative style in this way, to adapt to the requirements of the context in which they find themselves, it follows that this will result in observable phenomena. It should be possible to extract a signal from the individual's changing use of language, from which we can (given sufficient data) infer something of the context of each communication. However this is an inherently noisy signal, as relationship negotiation is only one aspect of the purpose of any given communication, and multiple relationship categories may be constructed in parallel.

Traditional n-gram approaches to language modelling, as originally proposed by Weaver (1955) for machine translation, have subsequently been adopted for use in other language processing tasks. While often successful, n-gram models operate at the surface level of the language. Influences of topic and style are conflated into a single model, as an n-gram model captures both semantic content (via specific lexical choice) and stylistic content (e.g. word order, sentence complexity, function word distribution). Since it is possible to discuss varied topics with the same audience, or the same topic with different groups, we hypothesise that it would be worthwhile to draw a distinction between these two aspects. You may discuss your work (or your weekend plans) with your boss, your colleague, or your spouse, therefore we predict that stylistic choices will likely be more indicative of relationship than the selection of a specific topic, although topic distribution may itself be informative. Additionally, it is easier for individuals to deliberately adjust their language at the level of lexical choice and word order: some stylistic features, such as function word distribution, are subconscious and therefore any resulting signal is more likely to be genuine.

This thesis sets out to test the following primary hypothesis:

> **H1:** Stylistic choice is a key method of expressing relationship-building, therefore relationship information can be inferred from linguistic style.

This depends on the assumption that it is possible to meaningfully separate topical from stylistic features in language. Topical features are those with explicit semantic content, in particular lexical choice (nouns, verbs), while stylistic features can be more freely varied to affect the manner of communication (punctuation, function word distribution). The question of whether these two aspects can be separated, or to what extent they are naturally correlated, is something that linguists are only beginning to explore (e.g. Argamon et al. 2007; van der Wees et al. 2015); we shall explore this issue in more detail at a later stage.

We additionally set out to test the idea that when it comes to stylistic choice, differences from the individual norm will be more informative than absolute values. If this hypothesis is true, per-author standardisation of feature scores should result in higher classification accuracy. We record this as a secondary hypothesis:

> **H2:** Differences from the individual's normal behaviour will be more informative than absolute feature scores, for predicting relationships.

When considering the classification of relationships, it is possible to operate at various levels: at the message level, using only a single piece of content as evidence; at the relationship level, using an aggregation of all messages exchanged between a pair; or at the community level by taking into account the shape of the resulting graph of power predictions. As each of these levels makes additional information available to the task, we might reasonably conjecture that performance will improve with each stage. This gives us an additional two hypotheses to test:

> **H3:** Prediction of hierarchical relationships will be more effective at the pairwise level than at the level of individual messages.

> **H4:** Relationship classification in a hierarchical situation will be aided by consideration of whole-network characteristics.

Finally, we consider social power as a driver for linguistic style accommodation, and test the following:

> **H5:** Linguistic accommodation behaviour is partially motivated by relative social power, therefore greater accommodation of linguistic style will correlate with lower social power.

In order to test our hypotheses, we use data from three sources: email data from the Enron corpus, and two sets of dialogues (one face-to-face, and one computer mediated) generated in an experimental setting by giving participants a speed networking task to complete, with controlled discussion topics and role-based power manipulation. In all cases we examine the organisational hierarchy, as this is a category of relationship which imparts comparatively well-defined imbalances in social power, and we consider the impact of hierarchical position on language choices. If we are able to reconstruct a graph of social hierarchy, using stylistic features with a minimum of interference from topical content, this will verify H1 as it applies to the organisational sphere. (However, a failure to do so would not constitute proof that H1 is false, only that our attempt has been insufficient.)

1.1. Potential Applications

There are a variety of reasons for taking an interest in the social structure of a group, whether in a virtual or physical environment. Marketing experts have become very interested in identifying key individuals with influence in a network, in order to target their products accordingly. At the opposite end of the spectrum, investigators in large-scale criminal proceedings (such as the Enron scandal) need to construct detailed social network graphs to develop an understanding of the relationships within the community under investigation.

As email and other online communications becomes more pervasive, and users are encouraged by providers such as Gmail to archive rather than delete old emails, there is now a considerable quantity of information to be gleaned from datasets too large for a human to ever read. This is not often in the form of simple statements of fact such as those targeted by traditional information retrieval (IR) techniques. Indeed, a lot of personal, interpersonal, and social information is never explicitly stated; rather it exists as part of the mutually understood common ground, or is negotiated via non-explicit exchanges.

By studying the way this information is encoded, we may be able to build tools that can present even this implicit information. Relationship information in particular may be of use in mailbox management applications (perhaps to automatically flag messages from a user's senior colleagues), or to enable the construction of detailed social network graphs in police and fraud investigations.

1.2. Structure of this Thesis

Chapter 2 gives an overview of the theoretical background and related literature, while Chapter 3 outlines our methodological approach and introduces the three datasets. Chapters 4 gives message level classification results, Chapter 5 looks at the options for improving on these scores using pairwise aggregation and network-level techniques, and Chapter 6 sets out preliminary work on a method to measure linguistic accommodation using vector techniques. Chapter 7 looks at the feature distributions of our data in more detail, and Chapter 8 summarises our findings and sets out some proposals for future work.

Work in Chapter 6, and the data collection described in sections 3.1.2 and 3.1.3, was undertaken in collaboration with Kate Muir, Adam Joinson, Simon Jones (University of the West of England), and Nigel Dewdney (University of Sheffield).

2. Theoretical Background

This work is grounded in the theories of socially-constructed identity (Tajfel & Turner 1979; Turner et al. 1987) and framing (Goffman 1974). These psychological and sociological theories account for social interaction as an ongoing negotiation between the participating parties, encompassing the participants' self-identification, personal relationships, and social roles.

Our approach is informed by stylometry (Lutosławski 1897), linguistic politeness (Grice 1975; Brown & Levinson 1987), and communication accommodation theory (Giles 1973). Using a tightly constrained set of stylistic features has a potential advantage over n-gram models, leading to much smaller models with more explanatory power.

The task of categorising relationships and reconstructing social hierarchies also builds conceptually upon previous work in social network analysis (SNA) of communications graphs (Rowe et al. 2007; Gallagher 2010).

We will now briefly explore these contributing fields.

2.1. Social Constructs of Identity, Framing, and Politeness

The social approach to identity proposes that, rather than being something fixed and personal, an 'identity' is flexibly and socially constructed with reference to the communities in which an individual is engaged (Tajfel & Turner 1979). This captures a lot of intuition: the same person will tend to behave differently with colleagues at work, with friends in the bar, or when visiting elderly relatives. The social identity approach also interfaces neatly with the concept of framing (Goffman 1974): not only do people behave differently with different communities, they behave appropriately to the situation in which they find themselves. If you take your grandmother to the bar, you must negotiate a suitable path between grandma-appropriate and bar-appropriate behaviour.

Although theories of interconnectedness can be traced back through philosophy as far as early buddhism, Cooley (1902) is one of the first modern sociologists to explicitly lay out a theory of interdependence between individuals and the communities they comprise, noting that "a separate individual is an abstraction unknown to experience, and so likewise is society when regarded as something apart from individuals." This theory is extended by later contributors such as Mead (1934) and Blumer (1986) in the literature of symbolic interaction, the fundamental premise of which is that people respond to their environment (including other people) according to the meaning they find there, and that this meaning is itself a product of communal negotiation.

One particular source of social 'meaning' is that derived from individuals' respective membership of social groups, an aspect which has been explored in depth by later theorists. Tajfel & Turner (1979) initially proposed social identity theory to complement Campbell's (1965) realistic group conflict theory, which in turn was put forward as an explanation of conflict between groups. Social identity theory draws a distinction between *interpersonal* and *intergroup* dynamics: the interpersonal element of an interaction is determined by two individuals and their specific relationship, while the intergroup element is conditioned by their respective social positions. It is this second component which wraps up the effects of stereotyping, in- and out-group behaviour, and related phenomena. In general, intergroup effects will be more salient between strangers, and as a relationship becomes closer, interpersonal elements will come to dominate.

Although initially set out to account for conflict between groups, the fundamental tenets of social identity theory are universally applicable to the building and maintenance of all human relationships. One consequence is that an individual's choices of presentation will vary depending on the group(s) in which she finds herself at the time, and whether she considers herself to be an insider or an outsider to this group. The theory of self-categorisation (Turner et al. 1987; Turner et al. 1994) builds on this idea, examining the changing salience of different group identities in different contexts. The likelihood of a particular self-category becoming salient depends on the level to which the individual identifies with that group, and the relevance of that category to the current scenario. Experiments such as those conducted by Onorato & Turner (2004) have shown that this can be manipulated, with deliberate priming used to affect which group memberships are activated during a given exchange; this flexibility to immediate context is one of the major differentiators of social identity theory from the earlier concepts of identity theory in sociology (Hogg et al. 1995).

It is worth noting that a 'group' in Turner's sense does not have to consist of a specific set of individuals who regularly (if ever) interact. Take, for example, a sole Christian family living in a largely Muslim community, who may at times (perhaps at Christmas, or during Ramadan) wish to position themselves as part of the wider group of Christians, in contrast to their Muslim neighbours. Nor must the groups to which an individual 'belongs' be considered mutually exclusive: there will be plenty of other times when it is more appropriate for our imaginary Christians to position themselves as simply members of the local community along with their neighbours, irrespective of religion, and in contrast to those who live elsewhere.

Along with the individual relationships between people, their respective group memberships — and the salience of these groups in the current interaction context — will impact on the way that interactions are realised linguistically. For example, Hogg & Reid (2006) reflect on the role of linguistic choice as used by leaders to cement their position within a group, and to shift group norms to better reinforce their own power.

The social setting, however, comprises more than just people. The concept of a 'frame' of interaction, as proposed by Goffman (1974), is a way of characterising the general context in which an interaction takes place. The idea of a frame is that different contexts of interaction require different things of the participants. For example, a restaurant frame would naturally include ordering items from a menu, eating food (probably served as multiple courses), conversing with one's companions, and paying the eventual bill. Different phases of this scenario place different requirements on the participants. Similarly, within the restaurant frame, the roles of diner and waiter require different typical behaviours — even though the same person could be a diner one day, and a waiter the next, in different contexts.

The interaction between frames and social groups is a subtle one. The example of taking one's grandmother to a lively bar is an extreme case which highlights some of the issues in a cartoon manner. But everyone will have personal examples of going to the pub with colleagues, or a spouse dropping in at work for a quick chat, or the waiter turning out to be a long-lost schoolfriend: times when a person from an unexpected group is involved in a frame where they are not usually seen.

When encountering an individual or group, in a frame in which they do not usually belong, we must choose to what extent we continue to fulfil the frame's generic expectations and to what extent we adapt our behaviour to the specific individuals concerned. To continue with the restaurant example, upon recognising our friend we would probably break out of the 'restaurant' frame to engage in a 'reunion' frame for a time. But if we are to succeed in our goals of eating a meal out, we will have to return to 'restaurant' mode to place our orders, and our waiter-friend will doubtless need to attend to other tables if they are to keep their job.

From this we can see that framing and grouping are not independent: expected groups are a part of the frame's expectation-setting. Violation of these group-related expectations can be jarring, just as we would be taken aback by other frame-related violations such as a restaurant's failure to supply food and drinks.

Selecting an appropriate level of politeness is one significant way in which language can be adjusted to be a more appropriate style for a given context, without materially affecting the topic or purpose of the communication.

The literature of politeness in language starts with Goffman's (1959) notion of 'face', and the theory that everyone chooses to present a particular version of themselves to society. This projected image changes depending on the circumstances of the interaction, and we can see that the concepts of frame and face are related, as the face one chooses to project is affected by the requirements of the current frame. In an individual's choice of language, they are determining what image to project, but they must also consider the face of those they interact with if communication is to succeed. By being more polite, the speaker reduces the face threat to her interlocutor.

In their seminal work on politeness, Brown & Levinson (1987) build on the concept of face, defining a face-threatening act as a social action which imposes on the recipient. Making a request is the prototypical face-threatening act, but any attempt to communicate carries a tacit request for the audience's attention, and therefore places a demand on their time and cognitive effort. The level of imposition this entails depends on the circumstances. Politeness strategies are used as mitigation, to reduce the face threat when performing such an action.

Brown & Levinson identify three features which should impact on the language choices made when making a request or taking any other face-threatening action:

— the degree of imposition

— the symmetric relation: the social distance between the participants (how well do they know one another?)
— the asymmetric relation: the power balance in the relationship, derived from the relative status of the participants (whether this status is professional or societal in origin)

The interaction of these features affects the level of politeness required to compensate for the threat to face. To give an example of extremes, making a significant request of a stranger requires more politeness than making a small request of a close friend. Strategies for increasing politeness include making a request indirectly ("off the record"), or padding it with expressions of positive politeness (such as saying 'please' or using formal address forms) or negative politeness (such as hedging or apologising). These strategies can be combined in a variety of ways, giving speakers the flexibility to express varying levels of politeness.

The level of politeness appropriate to a particular communication is contextually dependent, and more is not always better: being excessively polite can be just as inappropriate as being impolite when politeness is required (Locher 2004; Watts 2005). It sounds like a paradox to propose that it can be impolite to be too polite; it appears from this that we are overloading the concept of 'politeness'. We must distinguish the process of "employing politeness strategies" from the end result of "being polite".

These various theories all express different facets of the same basic principle: that individuals moderate their behaviour (and, consequently, language) according to context. This context is comprised of various aspects including physical environment, social groups, social roles, and contextual expectations. It follows that linguistic choices will be conditioned by some combination of these factors.

2.2. Factors Affecting Linguistic Choice

We have briefly examined the social theory of identity, and the concept of framing. Having thus established that people adapt their behaviour depending on the groups to which they belong and with whom they interact, we then come to consider how this may be expressed through their linguistic selections. There is a vast literature of conversational analysis spanning various disciplines, most of which we will not repeat here; Eggins & Slade (1997) give a good overview of the contributing disciplines.

If behaviour is generally informed by contextual factors, and language is a form of behaviour, it follows that the choices made in the course of linguistic production will likewise be influenced by a combination of contextual elements. These factors may be intrinsic to the speaker or extrinsic; they may be static or ever-changing; they may exert greater or lesser influence, independently or in combination.

Briefly, the factors affecting linguistic choice can be summarised by the following questions: who is speaking, to whom, with what purpose, and in what context?

Intrinsic aspects of the speaker which may prove salient include psychological traits, emotional state, gender, age, social background, native language, and level of education. For the purposes of most studies, these are treated as inherent attributes of an individual and assumed to be, if not absolutely unchanging, at least stable for the duration of the study in question. In practice it is evident that very little about an individual is static: emotional state changes quickly, education is a lifelong process, and even age is changing at a constant (if slow) rate.

The intended audience, and the speaker's relationship to them, is one of the most important aspects of the communicative context. The closer the speaker and audience, the more common ground they might expect to share – contrast two strangers who must first work out whether they share a common language, with twins who have grown up together and share a thousand little inside jokes.

The purpose of the conversation will determine the frame of the communication, or else be determined by the demands of an existing frame. Different modes and purposes of communication are distinguished by authors and readers alike as resulting in a 'style', defined by Karlgren (2010) as a "consistent and distinguishable tendency to make [certain] linguistic choices". Casual small talk differs in measurable ways from the focused exchange of a meeting, and even within similar contexts, the content of the specific conversation may affect the tone. One commonly studied example of this is the case of requests: a request is an imposition, and therefore, requests tend to be more polite than equivalent non-requesting communications, and requests of higher imposition require more politeness (Duthler 2006, Peterson et al. 2011).

Linguistic choices may also be influenced by the environment in which communication takes place, whether this is physical or virtual (noting that even electronic communications are composed in a physical environment). Choosing an email over a phone call gives the author scope to edit their words before hitting 'send', but there's a difference between sitting down comfortably at your desk, and trying to fire off a quick message on your phone while running late for work.

It should be obvious that these various categories are not independent. When you set out to communicate, you will choose who to speak with depending on what you wish to achieve; how you feel about them will impact your emotional state during the conversation; and how you contact them may depend upon the nature of your relationship or on your communicative goals. For simplicity, however, these aspects are usually studied separately.

We will now take a brief look at previous research which has investigated each of these categories in a little more depth.

2.2.1. The Speaker's Personal Attributes

Studies have considered correlations between linguistic selections and a range of personal attributes such as gender (Argamon et al. 2007), age (Rosenthal & McKeown 2011), and personality (Oberlander & Gill 2006). It should be obvious that different individuals tend to have different preferred styles, and that in some cases their stylistic choices might be attributable to a specific aspect of their individual circumstances: for example, a primary school teacher who has become adept at replacing 'shit' with 'sugar' in a work context, and who continues to do so at home. More often than not, however, there are numerous linguistic communities exerting their influence on every speaker, at any given time, and the patterns of difference become apparent only when examining the data at population-level.

A lot of work in this field originates in forensics (see Grant (2007) for an overview), with the requirement to identify or characterise individuals based on the evidence of their language, such as in cases where anonymous threats are made. The forensic scenario has the added complication that the author may be trying to obfuscate or disguise his style, which we assume is not relevant to non-criminal situations. Authorship identification has also been a topic of some interest in literary and historical circles, where related techniques are used to attribute (or deny) long-lost manuscripts to famous writers, dating back to Mosteller & Wallace's (1963) undertaking to assign the contested Federalist Papers to their possible authors by way of function words. These techniques tend to produce reliable results only when approached as closed-class categorisation tasks, and as the interest is in specific identity rather than attributes, all features are generally taken together to model an individual author's style without regard for intersecting categories of social identity (Stamatatos 2009). More recent work, however, has examined the linguistic characteristics of a number of facets of individual identity (Argamon et al. 2007; Nguyen et al. 2014; Rosenthal & McKeown 2011).

Gender

There have been numerous studies of gender differences in language use, dating back to the early days of sociolinguistcs; automated classification by gender begins with Koppel et al's (2002) study, in which they report classification accuracy above 80% on documents from the British National Corpus, using a combination of function words, part-of-speech n-grams, and punctuation features.

In their blog study, Argamon et al. (2007) examine 140 million words of text and find that articles and prepositions are significantly associated with male language, while women use more personal pronouns, conjunctions, and auxiliaries. They also find topical distinctions, and observe a lack of independence between topical and stylistic features. In this corpus women use proportionally more swearwords, which is at odds with the general perception of swearing, as well as contrasting with Rayson et al.'s (1997) older study of distributions in the British National Corpus. Another interesting finding was the similarity between age-signifying and gender-signifying words: in general, the words more commonly used by women were also the words more commonly used by younger bloggers. These observations could be an artefact of the way in which men and women use blogs: within this corpus, it was observed that women and younger bloggers tended to focus on personal topics, while men and older bloggers tended towards an external focus and more news-like writing.

More recently, Twitter and Facebook have become popular sources of informal textual data. Nguyen et al. (2014) gathered data for their TweetGenie system by setting up their experiment in the form of an online tool to predict gender and age from Twitter streams, and encouraging the general public to provide feedback on the accuracy of results. Their accuracy of gender predictions for Dutch was assessed at 94% based on the user-submitted confirmations (or corrections), and 82% based on a random sample of 50 users whose profiles have been manually annotated by the research team. Sap et al. (2014) create a weighted lexicon of gender-signifying words using regression and classification models. They report classification accuracy of 91.7% on Facebook messages and 85.6% on Twitter data, using lexicons derived from Facebook data; the inverse, using a (smaller) Twitter dataset to train the lexicon, resulted in classification accuracy of 88.9% on Twitter data and 81.6% on Facebook messages. A combined lexicon derived from Facebook, blog, and Twitter data gave the best result overall, with accuracy of 91.9% for Facebook and 90.0% for Twitter.

Bamman et al.'s (2012) investigation of Twitter data stands out for providing a more nuanced analysis than most. In particular, they identify that by looking at language exclusively through the lens of gender, researchers are limiting themselves to findings that directly support or oppose the fundamental hypothesis of binary gendered language. By clustering tweets according to content, and then examining gender distribution within the resulting clusters, they are able to shed light on the distribution of what would appear as 'outliers' in a traditional categorisation model. In particular, their findings regarding interactions within single-sex and mixed-sex groups highlight that, rather than being randomly distributed, users of less-typically-gendered language tend to interact in more heterogenous social groups. This supports a social constructionist view of gender, and meshes neatly with the social theory of identity in general, giving rise to an interpretation where 'gender' is just another group to identify with, the salience of which may wax and wane depending on circumstance. For comparison with other studies, they report accuracy of 88% using their clusters to inform a binary classification task.

Flekova et al. (2016) take a different approach to the question of classifying gender, undertaking an experiment to gather human judgement of gender from tweets, and examining the systematic biases which lead people to make the wrong judgements. They found that individual human performance had an overall accuracy of 75.5% with pairwise inter-annotator agreement of 70.0%, which is an interesting result to consider alongside the performance of automated systems described above. They examine textual features, using both unigram and topic-based models, to highlight those which correlate with biased judgement: annotators were shown to have given undue weight to topics such as household, celebration, and family (presumed female), and sports, technology, and politics (presumed male).

Age

Research into age-related language use (such as Rosenthal & McKeown 2011; Argamon et al. 2007) is typically muddled by a lack of distinction between features of *current age* and features of *generation*. For example, some topic-related lexical features are likely to be a function of absolute age: students are the most likely to talk about school, whereas discussion of marriage, mortgages, and children would tend to be more prevalent amongst people who are themselves buying houses or starting a family. These are related to the frames which are most typical for individuals of a given age. Other features, such as informal slang, are more likely to belong to a generational band and move up the age range as that generation grows older, while continuing to use (some proportion of) the popular vocabulary of their youth. Under standard unigram and n-gram models, these two aspects will be conflated, unless text is taken from the same generation of individuals when they are at different ages.

As this requires consistent data collection over a long period, such longitudinal studies are rare, and seldom used for linguistic research. One exception is Pennebaker & Stone's (2003) study of authors and playwrights, whose progression of published works provides a natural (if unconventional) dataset. They report that with increasing age individuals tend to use more positive language, fewer negative words, and fewer first-person singular pronouns.

Rosenthal & McKeown (2011) categorise LiveJournal blog posts according to the date of birth of the writer. They examine both stylistic and content (topic) features, finding that features such as emoticons and slang are more prevalent among younger writers, and also demonstrating predictive power from a bag-of-words model. Although they do not explicitly separate out the effects of current age versus generation, some stage-of-life features do come out clearly in the significant n-grams, for example in references to housing and school.

A model for language change in a somewhat generational style, exhibited over a shorter time period, can be seen in Danescu-Niculescu-Mizil et al.'s (2013b) study of language change in online forums. They find that in the case of synonyms gaining and losing popularity, forum users tend to prefer the terms that were standard when they joined the community, and gradually stop adapting to changing community norms. This might reasonably be extrapolated to describe linguistic development in general: most adults don't continually adapt to use the latest teenage slang.

Argamon et al. (2007) conduct an extensive analysis of blog posts, with factor analysis of the most common words, and found distinctions between age bands in both topic and style (function) words.

Personality and Mental State

Oberlander & Gill (2006) examine the correlation between linguistic choices and psychological traits such as extraversion and neuroticism, both of which have previously been shown to affect language use (Pennebaker & King, 1999; Dewaele & Furnham, 1999; Argamon et al, 2005) and which are assumed to be comparatively stable. They test a number of hypotheses relating to personality and language by eliciting an email corpus from 105 volunteers, who were also asked to take a three-factor personality test (scoring for extraversion, neuroticism, and psychoticism). Senders were stratified according to their personality profiles, to exclude those who were extreme on more than one dimension, and text from each stratum was analysed using both lexical and part-of-speech n-grams to identify distinctive features. They report that "Extraversion is associated with linguistic features involving fluency, positivity and implicitness; Neuroticism with self-concern, negativity and implicitness; and Psychoticism with creativity and detachment."

Nowson & Oberlander (2007) examine the question of inferring personality traits from blog text, and compare the results from two corpora. Their small, 'clean' corpus is labelled with personality data based on asking participants to complete a full OCEAN personality inventory, while for contrast they also collected a large, 'messy' blog corpus where personality categories were based on individuals' self-reporting of an internet meme personality quiz. They found that, using a model trained on the smaller corpus, results were less reliable when tested on the large corpus, but that there was enough evidence of successful classification that such 'messy' data could plausibly be used in future for classification tasks, thereby saving a lot of expense of data preparation and participant interviews.

Gill et al. (2009) found links between personality type and motivation for blogging, which was tied to topic and sentiment, with high extravert bloggers using the medium to share details of their lives and express emotions, while high neurotic bloggers tended to write negative and cathartic posts. Iacobelli et al. (2011) found that word-stem bigrams outperformed LIWC dictionaries for personality classification on blog text, and they identified some characteristic indicators of openness (religion for low scorers), conscientiousness (discussion of planning and evaluation for high scorers), extraversion (swearing and positive emotion for high scorers), agreeableness (positive sentiment for high scorers), and neuroticism (problems and negative emotion words for high scorers).

Recent work from the field of psychology has also found correlations between function word use and psychological state (Chung & Pennebaker, 2007) and mental illness (Coppersmith et al., 2014).

Native Language

When writing or speaking in a second language (L2), the native language (L1) of an author is likely to leave a mark, even if the production is generally fluent. In writing, this may be observed as variation in lexical choice or syntactic patterns, or as the appearance of various classes of 'error'.

Koppel et al. (2005) use three classes of feature: function words, character n-grams, and error analysis. They define various error classes including typographical substitution and transposition, errors of agreement, missing or repeated words, neologisms, and rare part-of-speech collocations. They found that including errors as an additional feature set improved performance of models built using n-grams, function words, or a combination of the two, although errors alone did not generate best performance. They also identified some interesting features that were more prevalent only for a subset of individual authors, such as the greater prevalence of minor spelling errors (repeated letters, single-letter substitions) in Spanish-authored text.

L1 interference can come from a variety of sources, from phonetics to syntax. Tsur & Rappoport (2007) found some evidence that L1 phonology may affect L2 lexical selection, even in written media, with individuals avoiding words that are phonologically difficult for them, e.g. those containing sound combinations that do not exist in their native language. Wong & Dras (2011) examine syntactic variation, using parse trees to generate features, and demonstrate an improvement over Koppel et al's classification accuracy by including context free grammar (CFG) production rules. In some languages there exist specific grammatical constructs for expressing deference and politeness; English is relatively limited in this regard. Studies of cross-cultural and interlanguage pragmatics have found that L2 speakers will often use pragmatic strategies from their native language(s), or mis-apply L2 pragmatic strategies that they have been taught (see Kasper & Rose (2002) for a comprehensive overview).

The early studies of native language interference all used the International Corpus of Learner English (ICLE; Granger 1998) for their work. This is a collection of learner essays, and as such, has a number of inescapable limitations: the production environment is a classroom context, topics were often set by the tutors (who collected the data for the corpus), and the students are far from being fluent non-native speakers so errors are particularly prevalent.

Brooke & Hirst (2012) address concerns about small, single-corpus studies by gathering data from a popular language practice website. This data is more divergent than the ICLE essays, as users can write about whatever they choose, at any length; users also span the full range of fluency. They then combine this with ICLE essays and a subset of the Cambridge Learner Corpus to create a larger dataset. Brooke & Hirst's features include function words, n-grams (character, word, and part-of-speech), dependencies, and CFG production rules. They demonstrate that models trained on ICLE alone do not perform well on out-of-domain data, but Bykh & Meurers (2012) contest this claim using a combination of lexical, part-of-speech, and hybrid n-grams as features.

2.2.2. Audience and Relationship

The intended audience of a communication, and the speaker's relationship to that audience, is another factor that will significantly affect their choice of language. This encompasses the symmetric and asymmetric relations in Brown & Levinson's politeness theory: the social distance between the two participants, and how they stand in relation to one another in terms of status and power. The size of the audience, and the communities containing the communicants, are also salient here.

The author may have further knowledge about the audience which impacts on their choice of language in specific ways. If the audience is known to have a lower educational level, for example, or to be specialists in a different field, then the author may choose to avoid technical language. There are potentially infinite dimensions to this, and it will not be possible to account for all of them in a systematic manner.

Social Distance

Social distance is what Brown & Levinson (1987) refer to as the symmetric relation. The 'distance' metaphor is firmly embedded into ordinary language: one can hardly talk about one's great aunt's second cousin without the phrase 'distant relatives' springing to mind, similarly we have 'close friends' and 'immediate family'.

Peterson et al. (2011) set out to investigate the applicability of Brown & Levinson's politeness theory to email by looking for correlations between informal features in text, and the level of politeness predicted by the theory. The features which they use to identify informal text include informal word lists, punctuation features (such as use of exclamation marks, or missing sentence-final punctuation), and case features (such as lowercase sentences). This was a thorough, systematic study using messages from the Enron email corpus, and found that informality features are distributed largely as predicted by politeness theory.

Social distance is proxied by two distinct measures in Peterson et al's study. First, they use the personal or business nature of a message, applying Jabbari et al.'s (2006) partitioning of the Enron corpus. Secondly, they look at the level of contact observed in the corpus. In both cases, they found that lower social distance (defined as the exchange of personal messages, or a higher degree of interaction) correlated with an increase in informal features, precisely as predicted.

Social Power

Brown & Levinson's asymmetric relation addresses the question of which participant has more power in an interaction. This power may be due to organisational structures, subject matter expertise, social class, age, gender, or any number of other factors. Such power sources need not be explicitly noted by the participants, and often, two or more factors may be in conflict: for example, a worker and their much younger boss, or a technical expert who is awkward and 'unmanageable' but holds power as a result of their knowledge. This fits in with our earlier observation that power within a social context is situational, and negotiated.

French and Raven's (1959) seminal paper sets out a taxonomy of social power. This is framed from the perspective of the person over whom power is being exerted, and is based around potential outcomes (the ability of those with power to reward and coerce), as well as attributes of the powerful (likeability, expertise, and legitimacy). Although this is by no means the only model, it provides a useful starting point for thinking about sources of power.

Kacewitz et al. (2013) examine pronoun use across a number of datasets with known hierarchies. They find a consistent difference in the distribution of personal pronouns, with lower-status individuals using more first person singular pronouns, while those of higher status tended to prefer first person plural, and second person. The implication is that those of lower status tend to be more focused on themselves, while in positions of leadership a greater outward focus is required. The study addresses only correlation; the question of causality is left to future work.

In a discourse completion test, Morand (2000) finds that more politeness strategies are employed when addressing individuals of higher status, when other variables are controlled. Negative politeness strategies such as hedging and questions are particularly preferred.

Sexton & Helmreich (2000) examined the difference in language use between the members of a flight crew. Of the three roles examined (Captain, First Officer, and Flight Engineer), Captains were found to use 'we' more often than the others, and to use more words in general. Engineers were found to use longer words, possibly due to a function of their job being to describe technical information to their colleagues. This is a clear demonstration that, in at least one professional field, stylistic variation can be used to distinguish social roles. We are interested in whether this result can be generalised to an office-based context.

Panteli (2002) studied the use of email within two academic departments at the same university, one science department and one social science. In each case, the study examined the mailbox of one researcher within the department, who was recently appointed to a relatively junior position. Other message participants were classified as 'professor,' 'lecturer,' 'researcher' or 'administrative' depending on their role. The messages were assessed for surface features such as signatures and greetings, as well as a deeper analysis using 'deconstruction' techniques.

She found that higher-ranking staff were less likely to include any form of address, and more likely to send messages that were short and unstructured. Within the science department, where email communication was more prevalent, she found that research assistants were likely to use 'chat' abbreviations and nicknames when talking amongst themselves, but would adopt more formal language when sending messages to professors. This is a clear indication of how an individual may change their style (consciously or otherwise) to suit the audience.

Other examples are more subtle, for instance, an example is given of a response to an email which blatantly neglects to answer the question posed, even though the sender is clearly in a position to provide the answer. In this instance the question was asked by a junior staff member, and the (inadequate) response is from the head of the department: one imagines that the inverse situation is less likely to occur. In another message, one senior academic refers to a colleague "sending" a junior colleague to see someone. These behaviours show that existing office hierarchies are not erased by the medium of email, and may indeed be evidenced in linguistic behaviours.

Peterson et al. (2011) modelled relative power in the Enron corpus by using the organizational ranks of participants; a subset of 3,999 messages was identified, consisting of one-to-one messages where the rank of both sender and recipient is known. In general, bigger difference in rank was correlated with less use of informal text, which is in line with theoretical predictions.

Danescu-Niculescu-Mizil et al. (2013a) take a slightly different approach to the case of social power, looking at politeness as a predictor of success within a community. They show that politeness is a precursor to promotion, at least in a community-approval model such as becoming a Wikipedia admin: users who employ more politeness strategies are more likely to succeed in their social goals.

Prabhakaran et al. (2013) annotate a subset of the Enron corpus for power, based on human judgement. They annotate for hierarchichal power (derived from the org chart), situational power (having authority for a given task), influence (attempting to persuade others), and 'controllers' of the communication (those who are proactive in steering the conversational goals, rather than simply reactive). They then proceed to use these annotations to study the different behavioural manifestations of these power sources in email (Prabhakaran & Rambow 2013).

Even in fictional scenarios such as movie scripts, linguistic choices can be seen to reflect power differentials (Buitkienë 2006). The fact that these effects are observed in created as well as natural dialogue shows that this is an important facet of human behaviour.

Audience Size

Another feature of the audience is its size: addressing a large group is a different proposition to an intimate, one-to-one discussion, even if both circumstances involve individuals who otherwise stand in the same relation to oneself.

Pérez Sabater et al. (2008) found the use of salutations was the feature which exhibited most variation between one-to-one and group messages in their data: group messages were found to almost always include a formal salutation, while one-to-one messages exhibited more variation.

As part of their study into factors affecting use of polite and informal language, Peterson et al. (2011) identified a correlation between number of recipients and level of formality. Informality was found to decrease as the number of recipients increases, up to ten recipients, at which point informality increases again. This increased informality in large-group emails is an interesting result, worthy of further study, in particular to explore whether this is a result of large group emails tending to cover different (more social) topics such as group outings.

While email can be considered an 'electronic letter', there are some aspects of online interaction which have no parallel in pre-internet communications. Skovholt & Svennevig (2006) examine the phenomenon of 'CC' use in around 700 email exchanges, comparing recipient roles (To and CC) to explicit named references in the text, and examined the different behaviours underlying the use of the 'CC' field.

One common use was to add someone as a 'witness' to a communication which did not otherwise involve them. Although the authors do not proceed to discuss this, there may be implications for identifying the social roles of individuals who are commonly included in this way, as well as for drawing inferences about the circumstances in which senders feel the need to have their communications thus witnessed.

There is also an open question of how the existence of a broadly passive, online audience may influence communicative style, for example in the case of ostensbily one-to-one communications (such as twitter 'replies') which nevertheless take place within the public sphere. Research to date has not really examined this area.

National and Organisational Cultures

Relationships do not only exist between pairs of unrelated individuals. Social groups have their own cultures, at the community, organisational, regional, or even the national level, which may translate into different behavioural and linguistic norms. Even online environments have their own virtual cultures, for example forums with their own shared jargon and social norms.

Waldvogel (2007) examined greetings and closings in 515 emails collected from two organisations in New Zealand. In one organisation (SCT, an educational establishment), staff morale was low due to recent staff restructuring, while the other (Revelinu, a manufacturing plant) is described as having a close-knit and friendly culture. For each organisation, two corpora were collected: one data set constituting a week's worth of email from an individual at 'senior manager' level, and one exchange of messages focused on a specific topic. The study concentrated exclusively on messages where both sender and recipient were within the organisation in question, in order to focus on intra-organisational culture. Waldvogel's hypothesis was that the choice of greetings and closings should reflect the organisational culture, as well as the individual relationships between staff members, and that significant differences should therefore be observable between these two environments.

The study found that the majority of messages within SCT began without any greeting element at all (59%). Messages within Revelinu, on the other hand, usually featured a specific greeting word such as 'hi' or 'dear' (58%), and this was typically accompanied by the recipient's name. Beginning a message with only the recipient's name was a popular alternative in both organisations, accounting for 21% of SCT messages and 25% of Revelinu messages. Different patterns of message closings were also observed between the two organisations. In SCT the most common sign-off was the sender's name (38%), but 34% of messages featured neither closing words nor the senders name. Within Revelinu the vast majority (75%) of messages featured a specific closing word or phrase, usually accompanied by the sender's name. Only 10% of Revelinu messages did not include any closing.

Waldvogel also considered the impact of intersections with gender, hierarchy, and social distance. While some effects were observed, these were less significant than the effect of the prevailing organizational culture, and were not the main focus of the study. Nevertheless, the results are of interest. In both organizations, individuals of higher status were more likely to be addressed by name, and in SCT (where closings were less common), they were more likely to receive a closing of some form. Message senders were asked to rate their colleagues as 'close' or 'distant' which was used as the measure of social distance. Close colleagues were more likely to be addressed by first name only, while those with greater social distance were more likely to include a greeting word. At Revelinu, a similar effect was observed for closings, with distant relationships more likely to include both a closing phrase and the sender's name. The results for gender were reversed between organizations: at SCT, women were more likely to use names, greetings, and closings, while at Revelinu men were more likely to do so.

Waldvogel's study contrasted two New Zealand companies; in a study crossing national borders, Murphy & Levy (2006) asked a group of Australian and Korean academics to complete a short questionnaire on their experiences of email exchanges with foreign colleagues. The questions asked participants to consider their own messages as well as messages received from overseas. Around half of the respondents (from both cultures) claimed to express politeness differently when writing to a foreign colleague.

When asked about incoming mail, 27.5% of Australians and 40.0% of Koreans felt that they had received messages from overseas which were insufficiently polite; however, no comparable question was asked about email received from their own culture, so the significance of this result has not been established.

More interestingly, participants were also asked for examples of polite and impolite language, which provides some valuable qualitative data. Responses highlighted features such as incorrect use of titles, excessive brevity, and lack of expected greeting and closing formulations. There was some difference in emphasis between the two cultures, but many of the examples were cited by both nationalities, suggesting that this was less to do with cross-cultural differences, and more about a general perception of impoliteness.

Another comparison of linguistic strategies between speakers of different language backgrounds is found in Kankaanrata's (2005) PhD thesis, in which she considers English-language email in a multinational corporation, focusing on Finnish and Swedish staff. Emails in the corpus are categorised as 'noticeboard' (broadcast of notices for staff), 'dialogue' (interpersonal discussion) and 'postman' (sending of attachments, files, etc); the dialogue category is the only one of interest to us. Salutations with first names were found to be almost universally used by both nationalities. On the other hand, different politeness strategies were employed by Finnish and Swedish writers when making requests: Swedish writers were more likely to use indirect requests, while Finns were more likely to use a direct request with explicit politeness markers such as 'please'.

Pérez Sabater et al. (2008) examine 100 emails from an academic context, which are formal in general, with comparatively few non-standard features. The features included in their study are openings/closings, contractions, politeness indicators (such as 'thanks'), and non-standard linguistic features (such as spelling mistakes or grammatical errors). They found that non-native speakers tended to use more informal language, with more contractions and fewer politeness indicators, while simultaneously using more formal salutations and valedictions. This may indicate a different preference for expressing politeness, but given the heterogeneity of non-native speakers represented in the data, it is hard to generalise from these observations.

2.2.3. Purpose of Communication

The purpose of communication is obviously going to affect the choice of language in a direct manner: one cannot order a pizza by limiting one's discussion to the weather. Topic is perhaps the most obvious expression of purpose, tying directly into the frame of the communication.

Planning to meet up for coffee is more enjoyable than arranging an interview or to discuss a pay rise, and the overall register of language used is therefore likely to be less formal. It is therefore unsurprising to find that stylistic variation is not independent of topics under discussion (e.g. Golcher and Reznicek 2011). And we have already seen that topic is not independent of demographic features such as gender and age (Argamon et al. 2007; Rosenthal & McKeown 2011). In fact, Argamon et al. report that they are unable to separate topical from stylistic effects, in their blog corpus.

An early study demonstrating register variation across different genres of text was undertaken by Biber (1993), who also demonstrates a high success rate for predicting genre of English texts based on five 'dimensions' of register. The paper describes in detail the linguistic features used to identify these dimensions. Biber also summarizes key results from studies in other languages, to illustrate that his findings are not specific to English. Biber's proposed dimensions are derived from examining linguistic features which commonly co-occur. His five categories are as follows:

- Informational vs. involved production
- Narrative vs. non-narrative concerns
- Elaborated vs. situation-dependent reference
- Overt expression of persuasion
- Abstract vs. non-abstract style

Topic and genre are often conflated in studies of domain adaptation; in a machine translation context, van der Wees et al. (2015) begin to pick apart these two dimensions, but this is early work. They define topic as "the general subject of a document", and genre as a more complex, complementary concept encompassing "non-topical text properties function, style, and text type". For their study, van der Wees et al. produce a benchmark corpus of parallel text Arabic-English data, selected from newswire and user-generated (online comment) text covering the topics of economy, culture, health, politics, and security. By separating these two dimensions, they find that they can uncover more nuance in explaining machine translation errors, for example colloquial expressions that appear in user-generated comment but not in newswire. Overall, they found that the difference in performance was greater between genres than between topics.

In other work on genre classification, Santini (2007) proposes that web pages require the option of assigning multiple genres, as a single page may contain components with different registers and functions (e.g. articles, navigation menus, interactive elements, and multimedia). Santini also discusses the blending of genres that can occur when a specific document adheres to the conventions of more than one genre of text, and the phenomenon of individuals blurring genre boundaries to meet their personal requirements, both of which complicate the issue of genre classification.

The immediate goal of a message is also likely to influence the form used. If a message contains a face threat, then more politeness features would be expected, in line with Brown & Levinson (1987). Peterson et al. (2011) demonstrate that this expectation is met in the case of Enron messages containing requests: the presence of requests in a message, as determined by an automated request classifier, reduces the level of informality observed in the text. Conversely, a message intended as social 'glue' is more likely to display informal features, as informality can be used to build rapport and solidarity.

A request is only one example of a class of linguistic phenomena collectively known as speech acts. The essential idea behind the theory of speech acts is that when two (or more) people converse, there is more to that interaction than a simple exchange of factual statements. We do things with words – whether asking a question, making a promise, giving advice, or any of a number of possible options. Words are spoken with a purpose.

Searle (1969) begins his discussion of speech acts with the assumption that it is always possible for a speaker to put his precise meaning into words (although he acknowledges that in an extreme case it may be necessary to change to a different language, or invent new words). However he goes on to recognise that speakers most often do not go to such lengths, and therefore may rely for meaning on something beyond the surface form of the words. An utterance will often do more than simply carry its surface meaning. If I comment that "it's cold in here," for all that the utterance may be a statement of fact, in reality I'm probably asking you to do something about it – which depending on the context may be to close the window, pass me my jumper, put the heating on, or simply provide an appropriate level of sympathy.

A performative act may be considered felicitous if all the necessary external factors are in place, such that the intended action may indeed be performed, or infelicitous otherwise (for example, making a promise with no intention to keep it, or proclaiming an illegal marriage). There's no assumption in the literature that infelicity must be deliberate, or even that the speaker need necessarily be aware of whether or not the necessary requirements have been met.

The inauguration of Barack Obama as President of the United States gave speech act theorists pause for thought in 2009 when first Chief Justice Roberts (who was administering the oath), and then the incoming president himself, accidentally diverged from the constitutionally-mandated wording of the oath. This kind of oath is a classic example of an explicit, performative speech act. The oath begins "I do solemnly swear…", and by speaking the prescribed words, the president performs the act of 'swearing'. The new Obama administration was so concerned about the possible infelicity of the incoming president's mistake that, in order to eliminate any doubts about the validity of his presidency, he re-took the oath the following day.

Speech acts have come to be called 'dialogue acts' in the field of dialogue processing; we shall use the two terms interchangeably.

Bracewell et al. (2012) study social intentions in the digital sphere, moving beyond the traditional classes of speech acts into social goals of a more pragmatic nature, such as establishing [one's own] or challenging [another's] credibility, or expressing solidarity. These goals can be considered as orthogonal to more traditional, surface-level communicative goals such as asking questions or providing information. For example, a question can be used to cement one's own authority by showing up gaps in another's knowledge, in just the same way as a statement can convey subject-matter knowledge.

2.2.4. Production Environment

The physical environment surrounding the author at the time of message production will likely have an impact on their writing style. An author will produce different text if they are in their office, comfortable and unhurried, as opposed to if they are rushing to meet a pressing deadline or tapping out an email on the train. Many of these aspects are notoriously difficult to measure after the fact, and can only be controlled if data is collected in experimental conditions (which itself introduces other limitations).

Most studies control for the impact of medium by limiting their scope to a single source (e.g. dialogue acts in email, or linguistic accommodation in web forums). When comparative studies have been undertaken, it has generally been a qualitative examination of the differences, for example Baron's (1998) comparison of email to letter-writing and conversational speech, or Duthler's (2006) study of email and voicemail requests.

Twitter data provides a partial record of production context, as each tweet is accompanied by details of the device or client used to submit the message, such as Blackberry, iPhone, and PC (website). In a recent study of Twitter data, Gouws et al. (2011) show that significant differences can be identified between tweets submitted via different clients. Mobile users were found to use fewer out-of-vocabulary items than PC users, possibly due to predictive text/spelling correction, but certain formulations (such as 'u' for 'you') were used more often by mobile users.

Duthler (2006) compared politeness strategies employed in voicemail and email messages, for both high- and low-imposition requests. Their work sets out to test three hypotheses:

> H1: Imposing requests will be rated higher on measures of politeness than unimposing requests.
> H2: Electronic mail requests will be rated higher on measures of politeness than voicemail requests.
> H3: Less imposing requests made via email or voicemail will not differ on ratings of politeness. Highly imposing requests via email will be rated higher on measures of politeness than voicemail requests.
>
> — Duthler (2006)

This study elicited data by setting up an experiment in which student participants were required to request a meeting with a professor, using either email or voicemail. Participants were assigned at random to one of the four groups: low imposition by email, low imposition by voicemail, high imposition by email, high imposition by voicemail. The degree of imposition of the request was varied by changing the scenario: in the low-imposition case, the meeting was to be during regular office hours, and in the high-imposition case, the request was for a meeting outside of office hours. A preliminary study of 28 students showed that a meeting within office hours was graded as a low imposition, with a mean score of 1.36 on a scale of 1 to 7, while requesting a meeting outside of office hours was considered a comparatively high imposition (4.71).

Analysis of the resulting 148 messages found that email exhibited more variation in line with degree of imposition, as compared to the relative stability of voicemail (in support of H3). The other two hypotheses were partially supported. Adjunct phrases (defined as supporting phrases which were not part of the key task execution, for example greetings, closings, and small talk) were more common in email than in voicemail, and more common in imposing requests than in low-imposition requests. However, forms of address did not vary (on average) between email and voicemail, and were found to be, unexpectedly, more formal in low imposition requests than in the case of higher imposition. Adjunct phrases are considered to be a politeness strategy, but the number of adjunct phrases had an inverse relationship to 'overall politeness' as graded by two expert annotators. This is a surprising result, but may simply indicate that students choose to use either adjunct phrases or another politeness strategy to achieve their goals.

The greater variation observed in email messages may be due to having more time to think about (and edit) linguistic choices when composing an email, compared to leaving a voicemail. That email exhibits such variation makes it an ideal medium for examining politeness in relationships.

Morand & Ocker (2003) consider computer-mediated communication from a sociolinguistic perspective, and propose a number of theories for how politeness strategies and their implications may transfer from their spoken origins into the world of email communication. This work is based on a consideration of previous literature, and provides a number of hypotheses which could be tested by future work. Central to their hypothesis is the idea that both positive and negative politeness strategies are required in computer-mediated communication, just as in face-to-face communication, and moreover that these strategies perform the same roles in CMC. The authors map these considerations onto various other scenarios which have been studied in the CMC literature.

One interesting observation on the nature of computer-mediated interaction is the possibility that misunderstandings are more likely. Ambiguity is a common feature of language, and more ambiguous phrasing can often be used as a politeness strategy (such as hedging), but in text-based communication there is a relative dearth of contextual cues such as intonation and body language to assist with interpreting ambiguity. Without these additional cues, it can be harder to convey the pragmatic components of a message.

There is also less chance for interlocutors to adapt their politeness level as the 'conversation' progresses, due to longer turns with no possibility of interruptions. If you cause offence in an email, you will not know it unless the recipient chooses to reply, and by that stage it may be too late to mitigate the impact. This is in direct contrast to face-to-face communication, where it is easier to pick up non-verbal cues as you are speaking, allowing you to adapt your language accordingly (including offering up a speedy apology for any unintended insult).

2.3. Modelling Human Language

Our intuition, backed up by the various studies we have examined, is that language changes in response to a combination of contextual and environmental variables. We now turn our attention to the question of how best to capture the extent of these linguistic variations.

All language modelling is essentially an approximation, reducing layers of expressed and implied meaning into a (hopefully) simple model. The different underlying assumptions and dimensions used to construct the model will have an inevitable impact on the findings which can be uncovered.

2.3.1. N-Grams

One frequent approach in linguistic studies is to use word-level n-grams as input features for a model. N-grams are an extension of the 'bag of words' model, in which a document is treated simply as a collection of unrelated words, which are then counted. The value of n can be adjusted to capture a wider context: higher values of n result in models that are larger and more sparse, but which can describe the data more accurately.

For example, consider the following sentence:

> "This is a short example sentence."

Represented as overlapping trigrams (n-grams where n=3), this sentence would be divided as follows, with each trigram having a count of 1:

{this is a}, {is a short}, {a short example}, {short example sentence}

Across a whole document or document set, these distributions become more meaningful. Common phrases accrue higher counts, while rare vocabulary may appear only in trigrams of count 1. These raw scores can then be transformed into percentages.

N-grams are designed to capture elements of syntax and phrase-building, without requiring prior knowledge of a language or its structures. It is a brute-force approach to capturing the shape of a language.

The initial construction of n-gram models is dependent on having a training set of representative data from which to build the model: in common with other supervised methods, if a particular phenomenon is not in evidence in the training data, there is no chance of accounting for it, although there are various smoothing and back-off techniques to give an approximation. For example, the back-off probability of a previously unseen trigram may be calculated based on the probabilities of its constituent bigrams or unigrams. Chen & Goodman (1999) provide a good overview of smoothing techniques in the context of machine translation models.

There are many aspects of linguistic expression which are not captured by an n-gram model. Not least, many non-standard language features typical of informality are innovative forms – for example, affective lengthening, where letters are repeated for emphasis, such as 'reallllly'. Cases such as this where the individual author can choose how to vary from the standard orthography, are by definition likely to be rare, as different authors may choose different variations. N-gram models do not handle rare features well, and 'out of vocabulary' items (those never observed in the training set) will be ignored. The typical approach to an unseen phrase is to 'back off' to use the individual frequencies of the component words, but unseen words have no such back-off probabilities, and can only be approximated by a miniscule non-zero probability.

Additionally, it is not invariably the case that adjacent words are the most relevant to one another. It is possible to increase n to use a larger window, but only at the cost of a larger and sparser model. ‘The green car’ and ‘the red car’ probably have more in common than ‘the green car’ and ‘the green movement’, but this isn’t obvious from n-grams alone. Skip-grams (non-consecutive n-grams) can capture more contextual information for a given window size, but at the cost of increasing the model size (Guthrie et al. 2006). For our earlier example sentence, and allowing a skip of at most 1 token, the model now contains 10 rather than 4 trigrams. A hyphen is included here to show where each skip takes place.

{this is a}, {this — a short}, {this is — short}, {is a short}, {is — short example}, {is a — example}, {a short example}, { a — example sentence}, {a short — sentence}, {short example sentence}

Another alternative to word n-grams is part-of-speech n-grams. As part-of-speech patterns tend to be more universal (within a given language), this kind of modelling is less likely to encounter out-of-vocabulary issues, but with the consequence that the model may lack specificity.

{this is a}, {is a short}, {a short example}, {short example sentence}
{DT VBZ DT}, {VBZ DT JJ}, {DT JJ NN}, {JJ NN NN}

Bramsen et al. (2011) have also used word n-grams and part-of-speech n-grams in combination to capture more nuances, which they call “mixed” n-grams:

{this is a}, {is a short}, {a short example}, {short example sentence},
{this is DT }, {is a JJ}, {a short NN}, {short example NN},
{this VBZ a}, {is DT short}, {a JJ example}, {short NN sentence},
{DT is a}, {VBZ a short}, {DT short example}, {JJ example sentence},
{DT VBZ a}, {VBZ DT short}, {DT JJ example}, {JJ NN sentence},
{DT is DT}, {VBZ a JJ}, {DT short NN}, {JJ example NN},
{this VBZ DT}, {is DT JJ}, {a JJ NN}, {short NN NN},
{DT VBZ DT}, {VBZ DT JJ}, {DT JJ NN}, {JJ NN NN}

Although this combined methodology generates a much larger model, it gives a pleasing way of capturing the similarity between green and red cars: both are {the JJ car}. A more compact model may be generated by first building a full model, and then pruning it back to retain only the most significant subset of features.

These examples neatly illustrate the limitations of n-grams. To compensate for the naïvety of the underlying methodology additional layers of complexity (such as skip-grams or mixed n-grams) must be added, and consequently larger models used, if one wishes to capture desirable features that are expressed at deeper levels of the language. The size of these models therefore grows exponentially. If such models are subsequently pruned by an automated process, retention or elimination of particular feature will depend on significance within a particular dataset rather than any extrinsic measure of relevance to the task at hand, which can lead to overfitting.

Additionally, because n-gram models treat all words as equal, topic and style features are conflated. This can be seen in n-gram analyses such as Gilbert’s (2012) study of power and language in the Enron corpus: emotive language (‘shit’, ‘tremendous’) and function word phrases (‘is really’, ‘haven’t been’) sit alongside strongly business-oriented vocabulary (‘candidate’, ‘promotion’, ‘compensation’) in his table of significant phrases.

2.3.2. Stylometry and Linguistic Formality Metrics

Stylometry is the study of linguistic style, as distinct from topic. When someone communicates, whether in speech or in writing, they must make decisions about not only the information they need to convey, but also the manner in which they wish to convey it. Such choices may not always be made at the conscious level but it is a fundamental tenet of Gricean politeness theory (Grice 1975) that, in order to be polite, it is necessary to use an appropriate manner. This, however, is a broad requirement that can be fulfilled in an infinite variety of ways, leaving a lot of scope for individual stylistic variation.

The inception of stylometry is traditionally traced back to Lutosławski’s (1897) studies of Plato’s dialogues, in which he identifies the importance of aggregating minor features into a meaningful signal. His point is crucial, and bears repeating in full:

> “Very important peculiarities are very few, while accidental coincidences may be found by the thousand. And their accidental character, even if fully recognised as accidental, does not deprive them of chronological importance, if sufficient numbers of such accidental coincidences are taken into consideration. The single occurrence is accidental, though it may be exceedingly significant, as, for instance, the occurrence of μεθεξις in *Parmenides* and *Sophist*. But if one dialogue has twice as many accidental coincidences with the *Laws* as another, this result is no more accidental than the difference of mortality between England and Spain.”
>
> — Lutosławski (1897)

More recently, in computational studies, stylometry has been used in authorship identification tasks, for example attributing anonymous works to famous authors (Mosteller & Wallace 1963; Holmes 1994), or latterly in the field of forensics (Koppel et al. 2009, Zheng et al. 2006, Turrell 2010). These tasks tend to be closed-class: namely, the goal is to match a piece of unknown text to one of a small subset of possible authors, or to verify whether a particular suspect could plausibly have been the author of a text. Related techniques have also been used to group individuals into clusters, either supervised according to some previously percieved class (e.g. age or gender, as discussed above, e.g. Argamon et al. 2007), or unsupervised as an exploratory technique (Layton et al. 2011; Akiva & Koppel 2013; Arefin et al. 2014).

One popular framework for automating stylometric research is Linguistic Inquiry and Word Count (LIWC), a system designed to categorise texts along a number of axes, inspired by findings in the field of psychology (Pennebaker et al., 2007). The basic theory is simple, and sound: language is the main medium through which we are able to communicate, and it follows that an individual’s choice of language will correlate with their mental state. The actual implementation of LIWC is rather naïve, consisting of a dictionary of representative word stems for each category. These words have been hand-picked and validated by human annotators, but despite years of study devoted to word-sense disambiguation, it is still not possible to guarantee that a computer program will correctly identify the sense of a word being used. To generate scores for a text, the program simply examines the document word-by-word to count the proportion of its words that fall into each of the LIWC categories.

The categories include basic parts-of-speech such as articles and function words, topics like ‘death’ and ‘religion’, as well as more subjective groupings such as ‘positive emotion’, ‘anxiety’, and ‘certainty’. Although these lists have been judged to accord with the majority of human annotators (Tausczik & Pennebaker 2010), the authors themselves acknowledge that the system is not able to cope with all the subtleties of natural language. In particular, contextual information is disregarded, and there is no scope for handling irony, sarcasm, jokes, or deception.

LIWC’s scores are best considered as potential inputs to a more sophisticated system, and a fast-growing body of work has used LIWC features along with machine learning to predict personal attributes including personality types (Pennebaker & King 1999; Mehl et al. 2006; Iacobelli et al. 2011), gender (Newman et al., 2008), and age (Pennebaker and Stone 2003), as well as social attributes such as status within a hierarchy (Sexton & Helmreich 2000; Kacewicz et al. 2013). For each of these use cases, it is a different subset of the core LIWC feature set that proves most salient.

At the other end of the spectrum from function words, the origin of core vocabulary has also been used to analyse the formality of text, such as in DeForest & Johnson’s (2001) study of dialogue in Jane Austen’s work, where Latinate vocabulary was found to correlate with higher-status and male characters, while words of Germanic origins were more prevalent in the speech of women and characters of lower social status. Fang & Cao (2010) observe similar formality correlations in a study of the British National Corpus, with academic texts featuring more Latinate vocabulary.

There are many other ways of capturing stylometric information, but all methods ultimately result in a vector of stylistic features. Details of our specific implementation can be found in Section 3 (Methodology); we will now take a brief excursion into some relevant feature categories and surrounding research.

Common Ground and Deixis

One key decision to be made before embarking on any communication is that of how much detail to provide. This derives naturally from Grice’s (1975) notion of implicature, and his maxim of quantity, which states that interlocutors should provide just the right amount of information to be correctly understood.

The amount of detail that is “just right” will vary according to the shared experiences of those communicating. Childhood best friends with years of common experiences behind them may be able to convey a lot of information in few words (or wordlessly, sharing just a look), while strangers meeting for the first time will have to guess, attempting to pitch their words appropriately based on intuitions of their respective educational and cultural backgrounds, and assisted by the context of the communication, which is shared by definition.

This concept of common ground is the basis for Heylighen & Dewaele's (2002) definition of formality, expressed in terms of the amount of contextual information included in a text. In brief: the more formal the document, the more detail is included explicitly, and the less background knowledge is taken for granted. Under this definition, formal language is used primarily for the avoidance of ambiguity, and consequently more formal language is likely to be used in circumstances where there is less shared knowledge, or where there is a greater requirement to avoid misunderstandings.

From this definition, Heylighen & Dewaele proceed to define a practical and tractable measure of formality in language. The simplifying assumption is that certain parts of speech (such as verbs and pronouns) are inherently more deictic, which is to say, they rely on contextual interpretation for their meaning, whereas other classes (such as nouns) are concrete and do not depend on external context. These deictic words are therefore inherently more informal, by the above definition of formality, and F-score[1] is designed to be higher for more formal texts, due to the simple division of parts of speech into those which are inherently more and less deictic, as in Table 1. The F-score is calculated in the following manner from parts of speech:

F = (non-deictic frequency – deictic frequency + 100)/2

Where the frequency of each part of speech is the percentage of tokens which fall into the category in question.

Deictic Parts of Speech	Non-deictic Parts of Speech
Verbs	Nouns
Adverbs	Adjectives
Pronouns	Prepositions
Interjections	Articles

Table 1: Heylighen & Dewaele's (2002) division of parts of speech into deictic and non-deictic categories.

Conjunctions are explicitly noted as the one class which does not contribute to the categorisation, on either side.

Heylighen & Dewale applied this measure to texts from a variety of media and languages, both to assess its value and to see how well it transfers between languages. Initially, French data was collected from language students in three settings: an informal conversation, an oral exam, and a written essay. These sources produced overall F-scores of 44, 54, and 56 respectively, which ties in with our intuition of the relative formality of these genres.

From word counts of a Dutch corpus, written text (F=62) was found to be significantly more formal than spoken (F=42). This data was also divided by medium, for example scientific texts (F=66) were shown to be more formal than novels (F=52). For two distinct Italian corpora, novels score 58 and 64 respectively,

[1] Heylighen & Dewaele refer to their score as 'F-measure' but, due to the naming conflict with the more commonly used F-measure in statistical analysis, I shall use 'F-score' throughout.

and scientific papers 69 and 72. A similar pattern holds for other genres: the average F-score of each category in Italian is higher than the Dutch equivalent, but the ordering is roughly constant. As pronouns are optional in Italian, this could contribute to fact that Italian scores are typically higher.

In the same study, differences in average F-score are also reported between certain groups of people. Men, introverts and the highly educated are found to use more formal language on average than women, extraverts, and those with lower levels of education, respectively. These are presented as preliminary results in need of further investigation, but this highlights the fact that multiple factors may contribute to an individual's choice of style.

Lahiri et al. (2011) calculated Heylighen & Dewale's F-score at the sentence level, and compared this to human judgement of formality at the same level. They collected data from a variety of sources (blogs, news articles, academic papers, and forum threads) and calculated the F-score for each sentence. They found that they could distinguish these genres solely on the basis of formality, with the exception of blogs and news articles — a discrepancy explained by the selection of blogs from the Technorati list of 'Top 100 Blogs', many of which were news-like in style.

Human annotation was restricted to sentences from 50 blog posts (7.5k sentences), which were labelled as 'formal' or 'informal' to give a ground truth data set. A clear difference in F-score was observed between these formal and informal sentences, which supports the intuition that F-score can be applied at the sentence level. The same authors also measured the degree of inter-annotator agreement between untrained annotators on a formality task, with a less promising result: the raw agreement figure was high, but kappa statistic was poor. However, the study was specifically designed to measure 'inherent' human agreement, so no definition of formality was provided to the annotators. It is likely that inter-annotator agreement at this task could be increased by the provision of guidelines.

Abu Sheikha & Inkpen (2010) also studied formality at the sentence level, using sentences from a variety of corpora including Enron email and Reuters newswire. They investigated the predictivity of a number of features including formal and informal word lists (manually constructed), contractions, abbreviations, phrasal verbs, active voice, and average word length.

Based on this data set, a sentence-level classifier was trained using three different algorithms: Naïve Bayes, Support Vector Machines, and J48 Decision Trees. 10-fold cross validation was used for each classifier. Results were encouragingly high, with similar performance across the three algorithms. Decision Trees gave the best result, with an F-measure of 0.985, while the Naïve Bayes performed worst at 0.970. They found the most significant features to be increased use of first- and second-person pronouns in informal text, average word length, and the use of words from their custom informal word list; they also reported that removing the worst-performing features degraded classifier performance.

A major limitation of this work is that, despite conducting classification at the sentence level, they assumed whole data sources would be formal (newswire, academic texts) or informal (email, speech transcripts, personal letters).

Consequently there is no sentence-level ground truth, and no human judgement was used to produce gold standard data. The result of this is that, although the classification results were high, it is not obvious that the authors were exclusively measuring the effect (formality) which they had intended: rather, in practice, they are measuring their ability to correctly identify the source genre of a text.

Teddiman (2009) applies Heylighen & Dewaele's F-score to blog entries and comments, and compares the result to scores from other data sources. Blogs were collected from LiveJournal, which is generally used for personal, 'diary' style posts. Blog authors had a mean age of 24.7, and 65% were female; UK and US bloggers were compared, but as no significant differences were found between these two groups, the final corpus consists of combined data from both countries. Demographic data is not available for the comment authors.

For both blogs and comments, the average F-score was found to be 55.5, which falls between 'school essays' and 'biographies' when compared to the F-scores of BNC categories. In other respects, however, the part of speech distribution was found to be non-typical, sometimes closer to that observed for speech than for other textual genres. For example, in both datasets, use of emphatics such as 'just' and 'really' was found to be higher than expected.

Although the overall F-score was the same for blogs and comments, they do differ in other features. In particular, the distribution of second-person personal pronouns was much greater in the comments (again, this was closer to the values expected for speech). These results indicate that, while F-score is a good starting point for formality research, it is too much of a blunt instrument to differentiate such closely related genres of text.

Hedging

Hedging expressions such as 'sort of' or 'I think' are used to reduce the speaker's responsibility for the claims they are making. Hedging is an indirect politeness strategy as described by Brown & Levinson (1987), allowing a speaker to save face if their claim turns out to be wrong; as such, we might expect those with lower status and power to use more hedging words.

The literature of hedging begins with Lakoff's (1972) examination of fuzzy logic and semantics, in which he examines the relation between hedging expressions and propositional truth values.

Ganter & Strube (2009) observe that work to automate the identification of hedges had previously been limited to the biomedical domain. They proceed to examine topic-independent hedging strategies by use of tagged Wikipedia documents, where 'weasel words' have been marked for editing across articles covering a variety of subjects. This is one way of identifying a set of potential hedges, incorporating both words and phrases, although it remains to be seen how well their models transfer to other data sources.

Linguistic Innovation and Out of Vocabulary Terms

Linguistic innovation will often result in previously unseen tokens. As we have seen, n-gram models do not meaningfully handle unseen tokens, and likewise, wordlist-based approaches will of necessity only be able to identify the most common of 'innovative' forms: those which have become sufficiently standard to make it onto lists. Nevertheless, as we shall see, many instances of non-standard orthography can be mapped onto meaningful and identifiable classes.

One hypothesis is that the more informal the text, the more non-standard vocabulary items will be observed. Non-standard tokens can consist of spelling mistakes or typographical errors, but in modern usage (particularly online), non-standard spelling may also be a deliberate choice. Out of vocabulary tokens may also arise due to codeswitching (informal borrowing of a word from another language), technical jargon, or slang.

While most text-processing applications are interested in these tokens only so far as to normalise them (e.g. Tang et al 2005, Liu et al. 2012), we consider that the use of such forms conveys valuable stylistic information. As Eisenstein (2013) notes, information is encoded in the very choice of these innovative forms that traditional NLP seeks to eliminate, but text processing has traditionally taken the approach of either 'normalising' informal text to match newswire, or applying techniques for domain adaptation to transfer models between contexts.

Most of these cases are associated with informality in some form. In the case of accidental errors, we note that writers are less likely to carefully proof-read and spell-check an informal document. Slang is more commonly used in social and informal situations. Non-standard abbreviations, affective lengthening, and alphanumeric words also fall into the category of slang, of an innovative kind. Codeswitching requires a common second language, which is contextual knowledge in Heylighen & Dewaele's (2002) sense; the blending of languages and cultures online is resulting in a number of hybrid linguistic forms that did not exist in the pre-digital era (Lam 2008). Even technical jargon, while inherently work-related (and therefore, one assumes, more likely to occur in more formal contexts), is indicative of assumed common ground between writer and audience, implying a shared technical language: greater shared vocabulary is likely to correlate with lower symmetrical social distance.

Varnhagen et al. (2010) examined the interaction between 'new language' use and spelling ability in teenagers. While they found (unsurprisingly) that genuine misspellings were more common for those who achieved lower scores on a spelling test, the same was not necessarily true for abbreviations and other linguistic innovations. In particular, girls who were better at spelling tended to use more innovative forms. This indicates that misspellings and innovation may be considered a matter of stylistic choice, not simply a reflection of the author's level of education.

Affective lengthening is the phenomenon of repeating letters within a word, usually for the purposes of emphasis. This can take a variety of forms, and can typically be seen as a case of phonetics invading the written language. As such, it is a highly informal phenomenon.

Existing double-letters are ripe for lengthening, for example in cases such as ‘realllllly’ or ‘coooool’. Brody & Diakopoulos’ (2011) found that ‘nice’ was the most commonly lengthened word in a Twitter corpus of English; swear words also scored highly, and in general, a correlation was observed between emotionality and lengthening. In line with the phonetic origins of the phenomenon, vowels and continuant consonants (approximates, sibilants, liquids, and fricatives) are more likely to be lengthened; it’s hard to lengthen a plosive without reduplication. In more modern usage (such as Twitter) it is also sometimes observed that non-words such as acronyms are lengthened, which may be less phonetic in nature, e.g. ‘lolololol’ or ‘omggg’ (Schnoebelen 2012b). Myslín & Gries (2010) note that the phenomenon of lengthening is also observed in online Spanish, with similar constraints.

Brody & Diakopoulos’ (2011) used a language-independent method for identifying the canonical form of a lengthened form. Their technique relies on grouping together all tokens which differ only in the number of repetitions of a given letter, and using the most common instance as the canonical ‘correct’ form. Cross-checking with a dictionary, they found that of those sets which contained any acceptable dictionary word, the most common observed form differed from the dictionary form in only 2.63% of cases.

Another special case of non-dictionary words, alphanumeric words are used in some online settings to shorten text, e.g. ‘l8r’ for ‘later’ or ‘b4’ for ‘before’. As these words are not a part of standard language, they can be considered inherently informal. Although these forms originated to save space in length-constrained environments such as SMS messaging, they may also be used to signify informality in other online environments.

As they are not part of standard language, emoticons are inherently informal in nature. Emoticon selection is not inherently language-dependent, but the distinction between Eastern and Western styles should be noted, although Eastern-style emoticons are becoming increasingly used in English-language contexts.

Eastern examples:	o_0	>.>	-_-	(*_*)	
Western examples	:-)	;-)	:-D	=)	8)

Choice of emoticon style may be a valuable feature in its own right, particularly if someone varies from their personal norm to match the style more commonly used by their interlocutor. Schnoebelen (2012a) examines the range of emoticons in a dataset of American English tweets, and makes some interesting observations on the differences between nose-using and non-nose-using tweeters. For example, the inclusion of a nose seems to correlate with more correct spelling, fewer swear words, and longer messages.

If emoticons are found to be sufficiently prevalent in a given data set, the proportions of each ‘expression’ may also be a useful feature. Derks et al. (2007) found that, unsurprisingly, positive emoticons were more often used in positive contexts, and vice versa. An early study of emoticon interpretation found that human ratings of emotion in a message were influenced more strongly by the verbal content than by inclusion of an emoticon, but that emoticons could alter the overall sentiment of a message (Walther & D'Addario 2001). These aspects deserve further investigation in future, but in our datasets emoticons are very rarely used, and subdividing them is not useful.

Expletives

Swearing is another feature more likely to occur in informal conversation. In an analysis of the British National Corpus, the word 'fuck' appears twelve times more frequently in speech than writing, and it appears almost exclusively in a dialogic context (McEnery & Xiao 2004). Swearing has also been shown to have social functions, for example demonstrating solidarity between factory workers (Daly et al. 2004) and alignment between friends (Bernardi 2012), and even contributing to leadership function in the high-swearing context of a rugby team (Wilson 2012).

Beers Fägersten (2007) contrasts frequency studies, which demonstrate comparatively high occurrence of swearing, with offensiveness studies, which generally rate certain words as extremely offensive. They found differences between different words in the set under consideration, and additionally, differing perceptions between participants. In some cases, judgements of offensiveness differed along demographic lines (e.g. race or gender). Multilingual speakers have reported that in spontaneous production they prefer L1 swearwords, and therefore expletives may not match the surrounding language (Dewaele 2010).

2.3.3. Linguistic Accommodation

So far, we have limited our consideration of linguistic choices to absolutes: what is the prevalence of this or that feature? What are the linguistic correlates of a particular trait or context? But we may also find variation in *relative* linguistic behaviour to be informative. In particular, Communication Accommodation Theory suggests that individuals adapt their language to be more or less like their interlocutors, with predictable social outcomes (Giles et al. 1991).

Linguistic accommodation is the process whereby speakers adapt to more closely mirror those with whom they communicate. Mimicry appears to be a universal human tendency, so it is natural that this would also apply to language. The original studies of linguistic accommodation dealt with the concept of accent mobility in speech (Giles 1973), encompassing shifts of speech rate, pitch, and phonology, but the field has subsequently expanded to include lexical and syntactic matching. In a recent study of existing metrics, Carrick et al. (2016) observe that the concept of mimicry applies at all levels of the Interactive Alignment Model, from situational understanding through to fine details of phonetics and phonology, although there is not yet any widely accepted methodology for working across these various levels.

Giles et al. (1991) differentiate convergence (moving towards one's conversational partner), divergence (moving away) as two distinct strategies with differing purposes and outcomes. A recent meta-analysis (Soliz & Giles 2014) found robust support for the idea that accommodation is associated with positive evaluations of the speaker and the communication quality, and positive outcomes such as compliance and trust. Meanwhile non-accommodation, reluctant accommodation, and avoidant communication were all negatively correlated with positive outcomes to a significant degree.

Convergence is used to emphasise solidarity and similarity, and is generally associated with positive social outcomes such as increased perceptions of warmth, attractiveness, and competence. In a French-English bilingual setting, Giles et al. (1973) found that using one's second language as a form of accommodation was positively received, and frequently led to reciprocal behaviour. Perception of effort was also found to be a significant factor, with non-fluent second language use being the most rewarded by reciprocity. This effect is not limited to linguistic alignment. Similar results have been observed for non-verbal mimicry: Chartrand & Bargh (1999) found that mimicking non-verbal behaviours led to higher ratings of likeability, while Balienson & Yee (2005) show that even automated mimicry by a digital agent is effective at generating positive ratings of likeability and persuasiveness.

By contrast, divergence tends to be used to highlight social differences, with more negative social outcomes. Researchers in accent mobility experiments were able to generate divergent behaviour by emphasising and threatening national identities, for example an English-accented researcher asking insensitive and critical questions about the Welsh language led to a broadening of Welsh accents and an increase in Welsh vocabulary being used in responses (Bourhis & Giles 1977). Giles & Gasiorek (2014) propose that listeners' inferences about the speaker's intentions and motives will affect their perception of non-accommodative behaviour: although divergence and over-accommodation are generally received negatively, a listener may be more forgiving if they believe the speaker to have had a good reason, such as interrupting in order to convey important information rather than out of rudeness.

Bradac et al. (1988) consider the interaction of convergence and divergence with the status afforded to the base values of the features being studied — in this case, that of lexical diversity, where higher diversity is generally associated with more positive impressions. They generated a series of artificial dialogues, and manipulated the level of convergence or divergence by one of the characters. Students were asked to rate both their impression of the character, and their perceptions of how his language changed during the scenario. A positive correlation was observed between convergence and perceived status and competence, even when the convergence was to a less-prestigious norm; conversely, divergence was correlated with lower rankings of status and competence.

Similarly, in a job interview context, Willemyns et al. (1997) manipulated interviewer accent and measured the language used by applicants in response. They found that applicants were more likely to converge towards interviewers with a broad Australian accent by broadening their own accents, but that convergence towards a more 'cultivated' accent was not observed. Indeed, for some male applicants, when the interviewer was perceived as having a more cultivated accent and the applicant considered himself to have a naturally broader accent, the end result was linguistic divergence on the applicant's part. These findings demonstrate that convergence and divergence are not consistent for all individuals within a particular scenario, and that other factors (such as gender and social class identity) play a role in mediating individuals' accommodative behaviour.

Job interviews are a setting with a clear power differential, and in general, power has been observed as an important variable affecting the way accommodation is exhibited in speech. In one of the earliest studies in this area, Taylor et al. (1978) examined linguistic accommodation in a bilingual French-English factory setting, finding that individuals were more likely to converge to the language of those who were above them in the organisational hierarchy: foremen converged towards managers more than towards workers, while mid-level managers converged more towards their management than towards foremen or ordinary staff. In a university setting, Jones et al. (1999) found that relative social status was the primary determinant of accommodative behaviour in a study of student interactions with faculty and peers, although gender and ethnicity were also influential factors.

Legal contexts also tend towards inherent power imbalances. Gnisci (2005) considers the case of hostile examinations in a courtroom setting, studying a set of 47 transcripts from a single criminal case, which was a complex financial and political case. This study finds that witnesses and lawyers both use accommodation and maintenance strategies, with lawyers using maintenance more often, which is in accordance with their comparatively higher status in court. A later study, looking at transcripts of U.S. Supreme Court arguments (Danescu-Niculescu-Mizil et al. 2012), found that immediate lexical co-ordination was correlated with social power, with lawyers co-ordinating to Justices more than the other way around.

Linguistic accommodation can also have measurable economic impact. Van Baaren et al. (2003) conducted a study in a restaurant setting to measure the effect of linguistic mimicry by waitresses, where the level of tips was used as the success metric. Across two experiments, direct verbal mimicry (repeating the order verbatim back to customers) resulted in higher tips, both compared to the control (non-mimicry) condition and compared to the tipping baseline. In a retail environment, Jacob et al. (2011) found that salespeople using verbal and non-verbal mimicry increased the sales of MP3 players from 61.8% to 78.8%, and also increased the chances that the buyer would select the specific model recommended to them. Similar results have been observed in cross-cultural contexts, such as van den Berg's (1986) study of salespeople's linguistic accommodation in Taiwan.

In another experimental manipulation, Guéguen (2009) recruited young women to undertake verbal and non-verbal mimicry in a speed dating setting. The ranking given to these women by their male dating partners was then compared between the mimic and non-mimic conditions, with the result that women were ranked more highly, and rated more highly for sexual attractiveness and quality of interaction, when they had been deliberately mimicking their partner. Van Baaren et al. (2004) demonstrated that being mimicked by an experimenter led to more positive social behaviours, such as helping others and donating to charity.

Language Style Matching (LSM) is a commonly-used metric of linguistic co-ordination which measures the similarity (or distance) between two texts, as expressed through the frequency of function words (Gonzales et al. 2010). In contrast to experimental manipulation of mimicry by confederates, this is a measure of the naturalistic co-ordination occurring between two (or more) individuals. This technique has been applied across a variety of subject areas, from predicting the success of romantic relationships (Ireland et al 2011) to understanding small group dynamics (Gonzales et al. 2010) to analysing police interrogations (Richardson et al. 2014).

Gonzales et al. (2010) introduce LSM as a measure for calculating linguistic similarity based on nine categories of function words: auxiliary verbs, articles, common adverbs, personal pronouns, indefinite pronouns, prepositions, negations, conjunctions, and quantifiers. They use this metric to study two different conditions of group interaction, one face-to-face and one via an online chatroom discussion, where participants were placed in groups of 4 to 6 people and asked to collaborate on a problem solving task. LSM was found to predict self-reported group cohesiveness, with a positive correlation between linguistic similarity and self-reported cohesiveness; however, for task performance (measured by the proportion of questions answered correctly), LSM was found to be predictive only in face-to-face interactions, with a statistically insignificant negative correlation observed in the online condition.

Ireland et al. (2011) conducted two separate experiments into LSM as a predictor of romantic success, looking at relationship initiation and stability. In the first experiment, volunteers took part in a speed dating scenario, in which their conversations were transcribed and analysed; similarity in linguistic style was found to be a predictor of whether the participants in a speed-date mutually requested to see one another again. (Unlike in Guéguen's (2009) study, neither party was given instructions or information about linguistic mimicry.) The second experiment focused on existing relationships, and couples were asked to provide 10 days' worth of instant message chat logs for comparison. Again, LSM was found to be predictive of relationship success, with linguistically-similar couples proving more likely to be together three months after the initial data collection.

Richardson et al. (2014) examined the content of 64 police interrogations, using LSM to measure the similarity between linguistic style of interrogator and suspect. Additionally, they adopt a turn-by-turn variant of LSM to gain a more nuanced understanding of the conversational dynamics. They did not find any significant relationship between LSM and interrogation success (defined as achieving a confession), but turn-based LSM showed that during interrogations leading to a confession, suspects accommodate towards the style of the interrogator, more so than interrogators accommodate to suspects. In cases with no confession, no such distinction was observed. This study used the same nine-feature set as the previous studies (Gonzales et al. 2010, Ireland et al. 2011), but found that matching in four of the features was particularly predictive: auxiliary verbs, quantifiers, prepositions, and personal pronouns. These features were also used more, in absolute terms, by interrogators in those interviews which resulted in a confession.

Recently, as online media have become more prevalent, scholarly attention has turned towards measuring linguistic similarity in textual dialogues. The question of accommodation in online forums is examined by Nguyen & Rosé (2011), who look at the language changes over time as new users converge on forum norms. They measured the adoption of forum-specific slang and matching the informal, emotionally-involved style of the forum under consideration. They found that linguistic accommodation was a good predictor of whether a user would remain active over the period. In particular, longer-term forum users tended to use more abbreviations and more inclusive language. As the forum under consideration was a breast cancer discussion group, this study focused on a rather limited demographic, but it would not be unreasonable to assume that similar phenomena would be observed in other online communities — just as they are in speech.

Niederhoffer & Pennebaker (2002) look at linguistic similarity across a range of LIWC categories, in two different instant message conversation experiments (45 minute versus 15 minute exchanges). Measuring language use turn by turn, they observe that participants across both experiments tended to converge in their use of language, that is to say, the use of a particular feature by one party led to an increased likelihood of the same feature being used in response. Niederhoffer & Pennebaker also recorded both self-reported and independent judgements of interaction quality, however they found that there was no correlation between level of linguistic similarity and level of 'click' — the only exception being the case of positive emotion words, where it is more likely that these words were noticed by judges and taken as evidence of a positive interaction. Based on these results, the authors question the link between mimicry and positive results, hypothesising that a poor-quality interaction may still exhibit co-ordination of negative behaviours (for example, two people growing increasingly short with one another).

Riordan et al. (2013) conduct two studies into accommodation of message length and response latency in instant message conversations. In their first experiment, volunteer participants interacted with a stranger to debate a medical topic. A confederate participant was used to manipulate variables such as disagreement and non-verbal cues (such as innovative punctuation and capitalisation), but was unaware of the convergence aspect of the study. Overall, convergence was observed in both length and latency, although this effect grew less as more turns were exchanged, and there was less convergence in the disagreement condition. The second study asked pairs of existing friends to engage in task-based or social dialogue via instant message. Again, convergence was observed under both conditions, although in this case the rate of convergence increased as the discussions progressed, which may be a function of the participants being previously acquainted.

Huffaker et al. (2011) study the effect of linguistic accommodation on negotiation outcomes in groups, using information entropy to calculate language convergence at the word level. Groups of three participants were given a negotiation task to complete in an online chat room, with an unusual setup where individual private chat was also possible between each pair of particpants. The authors found that greater language convergence was strongly predictive of agreement in the negotiations. Similar results are reported in Taylor & Thomas' (2008) study of nine lengthy hostage negotiations, where higher linguistic similarity was found to correlate with successful negotiation outcomes.

Bunz & Campbell (2004) examine accommodation of politeness indicators in particular, in the context of email. They examined 121 messages, elicited in response to an email request. The data was elicited from students, in response to a message from a fictional visiting professor, so there was a constant power differential. The professor was given the neutral name of Chris, to try and eliminate the effects of any possible gender bias (although the degree to which this is likely to be effective may depend on the gender balance in the particular field of study). Four different versions of the original request were formulated, consisting of versions both with and without explicit politeness indicators, and explicit greetings/closings. The data was examined for any evidence of accommodation in the responses. Their results showed that participants accommodated in both cases: including a higher proportion of greetings in response to messages with greetings, as well as echoing politeness markers in the text.

Danescu-Niculescu-Mizil et al. (2012) examine linguistic co-ordination (specifically, of function word frequency) within Wikipedia discussion pages. They look specifically at ‘immediate’ coordination, how a speaker’s function word selection echoes that of the immediately preceding turn. Although Wikipedia is an egalitarian concept, there is an innate social hierarchy between the public (who can make basic edits) and those with administrative power; another interesting feature of this dataset is that users can be promoted to admin status via a community vote, so in many cases there is data from both before and after a user was promoted, as well as data from those who asked to become admins but were refused. This study finds a correlation between linguistic co-ordination and social power, with lower-status individuals exhibiting more linguistic co-ordination towards admins than towards non-admins. However, administrators accommodated more than theory might predict, and there are also some more nuanced findings regarding promotion to an admin role. In particular, those who were successfully promoted exhibited the same linguistic convergence behaviour before and after their appointment, indicating that higher levels of accommodation might have been a contributing factor to their success; conversely, those who tried but failed to secure promotion tended to have a less accommodative style.

Within the theoretical framework of Communcation Accommodation Theory, the psychological motivation for convergence is understood to be the drive for approval and affiliation (Giles et al. 1991). It follows that individuals with a greater need for social approval would be expected to exhibit greater convergence; in an early study, Natale (1975) found that this held true for convergence of temporal behaviour in spoken conversation.

Despite this theoretical grounding, the relationship between different personality types and accommodative behaviour has received only limited experimental validation to date. Ireland and Pennebaker (2010) report that people who scored highly for neuroticism tended to exhibit less style matching behaviour, while those with low conscientiousness and high extraversion tended to score more highly, although these correlations were weak. Kurzius (2015) tests the relationship between the OCEAN ‘Big Five’ personality traits and speech rate convergence, finding that extraversion and openness were significantly correlated with convergence. Self-monitoring is a personality trait concerned with an individual’s tendency to monitor their own behaviour and adapt their behaviour to different social contexts. Cheng & Chartrand (2003) found that high scores for self-monitoring were correlated with higher degrees of behavioural mimicry, and that this interacted with social position: in particular, those with high self-monitoring were more likely to engage in unconscious mimicry of a peer or someone of higher status, compared to when they interacted with someone of lower status. Chartrand & Bargh (1999) studied the impact of empathy on nonverbal mimicry, finding that the cognitive, deliberate aspect of empathy (perspective taking) correlated with increased mimicry, while there was no correlation with emotional feelings of empathic concern. These studies all indicate that individual personality is likely to interact with other variables of interest (such as social power) to affect the likelihood of accommodating within a conversation.

2.4. Reconstructing Social Hierarchies: Previous Work

Having examined some of the ways in which linguistic choice may be used to express group membership and social roles, we shall now go on to consider how this relates to the task of reconstructing social hierarchies.

Social network analysis (SNA) is a well-established discipline in sociology and latterly computer science. Most SNA approaches use graph theory to identify the most ‘central’ nodes in a network, or those ‘bridge’ nodes which connect two or more components of a graph. It is a matter of debate the extent to which such structural features map onto organisational (as opposed to informal, socially-defined) roles, but these techniques have generated some interesting results in this area, so it will benefit us to briefly consider them.

More recently, a small number of studies have considered the use of linguistic cues to social status, whether by measuring absolute feature scores (Bramsen et al. 2011) or relative features such as linguistic co-ordination (Danescu-Niculescu-Mizil et al. 2012).

This previous work gives a solid foundation upon which to build, and is a good reason for selecting a similar task. We will briefly examine some of the highlights of earlier work.

2.4.1. Social Network Analysis

Social network analysis (SNA) originated in sociology as a method of studying the social structure of groups using graph theory. As this attempts to solve the same problem that we are approaching, from a very different set of features, it will be valuable to consider some key results in this field to give a broader context to our work. From karate club friendships (Zachary 1977) to academic collaborations (Newman 2001), almost any form of social contact can be represented as a graph and studied accordingly; in communications studies, this has often taken the form of studying communications graphs to infer social conclusions (Gallagher 2010; Karagiannis & Vojnovic 2009; Rowe et al. 2007).

In a typical social graph, nodes of the graph represent individuals, and the presence of an edge between nodes indicates a relationship between those two people. Figure 1 gives an example of a simple graph. For example, in a graph of academic collaborations, a relationship of interest might be co-authoring a paper, where nodes would be authors and links would connect those authors who have co-authored. Co-authorship is a common topic of study, since academic authorship data is so easy to acquire.

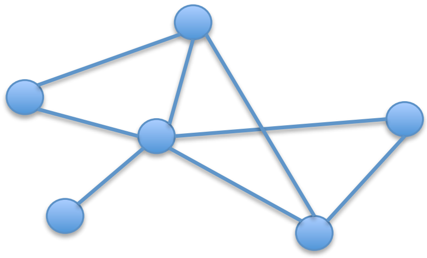

Figure 1: A simple example graph (unweighted, undirected)

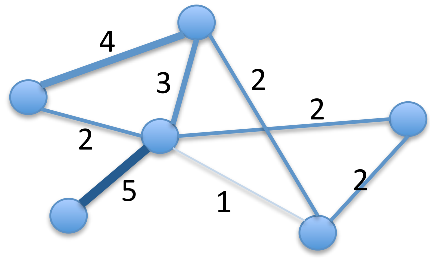

Figure 2: An example graph with weighted edges (arbitrary weights added to the edges of Figure 1).

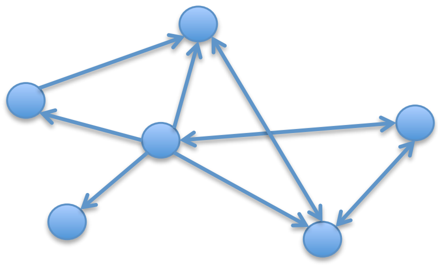

Figure 3: A directed graph (unweighted), generated by adding direction to the edges of Figure 1.

Of course, knowing that two authors once wrote a paper together is one thing; knowing that they have co-authored twenty papers gives quite a different perspective on the relationship. Weighted graphs allow for this kind of distinction to be captured by adding a weight to an edge. In the collaboration network, for instance, this could be the number of co-authored papers. Figure 2 gives an example of a weighted graph.

A collaboration network is undirected, which is to say that the edges simply connect two nodes, and the relationship is always symmetric: if A collaborates with B, then B collaborates with A. By contrast, consider a related academic graph, that of citations. The citation network is directed: A can cite B without B having even heard of A, let alone citing in return. Additionally, a directed graph can be weighted or unweighted, just the same as an undirected graph. Figure 3 provides an illustration of a directed graph.

Once a suitable graph has been constructed, a number of common algorithms can be used to identify properties of the graph which map to real-world correlates. Network density refers to the level of interconnectedness of the nodes, measured as the ratio of edges to possible edges (Wasserman & Faust 1994). For directed graphs, the degree of reciprocity, the proportion of edges A-B which have a corresponding edge B-A, may also be interesting (Garlaschelli & Loffredo 2004). Nodes may be notable because they are highly connected (degree centrality, Freeman 1979), or central to the graph (closeness centrality, Sabidussi 1966), or because they form a bridge between otherwise disconnected components (betweenness centrality, Freeman 1977). In the case of social networks, any of these features might reasonably equate to some form of social influence. Cliques, defined as sub-graphs in which every pair of nodes is linked by an edge, have an obvious analogue to the real-world phenomenon of the same name. Nodes which bridge two components of the graph may represent individuals with higher social mobility, or with a broader range of contacts and interests. And nodes with a disproportionate number of links may represent particularly gregarious individuals, or those whose work involves a significant component of professional networking.

Latterly, one common application of social network analysis has been the automated analysis of online communications networks on the basis of communication metadata: who contacts whom, with what frequency, and following what patterns? These are large and richly-featured graphs which have provided a productive seam of research (Gallagher 2010; Rowe et al. 2007).

When applied to the task of hierarchy detection from communications graphs, SNA uses graph-based algorithms and a combination of features including strength of ties (e.g. number of messages sent between a pair), initiation of threads and topics, and time to respond.

Rowe et al. (2007) infer a hierarchy over the Enron dataset using network analysis methods. Their algorithm includes the number of emails sent and received by a user, as well as the average response time for their messages, and graph features such as involvement in cliques (maximal complete subgraphs). Using these features, they generate a social importance score for each individual, by which users can be ranked.

They found that their techniques were particularly good at identifying the individuals at the top of a hierarchy, but less effective at replicating the published structure further down the organizational chart. This may be because the behaviour of those in lower-ranked jobs is more likely to diverge depending on differing job requirements.

For one example which is examined in more detail, they find that the head of a specific division is automatically identified as its most important member, but that the two next-most important are his administrative assistants. This may not reflect a traditional view of the organisational hierarchy, but represents a plausible model of information flow, if the administration staff act as gatekeepers for their head of division.

Gallagher (2010) applied Bayesian block modelling techniques to a graph with weighted edges, and applied this to the Enron corpus. This method divides the nodes of the graph (individual email participants) into groups according to their 'role' which is defined according to their patterns of interaction.

These roles are seen to group people into clusters which, at least in some instances, appear to correlate with their job titles or rank. The interactions between roles are also interesting, showing how different communities within Enron interact with one another in different (but often predictable) ways. As an unsupervised approach, this method exposes interesting structures within the network, without promising to group or rank staff in any particular manner.

Karagiannis & Vojnovic (2009) studied the enterprise email logs of a large organization (over 100,000 employees) over a three month period. This gave a sample size of 315 million emails. They examined reply behaviour, finding that replies constitute only a small proportion of messages sent. It follows that most messages do not receive a reply. They were able to identify several factors that appeared to influence whether a reply was sent. Messages with fewer recipients in the original message were more likely to receive a reply, as were larger messages (measured in bytes – so this may include attachments and quoted text). The time of day the message was sent was also a factor, along with whether the recipient had been using email recently (a proxy measure for whether they were online). Analysis of variance showed that best prediction of reply probability was obtained by using a combination of these features. Additionally, they found that if any reply was made, it was likely to be made quickly, but that on average recipients were likely to take longer to reply to someone further up the corporate hierarchy, suggesting that concern to send the *right* message outweighs the need to respond with haste when dealing with high-powered individuals.

Inspired by studies of timing in spoken conversation, Kalman et al. (2006) examined response latencies in three separate online datasets: corporate email (Enron), university discussion groups, and a public forum (Google Answers). Although the average response time for each forum was different, they found an underlying power-law distribution of reply times with a consistent pattern. This distribution held across all three datasets, in cases where a reply was sent at all; messages with no replies were ignored for the purposes of this study. When these results were compared to spoken conversations, which obviously have a much lower absolute latency, the pattern was still shown to be broadly consistent. Kalman et al. did not consider the social or organisational roles of senders and recipients in their work, but their findings nonetheless provide some interesting background.

One specific limitation of the Enron dataset is acknowledged: the differing timestamps on users' computers are the only source of timing information for the emails, leading to strange effects such as negative latency due to messages sent and received in different time zones. While examples were few (and the negative latencies discarded), it is not clear to what extent the timestamp differences would impact the validity of the results more broadly. The actual difference of time zones might also be expected to affect the absolute response time, as one user might send a message during another's non-working hours. However, the fact that the Enron results are broadly consistent with those from the other datasets implies that this has minimal impact on the overall results.

The average response latency for each dataset falls at or above the 80^{th} percentile, meaning that the majority of messages are sent with a below-average delay. The authors investigated whether this was an artefact of aggregation, by picking out a number of individuals' response latencies for comparison, and found that for most individual users the pattern of replies also followed a similar power law distribution.

This study supports Karagiannis & Vojnovic's finding that responses are more likely to occur in the period shortly after the message is sent, and shows that this pattern holds across a variety of online settings. The authors suggest a number of possible explanations for this phenomenon, without proposing methods to distinguish them.

2.4.2. Linguistically Motivated Modelling

Automatic classification of relationships based on message content is a field in its infancy, in comparison to metadata-based approaches. The term 'Social Power

Modelling' was coined by Bramsen et al. (2011) to describe the task of identifying relative social status based on language use.

Bramsen et al. (2011) trained a set of classifiers on the Enron corpus, to differentiate 'UpSpeak' (messages addressed up the hierarchy) from 'DownSpeak' (messages going down the hierarchy). They experimented with a variety of n-gram models, using unigrams, word bigrams, and part-of-speech bigrams. They also tried a variety of different classifiers, obtaining best results using SVMs. Training data was partitioned by author to avoid picking up on features due to idiolect; without this partitioning, they found that accuracy was artificially inflated due to effectively training and testing on language from the same individuals. Their best result on the partitioned dataset was obtained by using n-grams binned into sets, with information gain (WEKA's InfoGain implementation) used to filter out those sets which did not contribute much to the overall classification. This achieved an F-measure of 0.781 using a weighted test set. Without partitioning, their F-measure for 10-fold cross validation was 0.830.

Although Peterson et al. (2011) did not attempt to classify relationships based on their findings, their qualitative work provides a powerful indication that such classification should be possible using features related to linguistic formality. They found that messages going up the hierarchy are likely to be more polite than those going downwards.

Gilbert (2012) examines the language of the Enron corpus and uses statistical techniques to pick out key phrases which indicate hierarchical relationships. To avoid any influence from the unusual circumstances of the company's collapse, he uses a cut-off date of May 2001, six months before the investigation launched, and does not consider any later messages.

Gilbert's phrase model consists of unigrams, bigrams & trigrams. The model size is reduced by the deletion of any phrases which consist only of stopwords, and (to avoid undue influence from any one individual) deletion of any not used by at least three different message senders. Using penalized logistic regression and an SVM classifier, phrases indicative of 'upward' communication are identified. Of the 7,222 phrases used to construct the model, 974 have a statistically significant contribution to discrimination.

The paper contains a table of the hundred most discriminative phrases, encompassing a plethora of semantic spheres. Phrases particularly indicative of upwards communication include positive words ('tremendous', 'excellent', 'sounds good'), lots of neutral, functional nouns ('the report', 'worksheet', 'final draft', 'candidate'), and examples of indirect politeness ('can you get', 'can I get'). Contrasting this with the phrases more likely to be seen in downward or level communication, we see informality ('cool', 'man', 'fyi'), direct requests ('please send', 'give me',), and negativity ('forgot to', 'problem with', 'the confusion').

Gilbert also reports that fewer misspellings were observed in the 'upward' dataset, possibly indicating that individuals proof-read their own words more carefully when addressing their seniors; this is not developed more fully in the paper due to the limitations of n-gram language modelling.

Agarwal et al. (2012) contrast a content-based method with a social network analysis approach. For their simple SNA technique, they use degree centrality across an undirected, weighted graph of email contact (without distinguishing To, CC, and BCC fields). On their dataset, Agarwal et al. found that the graph approach outperformed an NLP approach based on Gilbert's (2012) trigram phrases.

However, one problem with using degree centrality, in the case of the Enron graph, is that degree will almost inevitably be higher for individuals from the 'core' group (those whose emails were subpoenaed) compared to others, and that the 'core' group tend to also be of higher rank. It follows that degree centrality may be expected to correlate with hierarchy more strongly in the Enron corpus than in a randomly selected email dataset.

Prabhakaran & Rambow (2014) attempt to address the problem of data sparsity by making a social power prediction based on a single 'thread'. They use a mixture of metadata and content features to predict the relationship between sender/recipient pairs within a thread, without concern for the total number of recipients involved in the conversation.

Metadata features used include the number of participants in a thread, who initiated each thread, the number of messages in the thread, average message length (in words), and the ratio of sent to received messages within a pair. Dialogue features include requests for action, requests for information, and overt displays of power as identified by an automated tagger. Against baselines of 52.54% (most common class) and 68.56% (bigram model), they found that dialog acts and structural features alone gave an accuray of 62.47%, while their best result of 73.03% accuracy was obtained from a combination of n-grams with a subset of structural features.

2.5. Discussion

In the literature of social identity (Tajfel & Turner 1979) and self-categorisation (Turner et al. 1987), framing (Goffman 1974) and face threat (Goffman 1959), we have found no shortage of evidence that individuals are social creatures, whose behaviour is constantly influenced by their relationships to those with whom they interact. Linguistic theory has subsequently set out criteria for politeness and appropriateness in interactions (Grice 1975; Brown & Levinson 1987), again informed by the speaker's relation to the audience.

Recent experiments in computational linguistics have examined the correlations between (im)polite language and social status (Peterson et al. 2011; Gilbert 2012) and given rise to predictive models based on linguistic features (Bramsen et al. 2011; Agarwal et al. 2012; Prabhakaran & Rambow 2014). Our work is motivated by a wish to improve upon this body of work by investigating what is possible using a much smaller, more tightly constrained feature set to build a model with more explanatory power. We therefore construct our primary hypothesis around the theory that stylistic choice, in particular, is a key method of expressing relationship-building, and we wish to test whether this enables us to build a predictive model based only on a small set of linguistic style features.

Our feature selection is inspired by work in authorship identification and author profiling (Koppel et al. 2002; Argamon et al. 2007; Rosenthal & McKeown 2011; Kacewitz et al. 2013), which is a field with a long history of using stylistic features to characterise individuals and demographic groups. We have also taken inspiration from qualitative studies into power-differential behaviour in email (Panteli 2002; Waldvogel 2007; Murphy & Levy 2006).

Additionally, our work on linguistic accommodation sits within the framework of communication accommodation theory (Giles 1973). We have seen that accommodation of vocabulary can predict continued engagement with online forums (Nguyen & Rosé 2011), as well as correlating with social power in Wikipedia discussions and courtroom settings (Danescu-Niculescu-Mizil et al. 2012). Bunz & Campbell (2004) demonstrated that students would accommodate their use of politeness indicators to an email from a professor, another instance of unequal social power. We are inspired by the success of approaches using function word distributions, via language style matching (Gonzales et al 2010), to capture important facets of relationship negotiation behaviour across contexts as diverse as speed dating (Ireland et al 2011) and police interrogations (Richardson et al 2014), but we recognise the inherent limitations of the currently predominant LSM approach, and will set out a more sophisticated technique using movement in vector space to capture nuances of accommodation as a dynamic process.

Chapter 3 lays out the details of our implementation, as informed by this prior work.

3. Methodology

Having considered the theoretical background, we will now turn our attention to the more mundane (but equally important) concerns of methodology and implementation.

Our primary hypothesis is that stylistic choice is a key method of expressing relationship-building, and therefore it should be possible to infer relationship information on the basis of linguistic style. In order to test H1, we adopt the task of classifying messages based on the hierarchical relationships between sender and recipient. This is a subset of the generalised relationship categorisation task, with the advantage that some organisational ground truth data is available (such as the Enron corporate hierarchy). Social power is also a facet of relationship-building which is particularly susceptible to experimental manipulation: unlike longstanding friendships, romantic, or familial relationships, individuals are accustomed to finding themselves in new relationships of social power throughout the course of daily life, and continuously adapt their behaviour accordingly.

Having settled upon this task, our first requirement was to identify a suitable set of data. The Enron corpus (Klimt & Yang 2004) was an obvious starting point, as it is the standard used in previous research (e.g. Bramsen et al., 2011), and some information on the corporate hierarchy is freely available. There are not many other sources of readily-available data in which both hierarchical and peer-level communications can be observed: in most data where there are clear power differentials, such as court transcripts, there are few opportunities for peer-to-peer communication, and conversely, in online environments where peer communication is common, the sources of power and authority are less well defined. We therefore complement the Enron corpus with two sets of transcripts gathered under experimental conditions, in which we deliberately manipulated social roles within the experimental set up (Muir et al. 2016a; Muir et al. 2016b). These consist of one set of transcribed speech, from face-to-face interactions, and one set of instant message transcripts from a computer-mediated environment. Data collection and preparation is described in more detail in section 3.1.

Once the data has been gathered and appropriately cleaned, the next stage is to extract features to be used in the analysis. Appropriate features are selected from the literature; standard feature extraction toolkits exist for some cases (such as parts of speech), but for many features of linguistic innovation there are no existing open source libraries. Feature selection is described in section 3.2, and section 3.3 describes the per-author standardisation we will undertake to test H2, our hypothesis that measuring the range of variation within an individual's behaviour will give a stronger signal than utilising only the differences between people.

Machine learning is used to construct the classifiers we will use to test our hypotheses. Assessing different machine learning methods is not a focus of this work, but in order to use existing tools appropriately we must understand their relative merits; section 3.4 describes our approach.

We then wish to examine some possible methods of improving our classification accuracy by using all of the available information in concert. We hypothesise that prediction of hierarchical relationships will be more effective at the pairwise level than at the level of individual messages (H3) and that classification will be aided by consideration of whole-network characteristics (H4). Sections 3.5 and 3.6 describe our approaches to these tasks.

Finally, we wish to investigate the impact of social role on linguistic accommodation in order to test H5. We define a novel metric for measuring linguistic style accommodation using a vector space model, described in section 3.7, and present the results of our initial experiments using this technique.

This chapter goes into more detail on these aspects of implementation.

3.1. Data Selection and Preparation

In selecting appropriate datasets for research, a number of factors must be considered. In this instance, as we will be endeavouring to use stylometric measures to infer hierarchies, we need to reduce the interference of other factors.

External factors, including the medium being used, will have an impact on the construction of a message, as discussed in section 2.2. The same essential content may be conveyed differently in different contexts. For example, Duthler's (2006) comparison of requests showed that different politeness features are exhibited in email and voicemail contexts, even when other variables are controlled.

There are several distinct axes along which the 'environment' of a communication can be categorised, any of which may have an impact on the style used:

- Is the discussion space public (e.g. a web forum) or private (e.g. email)? (Hill 1995)
- How many people are involved? Communications can be one-to-one (personal), one-to-many (broadcast), many-to-many (group discussion), or many-to-one (petition). (Bell 1984; Peterson et al. 2011)
- Is the medium spoken or written? (Baron 1998; Dulther 2006)
- Is communication real-time or asynchronous? Historically, spoken media have been synchronous and written media have been asynchronous, but this is not inevitably true: voicemail messages are asynchronous, while online, text-based chat can be instant. (Sotillo 2000)
- Are participants geographically co-located or distant? (Doherty-Sneddon et al. 1997; Setlock et al. 2004)
- Is the interaction method dialogic (interactive, encouraging response and exchange of views, such as an email, discussion forum, or conversation) or didactic (imparting information or opinion with no expectation of response, such as a poster or prepared speech)? (Tannen 1984; Kankaanrata 2005)
- Are there any constraints on the message format or device used? For example, length restrictions of Twitter and SMS, errors introduced by predictive text. (Gouws et al. 2011)

To reduce the amount of interference between these various factors, it is important to select a broadly homogenous dataset. Additionally, in order to ensure we are not over-fitting our techniques to one particular circumstance, we will use three independent sources of data, from different production contexts and involving different participants. In all contexts we decided to limit our study to personal communications, in dialogic one-to-one settings, in order to focus on the impact of interpersonal relationships on language use without the complication of audience size or bystander effects.

The Enron email corpus (Klimt & Yang 2004) is frequently used in communications research, which makes it a good starting point for comparison with earlier work. Email is always written and private, unconstrained in length, and generally assumed to be asynchronous.

The Muir-Joinson Speed Networking Speech corpus (Muir et al. 2016a; Muir et al. 2016c) is a set of transcriptions from a speed networking business scenario. The data differs from the Enron corpus in a number of important dimensions: as well as being captured under experimental conditions, the data is spoken, synchronous, and face-to-face.

The Muir-Joinson Speed Networking CMC corpus (Muir et al. 2016b; Muir et al. 2016c) is a set of online chat transcripts, generated by re-running the same speed networking experiment using a computer-mediated environment.

A summary of key differences and similarities is as follows:

<table>
<tr><th>Enron Emails</th><th>Muir-Joinson Speech</th><th>Muir-Joinson CMC</th></tr>
<tr><td>Textual data from emails</td><td>Transcribed speech</td><td>Chatroom transcripts</td></tr>
<tr><td>Naturally occurring content</td><td colspan="2">Data elicited under experimental conditions</td></tr>
<tr><td>Unknown and varied production environments</td><td colspan="2">Controlled, consistent production environment</td></tr>
<tr><td>Genuine and involved corporate hierarchy</td><td colspan="2">Simple hierarchical condition created by role-playing scenario</td></tr>
<tr><td>Messages of varying length, no constraints</td><td colspan="2">Turns tend to be short utterances; overall time constraint for the exchange</td></tr>
<tr><td>Messages subpoenaed months or years after creation</td><td colspan="2">Participants knew in advance that their speech would be recorded</td></tr>
<tr><td>Authors from across the Enron community, varying in age and seniority</td><td colspan="2">Participants were mostly students, of similar age and educational level</td></tr>
<tr><td colspan="3">Varying nationalities and linguistic backgrounds</td></tr>
<tr><td>Potential mixture of dialogic and didactic messages</td><td colspan="2">Dialogic setting</td></tr>
</table>

The following sections contain more detail on the collection, cleaning, and preparation of each data set.

3.1.1. Enron Corpus Subset

The Enron email corpus is a standard dataset for communications research, and as such it has been widely studied by a large number of scholars and for a variety of purposes. Enron email is an obvious starting point for sociolinguistic research, if only due to the extensive work – both qualitative and quantitative – that has already been conducted using this data, giving a sizeable body of earlier work for comparison. Although few authors have explicitly examined the interaction of style and power within Enron, a number (such as Bramsen et al. 2011; Gilbert 2012) have addressed closely related problems.

We used the CMU version of the corpus (Klimt & Yang 2004), which has undergone some removal of duplicate messages. This corpus contains around 200,000 message files from the mailboxes of 158 Enron employees. We undertook further filtering and processing of the messages in order to produce a suitable dataset for this study, for example limited reconstruction of messages from quoted text, and deduplication of identical messages which are sometimes included twice due to the complex folder structures. We stripped headers, quoted text, and other non-relevant content to leave only the new linguistic content of each messsage, using an SVM zoning approach as per Lampert et al. (2009).

In order to focus directly on the personal relationship between author and recipient, we limited our consideration to single-recipient messages. Previous work has shown that message formality changes in line with number of recipients (Peterson et al. 2011), which makes it harder to infer the nature of one-to-one relationships from messages with multiple recipients. We therefore discarded any message with more than one recipient, even when these were in the CC or BCC fields. Although not a perfect filter, this should also increase the percentage of dialogic messages, as didactic use of email tends to take the form of newsletters with a wide distribution.

Aside from a large volume of messages, the other crucial requirement for this study was some ground truth data on the underlying organisational hierarchy. Information on the position of individuals within the Enron hierarchy was derived from Peterson et al.'s (2011) work. Individuals are labelled with a rank from 0 (employee or trader) to 4 (CEO). Peterson et al. discarded all messages involving rank 0 individuals due to the diversity of roles encompassed by this label, however we chose to include them for the sake of retaining a larger dataset. We removed only the small number of individuals identified as the in-house lawyers, whose language has previously been shown to be atypical (Wright & Gomez 2012).

In order to examine the influence of relative power on communicative style, we used only messages where both sender and recipient were of known rank according to Peterson et al. Our reduced dataset consists of messages from 175 senders, addressed to 140 recipients. These individuals are distributed across the hierarchy as follows:

Rank 0	73
Rank 1	17
Rank 2	10
Rank 3	63
Rank 4	12

Rank 0 represents ‘ordinary’ employees, while rank 4 consists of CEO-level upper management. This distribution is clearly not representative of the true distribution of roles in the organisation. Sampling is biased by inclusion in the original corpus only of those individuals considered to be of interest to the investigation into Enron’s malpractice, and those with whom they directly communicate. Such persons of interest are more likely to be those who had more power within the organisation. There is consequently a significant under-representation of staff from lower levels of the corporation, in our dataset as in the wider Enron corpus (Shetty & Adibi 2004).

As with any naturally occurring data, attempting to reduce all the complexities of human relationships to a single numbered scale is guaranteed to be overly simplistic. Actual relationships are not defined by a simple position in an org chart: the water may be muddied by friendships, rivalries, and romances. Even within an exclusively business context, the perceived status differences between different departments or specialisms could have an impact. However, we are not able to account for these various nuances in a systematic manner: we can only note their likely influence, and proceed with caution.

Our final dataset contains 5,767 email messages. As this is only a fraction of the original Enron corpus, we will not be able to assume that conclusions would necessarily apply to the whole dataset, particularly as our selection criterion depends on including users whose role is known – and who are therefore likely to overrepresent senior roles in the organisation. However, this size of data could easily be all that was available in a smaller investigation. We label messages as ‘upward’ (addressed to a recipient of higher rank), ‘downward’ (addressed to a recipient of lower rank), or ‘level’ (between two individuals of the same rank). The overall distribution consists of 1916 upward messages, 2374 level, and 1477 downward.

As the next step in our data preparation, we divide the data into 10 static partitions in order to perform 10-fold cross validation in a repeatable manner across a variety of experimental conditions. We use three different divisions of the data. Prior knowledge about an individual’s personal style was hypothesised as likely to be helpful in building a model. In order to test the impact of this variable, versions of the corpus were partitioned in three distinct manners:

- completely at random
- split by sender-recipient pair
- split by sender

In the random case, each message was allocated to one of the ten partitions using a random number generator, meaning that training and test data could contain messages from the same sender, to the same recipient, without penalty. This was the easiest case for the classifier, and the results are expected to be over-optimistic. There are very few real-life situations in which it would be useful to predict a relationship using a classifier built from prior knowledge of that same relationship (exceptions may be found where, unbeknownst to the system, an individual is corresponding from more than one email address).

In the case of partitioning by sender-recipient pair, the folds were engineered so as to avoid the training and test data sets containing any messages from the same sender, to the same recipient: all messages from A to B were assigned to the same fold. This was achieved by generating a list of all sender-recipient pairs, randomly partitioning the list into ten groups, and allocating all messages from A to B into the same fold. Ten-fold cross validation was then conducted using these partitions. This is expected to be the most likely analogue to real-world scenarios, representing situation where the some of an individual's relationships are known, and others unknown.

In the case of partitioning by sender, the most cautious partitioning, the training and test data were guaranteed not to contain any messages from the same sender, regardless of recipient identity. This was achieved by random partitioning of the list of senders. Since some senders are far more prolific than others, this results in highly unbalanced text volumes between the partitions, which was not the case for the other two cases. However, attempting to balance the distribution of senders by message volume has the potential to skew results, as prolificness may correlate with other factors of interest (such as status) and we do not wish to accidentally balance our dataset for these factors. Therefore, we accept that the training and test sets for our partitions are of different sizes in this case, for different folds, and anticipate greater variability in the results between folds on this basis; however, by keeping the partitions constant between different experiments, we can at least ensure that this imbalance is kept constant under different conditions. Results from the sender-partitioned case are expected to be lower than a real-world scenario, as it is unlikely that we would have full knowledge of relationships for a closed set of senders and recipients, and yet wish to use this to model a disconnected-yet-comparable set. (In cases where there is no prior information, there is no ground truth data with which to build a model at all, and supervised machine learning methods are not applicable.)

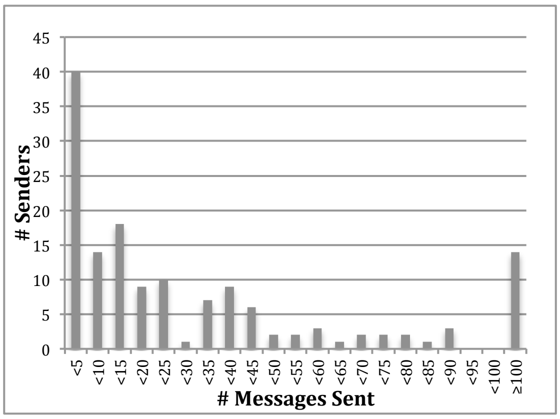

Figure 4: Distribution of messages in the Enron dataset.

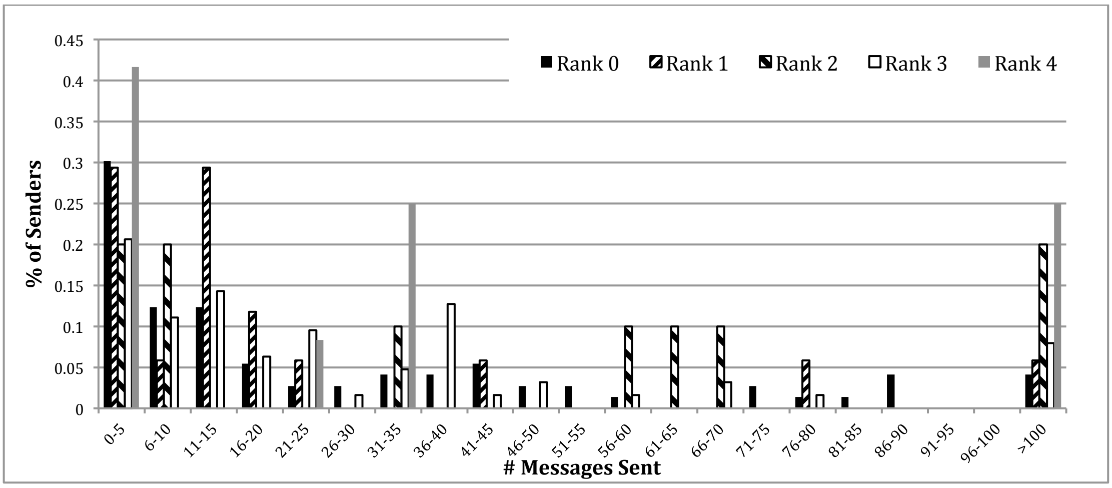

Figure 5: Distribution of messages in the Enron corpus, broken down by sender rank.

It is interesting to note the different levels of productivity by sender (illustrated in Figure 4). Almost a quarter of senders (22.8%) have sent five or fewer messages. At the opposite end of the spectrum, we have a set of fourteen hyper-productive individuals, who have produced more than 100 messages each (up to a maximum of 259 messages from the most prolific sender, who alone is responsible for 4.5% of all messages).

Further, after partitioning the data by the rank of the sender (Figure 5), we are able to see the distribution of senders across ranks. We see that most ranks skew towards the lower end (as with the data overall), but there is no clear pattern explaining the role of high-sending individuals. The most senior individuals, those given rank 4, tend towards extremes: of the 12 senders of this rank, three fall into the most prolific bracket, while five others have sent five or fewer messages.

It is important to note, when studying these distributions, that this is not necessarily representative of these individuals' overall behaviour, as these numbers relate only to the messages in our pruned dataset (single recipient messages, between individuals of known rank). Additionally, as with all Enron statistics, the dataset is skewed by those who communicate with persons of interest to the fraud investigation; in this case we expect this to be particularly biased against the inclusion of more junior staff. We examine these numbers only to inform our later experiments, not to draw any conclusions about the typical message-sending behaviour of individuals of a given organisational rank.

3.1.2. Muir-Joinson Speech Corpus

The Muir-Joinson speech corpus is a collection of dyadic interactions which we recorded and transcribed as part of an applied psychology experiment into power-differential behaviour (Muir et al. 2016a). This dataset consists of a series of one-to-one interactions, with a common task focus, in either the presence or absence of a power imbalance. We have published an anonymised copy of this data via LDC (Muir 2016c) in order to allow for replicability, and for future work to build on our efforts.

We recruited a group of student volunteers and gave them a discussion task to complete, in a simulated business environment. In the experimental group (the hierarchical condition) the participants were divided into judges and workers, and the workers were given ideas to pitch to the judges. Following a brief discussion, the judges could then choose whether or not to invest in each concept. Members of the control group (the non-hierarchical condition) were simply asked to discuss potential inventions between themselves with an eye to potential collaborations, but without either party making an investment decision. In both conditions, participants rotated through multiple conversation partners using a 'speed networking' model to generate a number of independent one-to-one interactions lasting five minutes each. These exchanges were recorded, and the recordings were transcribed.

41 participants took part in the study, and a range of supplementary information was collected on each participant, including demographic information and profiling of major personality traits. Most of the participants (82.9%) were undergraduate students from the University of the West of England. The remainder was made up of postgraduate students and non-students. Participants' ages ranged from 18 to 25. Female subjects made up 70.7% of the population, and 75.6% listed their ethnic origin as British.

The experimental group was composed of twelve participants assigned to the 'judge' role and twelve 'workers'. The remaining 17 participants were assigned to the control condition, and divided arbitrarily into 'A' and 'B' groups. With a couple of exceptions, one interaction was recorded between each judge/worker pair in the hierarchical condition (142 conversations) and between each pair in the non-hierarchical condition (72 conversations). After each interaction individuals were asked to score their communicative partners against a number of axes, from communicative fluency to level of rapport.

The major disadvantage of this dataset is that it does not contain example utterances from the same individual participating under more than one role. A given volunteer was either a judge, a worker, or part of the control group, for the whole experiment. However, the data still exhibits significant stylistic differences between speakers in different roles.

The recorded conversations sum to 13,266 turns. This equates to a mean of 61.99 turns per dyad, μ=59.92 in the hierarchical condition versus μ=66.07 for non-hierarchical condition. The distribution of turns is illustrated in Figure 6. An examination of the distribution of utterances by sender status (Figure 8) shows a tendency for judges to skew lower than workers (although two individual judges are also responsible for the two highest individual turn counts), and for peer exchanges to show lower variance.

As this is a corpus of transcribed speech, it consists of more and shorter turns compared with the Enron corpus. The overall volume of text, however, is much lower (totalling 193,102 words, vs 505,086), and involves fewer individuals; we therefore opted to use five rather than ten folds for cross-validation.

As the data was elicited under controlled circumstances, we have reliable information concerning which participants were assigned to which social roles. The participants did not know one another in advance, so it is not necessary to account for existing social relationships crossing these hierarchical boundaries in unexpected ways. As such, the data should be cleaner than that in the Enron corpus.

The potential disadvantage of the experimental setup is that the participants' behaviour may have been affected by the artificial nature of the setting. However, as we will demonstrate, the data still exhibits significant stylistic differences between speakers in different roles. After the experiment, a manipulation check was conducted by asking participants to score the level of power they felt they had during the interactions: results indicated that judges felt the most powerful ($\mu = 3.7$, $\sigma = 1.1$), while those assigned to the worker role reported lower scores ($\mu = 2.8$, $\sigma = 1.1$), which is significantly different at the 95% confidence level. Interestingly, both non-hierarchical (control) groups rated their perceived power as less than either of the hierarchical groups ($\mu = 1.8$, $\sigma = 1.1$ and $\mu = 2.0$, $\sigma = 1.0$), which may be a consequence of participating in a scenario where their actions were not expected to change any of the outputs: they were not passing judgement on the ideas, but nor were they actively attempting to persuade or influence.

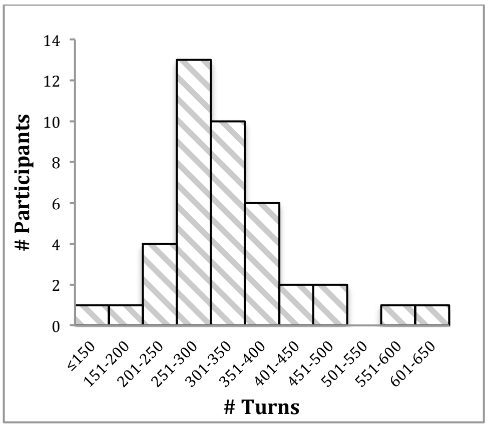

Figure 6: Distribution of turns in the Muir-Joinson speech corpus.

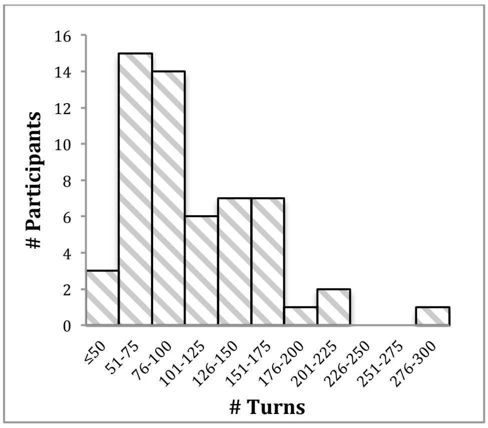

Figure 7: Distribution of turns in the Muir-Joinson chat corpus.

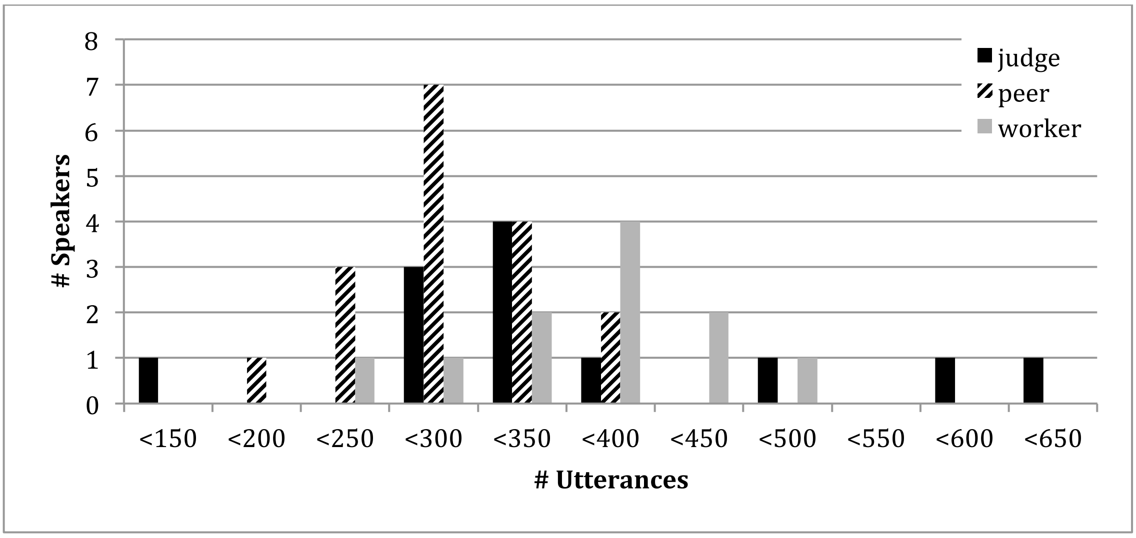

Figure 8: Distributions of turns by status, in the Muir-Joinson speech corpus.

3.1.3. Muir-Joinson CMC Corpus

To generate a third dataset in a different genre, we re-ran the speed networking task using computer-mediated communication (CMC), with participant discussions taking place in an online chat environment (Muir et al. 2016b). We left most of the experimental conditions unchanged, but instead of physically moving around the room, participants were allocated a computer and asked to communicate through the medium of an instant message chatroom, chatting with different participants in turn. In this way we hoped to control for the maximum number of variables, not only in terms of the production environment and experimental power manipulation, but also ensuring common conversational goals and the same topics under discussion. We have published an anonymised copy of this data via LDC (Muir 2016c) in order to allow for replicability, and for future work to build on our efforts.

The volunteers were a different group of individuals to those involved in the speech experiments. Although there would have been certain advantages to having both speech and chat data from the same individuals, the chat experiment was run at a later date, and we could not make use of volunteers who had already been debriefed from the earlier experiment.

A total of 54 volunteers were recruited for the CMC experiment. The participant ages were between 18 and 25 (mean 20.8), and the gender split was 28 male to 26 female (51.8% male). 26 participants were placed in the hierarchical (experimental) condition, with roles of judges and workers randomly assigned; the remaining 28 took part in the non-hierarchical (control) condition in two random groups.

We required participants to complete an assessment of each interaction, as in the face-to-face experiments described above, rating the success of the interaction, and the level of 'click' or rapport with their partner. At the end of the session, a manipulation check was conducted to capture participants' perceptions of personal power. This was highest for judges ($\mu = 4.5$, $\sigma = 0.7$), and lowest for workers ($\mu = 3.5$, $\sigma = 1.1$), with the control groups reporting that they felt almost as powerful as the judges ($\mu = 4.1$, $\sigma = 0.7$ for the first group, and $\mu = 4.4$, $\sigma = 0.7$ for the second). Interestingly, these self-reported scores are all higher than those given by the respective groups in the face-to-face experiment. Personality traits were also recorded for the CMC participants.

One early finding was that online chat generates significantly fewer turns than speech over the same time period: μ=107.6, σ=50.6 for chat, vs. μ=323.6, σ=90.8 for speech. The distribution of turns in the CMC data is illustrated in Figure 7. Typing a given sentence takes longer, for an average typist, than speaking the same words, and there is also the possibility for individuals to take additional time to self-edit in the written scenario. This has the unfortunate consequence that the chat data is less than a quarter of the size (in words) of the equivalent speech dataset, amounting to 45,825 words in total.

Although it is useful to have a short message data set for comparision, it should be noted that this data is not necessarily representative of other short message formats such as SMS and Twitter data, which have been shown to be used in different ways

(Munro & Manning 2012). Given the experimental scenario, our CMC corpus is most likely to be representative of business-focused chat channels such as Slack.

3.2. Feature Selection

Having identified suitable sources of data for testing our hypotheses, we next turn to the task of extracting stylometric features that we hope will, in some combination, provide a strong enough signal to enable classification of a hierarchical relationship, in order to support our primary hypothesis:

> **H1:** Stylistic choice is a key method of expressing relationship-building, therefore relationship information can be inferred from linguistic style.

We reject the default n-gram model for the reasons outlined in section 2.3.1, as n-grams conflate all aspects of language use into a single model, and do not meaningfully capture linguistic innovation or rare vocabulary, which are likely to be important for our task (Eisenstein 2013). We wish to clearly distinguish stylistic from topical (semantic) features, to avoid the influence of subject matter on our classifier; we will achieve this by selecting features that contain minimal semantic information, for example by generally excluding nouns and verbs from our language model.

To enable direct comparison with the Enron and CMC results, we will use the same feature set for speech transcripts, although we acknowledge the theoretical problems with applying textual features to transcribed speech. The concept of a 'sentence' is problematic in speech, and we do not have informal features such as varying capitalization, or emoticons, in speech data. However, we do not have a phonetic transcription of pitch, emphasis, etc. from the original speech recordings: we take the transcriber's choice of punctuation as a proxy, and rely on her consistency as all transcription was undertaken by a single professional. Because the data has been professionally transcribed, there is also less chance of typographical errors — and if such errors do exist, they are due to the transcriber rather than the participant. This reduces the amount of stylistic information available. However, slang terms would be transcribed as heard, which may still result in some out-of-vocabulary items by comparison to a standard dictionary.

3.2.1. Message Length Features

An individual makes many choices during a communication, and choices around language use are not limited to the words used or the syntax through which the message is conveyed. The length of the message itself, and how it's structured, contributes to the overall style and as such may also contain valuable cues.

Message length, measured by the number of characters, is perhaps the most basic of these metrics. More detailed structure can be observed in the number of sentences per paragraph (applicable only to written media), words per sentence (or utterance), and characters per word. Taken together, these scores give a proxy for complexity of language use; these are also components of popular text complexity metrics such as the Flesch-Kincaid readability tests (Kincaid et al. 1975) and the Gunning Fog Index (Gunning 1969).

3.2.2. Character Distribution Features

In written English, orthographic elements such as capitalisation and punctuation can be varied for emphasis without impacting the semantic content of a message. These stylistic choices result in variations of the distribution of various character classes. Character-level features are trivial to compute and (in most cases) applicable to all written languages, and a large number of character-level features were included for this reason.

Formal documents are likely to be in 'sentence case' with normal capitalization. What constitutes a normal distribution of capitals is language-dependent: in English, this equates to capitalization for the first word of a sentence, and for proper nouns.[2] As informality increases, the writer may choose to adopt a more flexible approach. Additional capitals may be used for emphasis (e.g. 'REALLY' or 'Very Nice'), or capitalization may be abandoned altogether in favour of lowercase text which is (fractionally) quicker to type. The proportion of alphabetic characters which are in uppercase was determined as a suitable single metric to capture case distribution.

The situation is similar for punctuation. Formal documents are likely to be punctuated according to the grammatical norms of the language in question: although this will vary by language and genre, and will also depend on individual style, it should be possible to identify divergence from the norm that is due to informality and innovation. Depending on the individual's style, informal documents may contain more symbol characters (due to emoticons or strings of '!'s) or fewer (if punctuation is simply omitted).

Classes of symbol characters were identified and categorised using Unicode categories. As well as the overall proportion of symbols in the text, the distribution of these types was considered.

The following categories of punctuation were specifically counted:

Comma
Full stop
Question mark
Exclamation mark
Colon
Semicolon
Brackets
Quotation marks
Hyphens
Hash symbols
Percent symbols
Ampersands
Currency symbols
Maths symbols

[2] This does not translate directly into all languages: in German, all nouns are capitalized, whereas in many scripts (such as Arabic, Chinese, and Thai), the concept of case does not apply.

For each of these categories, a feature score was calculated as a proportion of symbol characters falling into the specified class.

3.2.3. Lexical Distribution Features

Lexical choice is another obvious manner in which communicants convey their intent. This is also the area in which we must take most care to avoid conflating effects of topical and stylistic features, as individual words (especially nouns) will tend to carry the bulk of topical information.

In addition, some specific lexical choices are frequently indicators of more or less formal language. Lexical features are, for the most part, characterised as addressing the question of whether a given word fits into a given class, however that class is defined. Class frequencies are then calculated as the proportion of such words in the document.

Each document was tokenized using the Stanford toolkit in UIMA (Ferrucci & Lally 2004; Kano et al. 2010). For each word-level feature, each token (word) was then assessed against the relevant criterion to determine whether it is an example of the feature under consideration. In the majority of cases, this determination was by comparison to a lexicon of words in the class; specific feature implementations are detailed below. This feature set assumes that the language of the message has already been determined: for the current project, we assume all data is in English.

Parts of speech are one step away from the surface lexical level, capturing the structure at a higher level of abstraction rather than the distribution of individual words. Previous research has shown that part of speech distribution may vary depending on the register of the communication, and certain distributions may be more indicative of (in)formality. For example, Abu Sheikha & Inkpen (2010) found the distribution of first- and second-person pronouns to be the most important feature in determining text formality, in their study across a variety of media types, with a higher prevalence of these pronouns in informal text. However, their study defined email as informal by default, so their results may not agree with ours. Certain part-of-speech features have also been shown to be productive in studies using the LIWC dictionaries. Part of speech tagging was undertaken using the UIMA (Ferrucci & Lally 2004) framework and OpenNLP (OpenNLP 2011; Kano et al. 2010), using an English model.

F-Score. Heylighen & Dewaele's (2002) F-score is a linear combination of part of speech frequencies, normalized to fall between 0 and 100. If, as might be expected, the F-score is shown to capture the important information about part of speech distribution, then it would naturally be more efficient to simply store this result rather than retaining the individual POS frequencies. F-score is calculated according to the following formula, using information from the POS tags already generated:

$$F = (\text{non-deictic POS frequency} - \text{deictic POS frequency} + 100)/2$$

Out of vocabulary words are defined by comparison to a reference dictionary. Given a set M of the words in a message in language L, and set D which is a dictionary of words from L, there are a number of reasons that a word w may be in M but not D. For example, this may be due to codeswitching or technical jargon or linguistic innovation or simple typing errors. We used wordlists generated from the MySpell dictionary project (MySpell 2008), which is the dictionary used for spell-checking in Firefox. Such dictionaries are available in a variety of languages. Since the Enron data is sourced from a multinational corporation, and study volunteeers came from a variety of ethnic backgrounds, both US and UK dictionaries were used. A word was considered out-of-vocabulary only if it failed the spell-checker test in both US and UK English.

Out-of-vocabulary words were then tested to see if they fit into one of the distinct sub-categories previously identified.

Affective Lengthening is a form of linguistic innovation consisting of the repetition of one or more letters within a word for emphasis. Although certain words are commonly lengthened in this manner, there is no intrinsic limitation on which words can be emphasised, nor are there particular social conventions around the number of repetitions. A regular expression was used to identify three or more repeats of the same letter within a word. For tokens fitting this pattern, each repeated letter was replaced by single and double instances of that letter, and both of these were then tested against the dictionary. If a valid word was found by this method, the out-of-vocabulary token was considered to be an instance of affective lengthening. The score for this feature was calculated by dividing the number of lengthened words by the total number of words in the message.

Alphanumeric Words. A simple regular expression was used to identify tokens which contained both letters and numbers. The score for this feature was calculated by dividing the number of alphanumeric words by the total number of words in the message.

Emoticons, although not ‘words’ in the traditional sense, fall most naturally into the category of lexical features. Treating emoticons as a special class of word allows for a proportion to be calculated in the same way as for other word classes. A set of regular expressions were developed to identify common emoticons within a range of styles (see Appendix A.1 for details). Text was tokenized, and each token (word) was then checked against these patterns to assess whether the token represents an emoticon.

Deixis. Deictic expressions are those which depend on external reference in order to be properly interpreted. These may be temporal (‘now’, ‘yesterday’), geospatial (‘here’, ‘there’), or personal (‘he’, ‘you’). Texts featuring more deixis are considered to be less formal, because the use of deictic expressions assumes common knowledge that the audience will use to interpret the message. This feature overlaps heavily with Heylighen & Dewaele’s (2002) F-score, which was also calculated. However, F-score depends on reliable part of speech tagging, which does not exist for all languages and is not always effective for informal data. With a wordlist-based approach, a new language can be introduced by asking a speaker of that language to compile a list of examples. While this may not be perfect, it is nevertheless a more pragmatic approach for less commonly taught languages. For use in this work, a list of deictic expressions was compiled from examples in the literature; the full list can be found in Appendix A.2.

Hedges. A list of hedging expressions was compiled from examples in the literature and ESL sites. We took a very broad definition of hedging strategies, erring on the side of inclusion, to generate a wordlist against which tokens could be compared. This list can be found in Appendix A.3.

Politeness. In English this is a very limited feature, compared to some other languages, but we included words such as 'please', 'thanks', and 'sorry' which express explicit acts of politeness. Polite words used are listed in Appendix A.4.

Expletives. A list of English swearwords was harvested from online resources, and manually edited to give a condensed list of terms which are typically used as swear words. This was based on intuition rather than hard evidence, and erred on the side of excluding words with a common non-taboo use (e.g. 'bull'), on the basis that words which exist only as swearwords (e.g. where the literal as well as expletive sense is also considered taboo, such as 'fuck' and 'cunt') tend to represent the most extreme language. The list of expletives used can be found in Appendix A.5. Future work could investigate more rigorous methods of identifying whether a word is being used as an expletive in a given context; studies have shown that the wordlist-matching baseline can be slightly exceeded by a topic-modelling approach (Xiang et al. 2012, Sood et al. 2012).

Contractions. Contractions such as 'don't' and 'they've' are standard English, and as such, will pass the spellcheck test. Nevertheless, these are more informal than the uncontracted forms 'do not' and 'they have', and as such, their use is worth investigating. In English, contracted forms contain an apostrophe to indicate the missing letters, so a wordlist was generated by taking all entries in the dictionary file from (above) which contain an apostrophe, and then manually editing the list to remove any instances (primarily foreign loan-words) which were not in fact contractions. The final list can be found in Appendix A.6.

3.2.4. Tag Questions

Questions are an interesting discourse feature in their own right, but as discourse features tend towards the functional, we did not generally address them in this project. However, tag questions are a question type of particular interest, taking the distinctive form of 'You're coming, <u>aren't you?</u>' or 'He didn't, <u>did he?</u>' where a short question follows a declarative sentence form.[3]

Tag questions may be used in a literal sense, to request clarification, but they are frequently employed as hedges: in the hedging context they do not require an answer. It is in the case of hedging that they become a stylistic feature, and more likely to indicate features of relevance to our studies, such as deference.

[3] Tag questions also exist in a number of other languages, although they take several different forms, such as *¿no?* in Spanish and *nicht war?* in German.

Moore & Podesva (2009) examine the use of tag questions in a school environment, and find significant variation in the frequency and mode of tag use at both an individual and group level. This reinforces the general point that it is likely to be variation from the individual's usual style which says the most about their relationship to any particular interlocutor.

Ladd (1981) distinguishes *nuclear* and *postnuclear* tag questions as two distinct categories in speech, identified by their differing patterns of intonation and prosody, which typically have different semantic functions. For example, "Jane's coming, isn't she?" (postnuclear) as distinct from "Jane's coming isn't she" (nuclear), where the additional pause and questioning intonation indicates additional uncertainty in the postnuclear example. In particular, the postnuclear form is that which is more likely to convey hedging, and features a stronger break before the question part of the utterance. It is beyond the scope of the current project to computationally identify the pragmatic purpose of every utterance in text, so for now we consider all tag questions as equal, however in identifying tag questions we limit ourselves to the case where there is a comma before the question (indicating a break, and therefore more likely to be the nuclear variety). However, as we can never be certain that a tag question is being used for hedging, this was kept as a separate feature in its own right, rather than being incorporated into the count of hedges.

For the purposes of this project, we built a feature extraction engine to identify tag questions in English text, using a sequence of part-of-speech tags and punctuation marks. To identify the most common form of tag question in English, the required pattern is:

comma – verb – (optional) negation – personal pronoun – question mark

Example tag questions:

,	is	n't	it	?
,	are		you	?
,	did		he	?
,	was	n't	there	?

This was subsequently extended to include syntactic variants of the form 'did you not?' where the negation is expressed in full rather than as a contraction, requiring a change in word order.

The accuracy depends in part on the success of the underlying POS tagger. We used the OpenNLP tagger (OpenNLP 2011) as this gives a granular output when dealing with contractions, labelling the model verb and the negation separately, which is required for the above algorithm to succeed, whereas the Stanford tagger (Manning et al. 2014) labelled the entire contracted form as a single token. Using OpenNLP therefore enabled a more reliable result for tag questions.

To verify the output of the tagger, we ran an experiment using the Switchboard corpus, a data set which has been manually annotated for dialogue acts (Jurafsky et al. 1997). Tag questions are indicated with a ^g marker. The Switchboard corpus contains 384 files which contain the tag question marker ^g, and 575 individual lines are tagged as containing a tag question.

Of the tag questions returned by our annotator, all 137 were obviously good examples. More interesting are the instances which were missed by the annotator, but marked as tag questions by the human annotators responsible for creating the Switchboard annotations. 143 cases had a recognisable tag question form, but with patterns of punctuation that did not meet our specified constraints. In 83 cases, there was no comma preceding the tag question, and in 76 cases there was no question mark at the end (16 examples fell into both of these categories). As the Switchboard corpus is transcribed speech, the punctuation is down to the transcriber, and has nothing to do with the actual speaker. 227 examples were reduced tags, consisting of a single word (212, or 93%, of which are accounted for by the forms 'right?' or 'huh?'). At present we do not consider these forms in our analysis. Additionally, 44 examples were junk and contained no discernible question, and 25 consisted of a different question which did not meet our pre-defined syntactic patterns. Common examples of the latter case include 'you know?' 'you mean?' and 'do you think?'.

Parse failures are also responsible for some cases of a tag question being missed. If a sentence as a whole is not successfully tagged with parts of speech, the tag question annotator has nothing with which to work. However, attempting to correct for parser errors is out of scope for this project.

In the vast majority of cases, when tags were not identified from the Switchboard corpus data, this was due to the punctuation constraints which we have deliberately specified. It is worth a brief investigation into the effects of relaxing these rules. For example, foregoing the question mark requirement (while still requiring some sentence-final punctuation mark) retrieves 59 additional, genuine tag questions, and only two false positives. (198 total returned.) Removing the comma constraint results in 45 genuine tag questions being identified, and one false positive. In either case, this relaxation results in an increase to recall at minimum cost to precision. On the other hand, simultaneous removal of both constraints results in 246 false positives and a drop of precision to 51%, which is clearly not ideal for most purposes, even though the recall correspondingly increases to 91.4%.

	Recall (%)	Precision (%)	F1-score
Strict	48.9	100.0	65.68
relax question	70.0	99.0	82.01
relax comma	65.0	99.4	78.60
relax both	91.4	51.0	65.47
relax either	85.0	98.3	91.17

Table 2: Results of tag question detection on the Switchboard corpus.

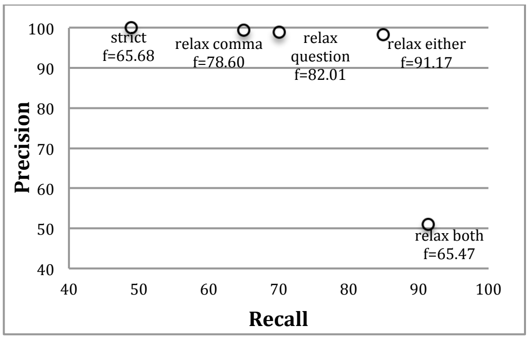

Figure 9: Detecting tag questions, precision versus recall.

As shown in Table 2 and Figure 9, if the goal is solely to maximise F-score of the tag questions annotated in the Switchboard data, the optimal balance between precision and recall is obtained by allowing a relaxation of either punctuation constraint, but not both.

However, we made a conscious decision to include only those tag questions which are preceded with a comma, in order to maximise the chance of identifying nuclear tag questions used for hedging, rather than for questioning. We can maintain this distinction and still gain a significant performance increase by allowing instances where the tag question is terminated by punctuation other than a question mark.

3.3. Feature Standardisation

In our examination of the theory, we have noted that different people use language in different ways, leading to personal idiolects. Our second hypothesis is that, in many cases, the difference between idiolectal preferences would be expected to overshadow the subtle variations generated as a particular individual reacts to situations of varying status and hierarchy:

> **H2:** Differences from the individual's normal behaviour will be more informative than absolute feature scores, for predicting relationships.

For example, consider a message containing a single mild swearword. This has a rather different meaning if the sender is someone who almost never swears, compared to if she is someone for whom half a dozen curse-words is the norm. For the first sender, the swearing is marked: it is likely to indicate a high level of informality, and possibly an unusual level of stress or excitement. For the second, even though swearing is generally an indicator of informality, this particular message may actually be at the more formal end of their communicative spectrum.

This example neatly illustrates why it is digression from the idiolect, rather than absolute value, which is likely to be most informative.

To measure the importance of individual variation, we take the approach of standardising each feature by sender, such that each sender's feature distribution will be adjusted have a mean of 0 and a standard deviation of 1. This is achieved as per Equation 1.

$$x_i^A \longrightarrow \frac{x_i^A - \mu_i^A}{\sigma_i^A}, \qquad i \in \{1, \ldots, n\}$$

(Equation 1)

where x_i^A represents a particular instance of person A using feature i, μ_i^A is the mean value of feature i across all communications originated by person A, σ_i^A is the corresponding standard deviation across all of A's communications, and n is the number of features for which standardisation is undertaken.

By adjusting each feature to a common distribution in this manner, we can more easily distinguish variations relative to the personal idiolect of the individual.

Having generated per-author standardised feature sets, we can then compare the results to results from the raw scores. If H2 is correct regarding the interaction of idiolect and audience-adjustment, we should see a consistent improvement from standardising feature scores in this manner.

3.4. Machine Learning Approaches

Machine learning techniques can be divided into supervised and unsupervised, depending on whether use is made of examples previously labelled with some catagorical ‘truth’ to which a model will be fitted (the supervised context), or whether the goal is to uncover structure within the data without any preconceptions (unsupervised approaches).

In this instance, since we have a number of labelled examples, supervised approaches will be the appropriate choice to enable us to predict classes and assess the accuracy of our predictions, as well as providing a model that could be used in an eventual application.

As we will need to explore a large number of variations in our feature sets, it is important to identify a suitable methodology that will allow comparison of results under different experimental conditions. Our interest is in the predictive power of our features; the central goal of our research is not to compare the efficacy of different classifiers or to fine-tune parameters to optimise classifier performance, so we will not devote much time to these matters, although this kind of optimisation would obviously be required for a practical application.

Additionally, we can concieve of our classification task in two distinct ways: as a three-way classification (upwards, downwards, or level communication) or a series of binary classifiers (whether a message represents an example of hierarchical or level communication; and if hierarchical, whether it is directed upwards or downwards). This gives three possible avenues of exploration.

We conducted an initial categorisation experiment to compare the performance of different classifiers using 10-fold cross-validation on a subset of the Enron corpus, using the WEKA framework (Hall et al. 2009). Note that this was not the same Enron dataset which we used for our eventual experiments, as we conducted this initial test at a very early stage, before we had cleaned and extracted our final dataset. We tested both three-way classification and the alternative binary classifiers. These initial results are plotted in Figure 10. This indicated that a Random Forest classifier would provide the best balance of accuracy and speed, providing the best performance in two of the three cases. It was therefore determined to use random forests for the bulk of the classification experiments. It may be appropriate to revisit this decision in future, in particular if an end-to-end application were to be built, at which point it would become more relevant to optimise performance using the finalised feature set.

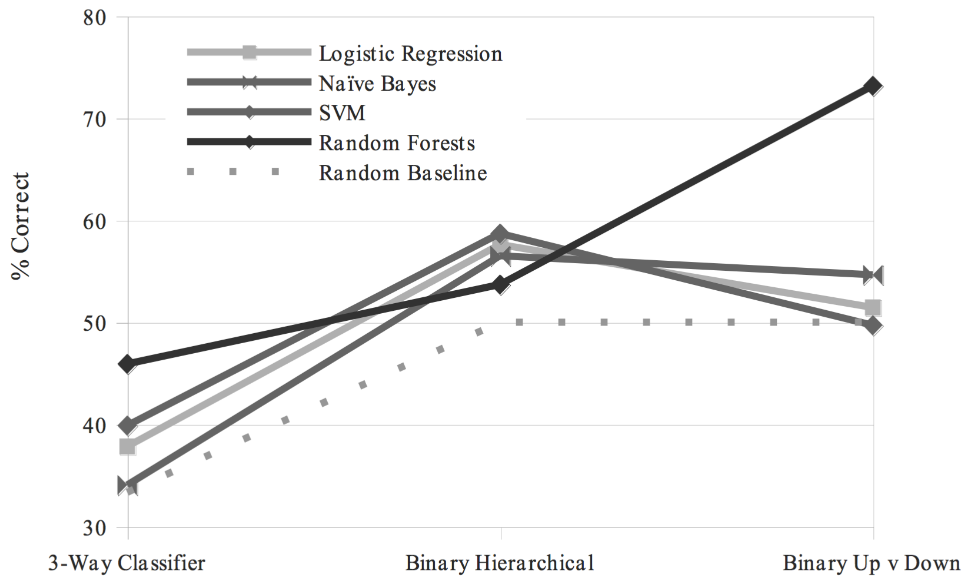

Figure 10: Comparison of performance of different classifiers

We also tested the stability of our results using different data partitions. Once we had completed cleaning the Enron data, we generated fifty additional pairwise-partitioned data sets from our corpus, and conducted 10-fold cross validation with each resulting set. We tested this using the 3-class task, using raw (unstandardised) scores. Results are plotted in Figure 11. Accuracy ranged from 36.62% up to 38.98%, with a mean of 37.74 and standard deviation of 0.58, while F-measure ranged from 0.340 to 0.368, with a mean of 0.353 and standard deviation of 0.006. From this we can see that our primary dataset, with accuracy 37.66% and F-measure of 0.356, is fairly representative, within one standard deviation of the mean in both dimensions.

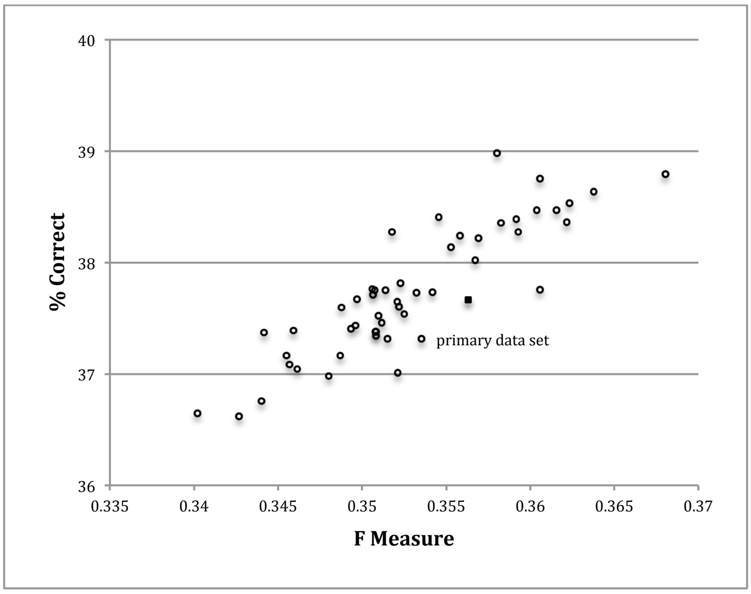

Figure 11: Stability of accuracy and F-measure results on the Enron corpus, using different data partitions.

3.5. Aggregation of Feature Scores

Our intuition is that a single prediction for each relationship will be more effective than individual predictions at the message level, due to the additional information that is available when predictions from multiple messages are used in combination, to provide a single prediction for each pair of interlocutors. This gives rise to our third hypothesis:

> **H3:** Prediction of hierarchical relationships will be more effective at the pairwise level than at the level of individual messages.

In order to test H3, therefore, we require a method of aggregating individual message predictions into a single classification for each relationship. The most basic method of combining scores is to use a 'voting' method. For example, given a set of twenty messages between A and B, we might have the following output from our individual message classification:

A to B	upwards: 7	downwards: 2	level: 1
B to A	upwards: 4	downwards: 6	level: 0

Based on these numbers, we would have one vote for an equal relationship, and 19 for a hierarchy. Looking further into the hierarchical evidence, we find 7 + 6 = 13 votes for A being subordinate to B, and 2 + 4 = 6 votes for B being subordinate to A.

In this case, if A is indeed B's subordinate, we have the potential to turn 65% message-level accuracy into a single correct prediction at the relationship level. Of course, the converse of this is that when we get it wrong, we will be degrading our overall performance.

For each relationship, taking the case with the highest number of supporting evidence points is the simplest method of aggregation. We will call this 'simple plurality voting', by analogy with electoral systems such as first-past-the-post, in which an absolute majority is not required.

It is also possible to reqiure a majority agreement, which we will call 'simple majority voting'. This will often result in fewer instances being classified, but with hopefully higher confidence.

In general, simple plurality voting and majority voting can both be considered as instances of a threshold-based system. For plurality, the applicable threshold is 0.333 (under three-way classification, the minimum number at which one class can outnumber both others), whereas for a majority, the threshold is 0.5. It should be possible to select a threshold value to adjust the balance between coverage and accuracy.

Another possibility is that the data may appear to rule out one class, without deciding clearly between the remaining two possibilities. For example, it could be clear that a relationship is hierarchical, but not in which direction; likewise it may be clear that A is not positioned above B in the hierarchy, but not whether the two are in fact peers or whether B is A's superior. We should investigate whether these less-precise assertions can be made with greater confidence.

For the speech and online chat data, we have balanced datasets due to the dialogic nature of the interactions: for all pairs, we have approximately the same number of utterances from A to B and from B to A. This is not true for the Enron data, which is often heavily unbalanced, whether due to actual differences in productivity or differences in corpus coverage. It is worth examining the effect of such imbalances on voting methods.

We will assess the impact of all these techniques on our data, as appropriate.

3.6. Identifying Power Flow in Graphs

Our fourth hypothesis, much like H3, concerns the idea that more information should enable more accuracte predictions. In this case, we turn our attention to the graph of predictions across the whole network, and the idea that we can be more confident in predictions which are logically consistent with other predictions in the same network.

> **H4:** Relationship classification in a hierarchical situation will be aided by consideration of whole-network characteristics.

To test H4, we consider the possibility of improving our predictions by looking for inconsistencies in power predictions. We will approach this by examining the overall shape of the graph resulting from the predictions.

There is a large body of existing work on training a prediction model based on the metadata of a communication graph (Rowe et al. 2007; Gallagher 2010; Agarwal et al. 2012), and we have no intention of replicating this work here. Rather than examining the underlying communication graph, we are rather interested in creating a graphical representation of our relationship predictions, and using this to determine whether our predictions are self-consistent.

In particular, a simple conceptualisation of power is as a transitive relation: if A is senior to B, and B is senior to C, then A is senior to C. Although this does not incorporate any of the complexities of human interactions, it should be broadly true in the context of a hierarchical organisation. Therefore, upon constructing a graph of our relationship predictions, if we encounter points in the graph where this transitivity is not upheld, this is likely to indicate either a mistake in our predictions or a relationship that is unusual in nature, perhaps due to the social sphere interfering with the professional.

Represented graphically, with directed edges to indicate seniority, we should observe the transitive relationships as illustrated in Figure 12, where if A is senior to B, and B is senior to C, then A is also senior to C. If our relationship-level predictions suggested that C was senior to A, then in graphical terms, we would have a loop (Figure 13), which is a strong hint that one (or more) of the directed links could be incorrect.

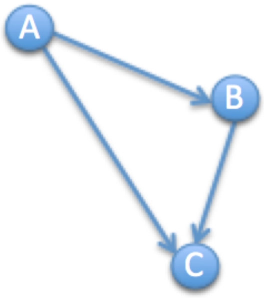

Figure 12: The expected shape of a power graph

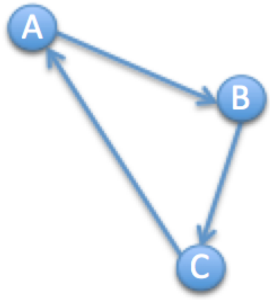

Figure 13: A loop, such as we would hope not to find in a power graph

In a small network, anomalies of this kind can be spotted by eye, but as the number of links in the network grows, this becomes increasingly unfeasible. One mathematical method for identifying such areas in the graph is to calculate flows and tensions (Kochol 2004). Flows are defined as directed cycles within the graph, and the tension of a graph is defined as the acyclic component which contains no such cycles, relative to any arbitrary reference orientation. For the tension component, the sum of the labels on any cycle within the graph will be zero. Any graph can be decomposed into a single tension component plus some number of flows; the edge weights retained by the tension component can be interpreted as giving the strength of the (directed) relationship, or in our terms, the level of relative seniority. In the context of an organisational hierarchy, we would expect to have greater confidence in relationships in the tension component, as these are mutually supportive.

The concept of flows and tensions naturally applies to a weighted, directed graph. In the case of an unweighted directed graph, we can use weights of 1 to indicate the direction of the link. Note that in a directed graph, a weight of -1 from A to B is equivalent to a weight of 1 from B to A. There are two plausible ways in which to construct power graphs from the predictions we have made in earlier chapters. The first and simplest option is to generate a single edge between each pair based on aggregate results, using input edge weights of 1 (upwards), -1 (downwards) or 0 (level). The second is to create weighted edges between each pair, with the weight based on individual message predictions. We will assess the output of both of these approaches.

Figure 14 shows a simple example of a tension-flow decomposition in a graph without any cyclic paths. Even though there are no awkward cycles in this graph to be mitigated, the resulting tension graph gives different weightings to the edges, due to the fact that there is a direct path from node A to node E, but that there is also a path A-D-E. This seems to back up the intuitive understanding that if A is senior to D, and D is senior to E, then not only is A senior to E, but A is correspondingly *more senior* to E than to D, as represented by the higher number on the link between A and E. Note that a flow component is also extracted; this is not to be taken as evidence that these links are not valid, as these nodes are still linked into the tension component.

Figure 15 demonstrates the effect of reversing a single link, to create a graph that is no longer naturally interpretable as a hierarchical system. The tension component no longer takes in all the nodes, once edges of weight zero are removed: note that nodes D and E are no longer part of the tension component, and the links A-D, D-E, and E-A have all been removed. In this case it is the omission of certain nodes/links from the tension component that is the most significant, although we can look at the related flow component to better understand the source of the complexity. In a non-trivial graph, complete removal of nodes from the tension component in this manner may be comparatively unlikely, as it results only for members of a perfectly-balanced loop. However, when links are omitted from the tension component, this is the equivalent of saying we cannot make any prediction for the power relation that would have been represented by the link in question.

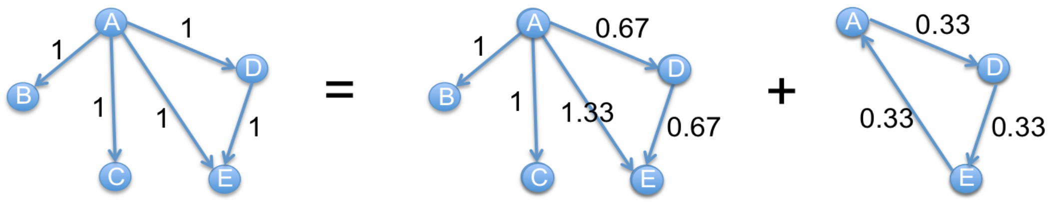

Figure 14: Decomposition of a simple graph into tension (left) and flow (right) components.

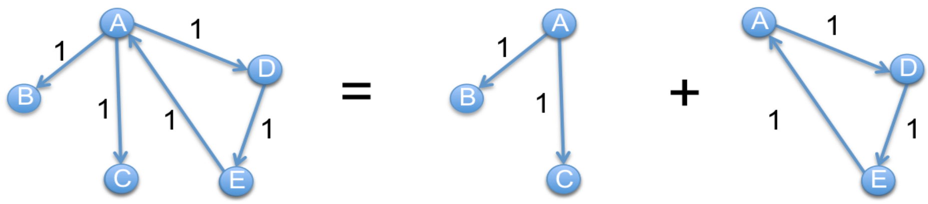

Figure 15: Graph decomposition into tension (left) and flow (right) components, with a loop.

We use Waters (2013) software implementation in Java to calculate flows and tensions for the graphs. We will use this to test H4 by calculating flows and tensions over the graphs which we will construct from our relationship-level predictions, in order to test our predictions for self-consistency. If H4 is supported, we would expect to find that self-consistent predictions (reflected by their scores in the tension component) give a more confident prediction than those predictions which contradict one another.

3.7. Linguistic Accommodation as Movement in Vector Space[4]

Our fifth hypothesis concerns variation of a speaker's linguistic style relative to their interlocutor, and derives from communication accommodation theory (Giles 1973) and previous work on linguistic accommodation (Bunz & Campbell 2004; Danescu-Niculescu-Mizil et al. 2012). Convergence in speech or non-verbal behaviours has resulted in a perception of increased similarity between interactants (Giles 2008; Welkowitz & Feldstein 1969), while convergence in non-verbal behaviours

---

[4] This section is based on work undertaken in collaboration with Kate Muir, Simon Jones, Adam Joinson (University of the West of England), and Nigel Dewdney (University of Sheffield) and was first published in Jones et al. (2014).

(mimicking body language, facial expressions, or gaze) has been associated with feelings of rapport between interactants (Lakin & Chartrand 2003). However, little prior work has focused on the factors which effect the likelihood of linguistic style accommodation taking place within any given interaction. We will examine the impact of social power in particular, in order to test our fifth hypothesis:

> **H5:** Linguistic accommodation behaviour is partially motivated by relative social power, therefore greater accommodation of linguistic style will correlate with lower social power.

To test H5, we will consider the relationship between accommodation of linguistic style and social power structures, using a vector space model of linguistic accommodation.

Language Style Matching (LSM) is currently the most prevalent method of measuring accommodation of linguistic style. To calculate LSM, one simply takes frequency of function word use for a pair of interlocutors, and measures the absolute distance between the two. Equation 2 gives the LSM calculation for prepositions, as used by Ireland et al. (2011), where $preps_1$ and $preps_2$ represent the count of prepositions used by the first and second speaker respectively.

$$LSM_{\text{preps}} = 1 - [(|preps_1 - preps_2|)/(preps_1 + preps_2 + 0.0001)]$$

(Equation 2)

LSM provides a single score per pair of individuals, which has been shown to correlate with social outcomes, but this metric does not capture the extent to which one or both of the individuals concerned may have adjusted their personal style in order to reach this position. As such, it cannot really be described as a measure of linguistic accommodation. Figure 16 gives a pictorial illustration of how symmetric accommodation (top line) between two parties could result in the same absolute LSM value as asymmetric accommodation (middle), or even divergence (bottom).

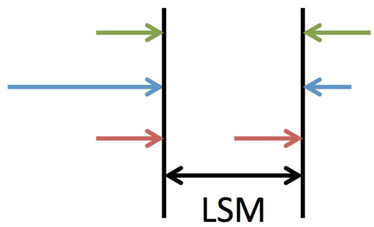

Figure 16: An illustration of how different accommodative behaviours could result in the same absolute score using Linguistic Style Matching (LSM).

Turn-by-turn measures of linguistic similarity (Niederhoffer & Pennebaker 2002; Danescu-Niculescu-Mizil et al. 2012) have been used to add a temporal component to the analysis. The theory behind this approach is that if LSM, or any other linguistic similarity metric, is calculated sequentially, the delta in these scores from turn to turn will represent an individual's movement towards or away from their interlocutor. This allows for a more detailed examination of conversational dynamics, in particular it is possible to see how linguistic co-ordination varies throughout the course of a dialogue. However, this method does not account for the fact that individuals may pre-emptively accommodate towards their perception of another's *likely* style (based on social group stereotypes) even during their first turn in the conversation, a phenomenon which has been previously observed in studies of behavioural mimicry (Willemyns et al. 1997).

Additionally, there is no existing method that accounts for the inherent tendency of an individual to accommodate (or not) in their use of language, which may be a salient feature when interpreting raw scores. If an individual has a tendency to adjust her language during every exchange to perfectly match her interlocutor, then her doing so is less significant, and we would be unable to draw conclusions about social power based on her behaviour. On the other hand, for someone who is not generally prone to accommodation, a numerically much smaller shift might be highly informative.

To address these limitations, we therefore propose that it is necessary to construct a metric which includes not only a model of an individual's baseline linguistic style, but which also captures the individual's tendency to accommodate.

We achieve this by means of the Zelig Quotient, a measure for normalizing linguistic variation using a vector space model (Jones et al. 2014). A vector space approach is a natural fit for this problem, since we have a set of point measurements (feature scores) and we are interested in how these change over time (movement in feature space). Since we wish to distinguish convergence and divergence, as well as over-accommodation, the direction of movement in vector space is important to our model.

The Zelig Quotient is named for Leonard Zelig, the central character of the Woody Allen film Zelig, who is described as "the human chameleon" due to his propensity for taking on the characteristics of other people. This is the logical extreme of accommodating to one's audience. An author who always adopts the language style of the intended reader is totally Zelig-like, whereas an author who does not adapt at all has zero likeness. Divergence represents opposite behaviour to Zelig (moving away from the audience), which is represented by negative Zelig scores. Over-accommodation occurs when the author adopts elements of linguistic style of their intended reader, but emphasises to the point of overuse; this is represented by scores greater than 1.

The Zelig Quotient is based on vector modelling, thus showing the extent to which an individual changes their linguistic style from their 'typical' or baseline style, to move either towards or away from each of their conversational partners. The average Zelig score across all conversational partners can then be used to demonstrate the individual's general tendency to accommodate their language use to that of others.

We wish our results to be comparable with earlier studies using linguistic style matching (LSM), and therefore we do not measure accommodation across the full range of stylistic features described in section 3.2. Rather, we use nine function word categories from LIWC (Pennebaker et al., 2007), as enumerated in Table 3, to enable direct comparison with earlier work (such as Ireland et al. 2011).

LIWC Category	**Examples**
Personal pronouns	I, his, their
Impersonal pronouns	it, that, anything
Conjunctions	and, but, because
Auxiliary verbs	shall, be, was
Articles	a, an, the
Prepositions	in, under, about
High frequency adverbs	very, rather, just
Negations	no, not, never
Quantifiers	much, few, lots

Table 3: LIWC features for measuring linguistic accommodation

We assume an author has a latent linguistic style resulting in a baseline value for each of our stylistic features. However, we expect variation in the observed values due to sampling, an author's natural variation, constraints of message content and format, etc. as well as any movement due to accommodation. We can estimate a baseline value for a specific feature, μ_f, by averaging over all the messages we have for an author *a*. Previous research has used a similar technique for establishing the baseline level of a lexical item in a dialog in order to study accommodation (Church 2000; Stenchikova and Stent 2007).

$$\mu_f(a) = \sum_{m=1}^{n_a} f_m(a)/n_a \qquad \text{(Equation 3)}$$

where *m* is a message, n_a is the number of messages sent by author *a* and $f_m(a)$ is the feature value in *m*.

We can further estimate the proportion of variance due to 'noise' and that due to accommodation by also calculating the average feature values on an author-reader *(a, r)* pairwise basis.

$$f(a,r) = \sum_{m=1}^{n_{ar}} f_m(a,r)/n_{ar} \qquad \text{(Equation 4)}$$

where n_{ar} is the number of messages written by author *a* to reader *r*.

Measuring the variance within a pair and then averaging over all pairs that author is party to gives an estimate of the natural variation in feature value for an author.

$$\sigma_f^2(a) = \frac{1}{R_a} \sum_r \sum_{m=1}^{n_{ar}} (f_m(a) - \mu_f(a))^2/n_{ar} \qquad \text{(Equation 5)}$$

The movement in a feature due to accommodation is simply taken to be the difference between the value seen within a communicative pair, and the author's baseline value.

Having calculated scores for several features, some of which may change more readily than others, we can consider authors as having corresponding points in an F-dimensional feature space described by the vector of feature values. For our purposes, this is a 9-dimensional space, but this technique could easily be generalised to measure accommodation in more (or entirely distinct) dimensions. The generalised phenomenon of accommodation can then be measured in terms of movement in this feature space, rather than movement in individual features. Note that to avoid bias towards the effect of particular features when considering overall movement, feature scales must be comparable.

We represent movement and distances between the author's baseline position, accommodated position, and the reader's position as vectors in the feature space:

$$\mu = \{\mu_1(a), \mu_2(a), ..., \mu_F(a)\}$$
$$a = \{f_1(ar), f_2(a, r), ..., f_F(a, r)\}$$
$$r = \{f_1(r), f_2(r), ..., f_F(r)\}$$

We use the law of cosines to yield the cosine of the angle between the vector connecting the reader to the author's baseline position, and that connecting the reader to the author's accommodated position. The angle will be greater than 90◦ if the author has over-accommodated, and will therefore have a negative cosine value. However, the dot product of these two vectors gives the cosine of the inner angle. Therefore, normalising by this gives a value of +/- 1 according to whether accommodation movement is less or more than the amount required to meet the reader.

Multiplying the accommodated distance by this +/-1 factor gives us a definition of an accommodation metric that expresses the accommodation as the change in directed distance from the reader, proportional to the amount required from the author's unaccommodated position. This may be greater than 1 (over-accommodation) or less than zero (divergence). In vector notation we define accommodation as:

$$Acc(a, r) = 1 - \left(\frac{|\vec{ar}|}{|\vec{\mu r}|}\right)\left(\frac{|\vec{\mu r}|^2 + |\vec{ar}|^2 - |\vec{\mu a}|^2}{2(\vec{\mu r} \cdot \vec{ar})}\right)$$

(Equation 6)

The dot product of $\overrightarrow{\mu r}$ and $\overrightarrow{ar}$ is zero if the two vectors are orthogonal. However this is matched by a zero value in the numerator and we take the final parentheses value in Equation 6 to be 1 in this case. In the other pathological case where $|\overrightarrow{\mu r}|$ is zero, the implication is that author and reader have the same preferred position, i.e. there is nothing meaningful to say about accommodation between the two.

Having estimated author to reader accommodation, we are now in a position to estimate how readily the author adapts to others, by averaging over the set of readers. This gives us our Zelig factor, Z.

$$Zelig(a) = \frac{1}{R_a}\sum_{r=1}^{R_a} 1 - \left(\frac{|\vec{ar}|}{|\vec{\mu r}|}\right)\left(\frac{|\vec{\mu r}|^2 + |\vec{ar}|^2 - |\vec{\mu a}|^2}{2(\vec{\mu r}\cdot\vec{ar})}\right)$$
(Equation 7)

A positive (+) Zelig Quotient signifies the author readily accommodates, with a Zelig Quotient of 1 indicating the author always adapts their linguistic style to that of their audience. A negative (-) Zelig Quotient suggests divergence in the author's linguistic style (moving away from the audience).

Significance of values can be estimated from the variance. Here we take movement beyond one standard deviation of the authors total message distribution. The significance of an author's Zelig Quotient then follows from averaging the variance seen over the pairs the author is party to.

$$Zelig_{min}(a) = \sqrt{\frac{1}{R_a}\sum_{r=1}^{R_a}\frac{\sum_{f=1}^{F}(\sigma_f(a) - \mu_f(a))^2}{\sum_{f=1}^{F}(f(r) - \mu_f(a))^2}}$$
(Equation 8)

This model assumes that there are latent baseline distributions for feature values but does not suggest a generative function. Further work will determine appropriate distribution models for features, to be parameterised from the estimation methods presented here.

Having developed a vector approach to linguistic accommodation, we are now in a position to test H5 by applying this approach to conversations with known social power dynamics.

3.8. Discussion

We have now laid out the underlying methodological approaches for testing our five hypotheses. The results of these experiments are discussed in the following chapters.

Chapter 4 considers the task of predicting social power from a single message, whether this is an email, a spoken utterance, or a turn in an online chat room (H1) and examines the effect of standardising feature scores on a per-author basis (H2). These experiments are repeated across three datasets to allow comparison across genres: we use a subset of the Enron email corpus, the Muir-Joinson speech corpus, and the Muir-Joinson CMC corpus.

Chapter 5 considers methods of aggregating information from multiple messages in an effort to increase confidence in our results, both through aggregation of predictions (H3) and through graph analytics (H4). The message predictions generated in Chapter 4 give the input to these methods. Chapter 6 takes a different approach to the intersection of linguistic style and social power, examining the relationship between an individual's power role and their tendency to accommodate their use of language (H5).

In combination, these experiments will provide an overview of the impact of social power on stylistic choice, and should develop a solid foundation for future work.

4. Message-Level Categorisation

Our primary hypothesis is that we can infer a social relationship from aspects of linguistic style:

> **H1:** Stylistic choice is a key method of expressing relationship-building, therefore relationship information can be inferred from linguistic style.

To test this, we use the datasets described in section 3.1 and the features described in section 3.2, and attempt to predict social power based on style features. The task of identifying relative power from text can be conceptualised as a categorisation problem. Relationships within an organisational hierarchy can be, simplistically, characterised as 'hierarchical', where a power differential exists, or 'level', where two colleagues are approximately equal in status. A given message may have been sent up the hierarchy to a more senior member of staff, down the hierarchy to a junior, or between two peers in an equal relationship. At the message level, then, this problem can be described as a three-class categorisation task.

It is also possible to consider three-way categorisation as two sequential binary classifiers: firstly to determine whether a message exists in a hierarchical or peer-to-peer context, and secondly for hierarchical messages, to determine the direction of the hierarchy: relative performance on these two tasks may give an indication as to where overall performance could be improved.

At the same time, we will consider our second hypothesis, which relates to individual variance of linguistic style:

> **H2:** Differences from the individual's normal behaviour will be more informative than absolute feature scores, for predicting relationships.

In order to test H2, we will run our categorisation experiments again using feature scores that have been standardised by author; if H2 is upheld, we expect to see a performance improvement under this condition.

We used the WEKA implementation of the Random Forest classifier (Hall et al. 2009) for all of the experiments in this chapter.

4.1. Enron Emails

Our first case study is of the Enron email corpus. During the data preparation phase, we prepared three separate sets of ten-fold data splits: one completely randomised at the message level, one split by sender-recipient pair, and one split by senders.

We perform ten-fold cross-validation for each of our three cases, using both raw and standardised feature scores in each case. The results are plotted in Figure 17.

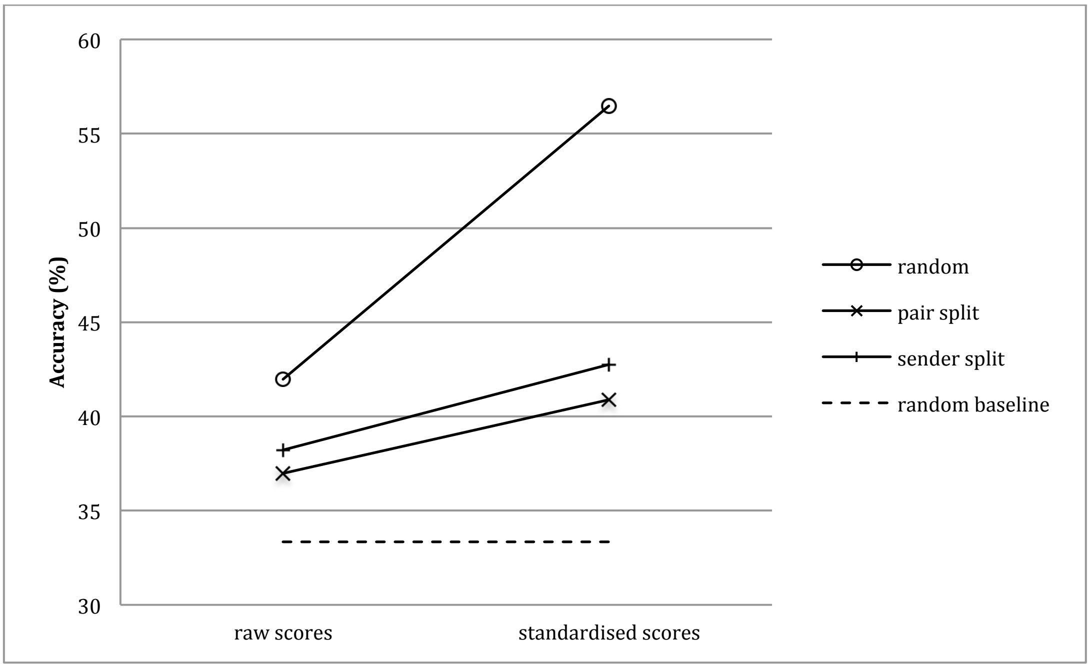

Figure 17: Enron corpus 3-way classification results, effects of standardisation under different data partitioning scenarios.

We expected to observe a general decrease in performance as the partitioning becomes more stringent, however, we note the surprising result that performance is actually slightly higher on the sender-partitioned data than the pairwise partitioning, for both raw and standardised data. It is not clear why this might be the case, and indeed this was not observed in preliminary tests (using an earlier selection of features), so this may be an artefact of the particular data partitioning.

It is interesting to note that the standardisation of features brings the pairwise- and sender-partitioned scores up to approximately the same level as the raw scores in the random case. This suggests that the standardisation may be compensating somewhat for the decrease in performance that follows from more rigorous partitioning. However, random partitioning also exhibits a significant increase when features are standardised, providing additional support for the validity of the standardisation approach.

We will focus primarily on the pairwise partitioning, as this represents the most likely case in a real-life scenario: one may have partial information about an organisation's hierarchy, and wish to build a model from this to predict the relative positions of unknown individuals within the same organisation. And we use the standardised feature scores, as these produce consistently better results. Our overall accuracy under this condition is 41.34%, which although not fantastic performance, is nevertheless above the random baseline of 33.3%, and narrowly beats the most common class baseline of 41.17%.

	upwards	downwards	level
upwards	**354**	93	1469
downwards	145	**294**	1038
level	392	270	**1712**

Table 4: Confusion matrix for pairwise partitioned, standardised data, 3-way classification

Examination of the confusion matrix for the pairwise-partitioned, standardised data clearly shows that the majority of mis-classification is down to messages being incorrectly placed into the 'level' category (Table 4). Although 'level' communications do make up the largest class of the training data, representing 41.17% of messages, the classifier applies the label to 73.16% of results. This bias was not corrected by experimenting with an artificially balanced dataset: the tendency to incorrectly classify messages as 'level' persists even when it is not the majority class in the training data, and may be because 'level' messages tend to be more neutral in style, in a manner that may also be employed as part of a range of communication strategies used by participants within a hierarchical relationship.

The three-way classification problem can alternatively be considered as the concatenation of two sequential binary classifiers: one to distinguish hierarchical from peer-level messages and the second, for messages with a power differential, to distinguish between those going upwards or downwards in the hierarchy. This view of the problem space allows for a more nuanced understanding of the contributing features and their interactions.

To train a classifier to distinguish hierarchical from level communications, we begin by relabelling the data, assigning all 'upward' and 'downward' messages into the 'hierarchical' category. This gives us a dataset of 3393 hierarchical messages, and 2374 level. Note that we use the same partitions for cross-validation as in the three-class task: we have simply relabelled the data. Our random baseline for this task is 50% and the most common class baseline is 58.8%.

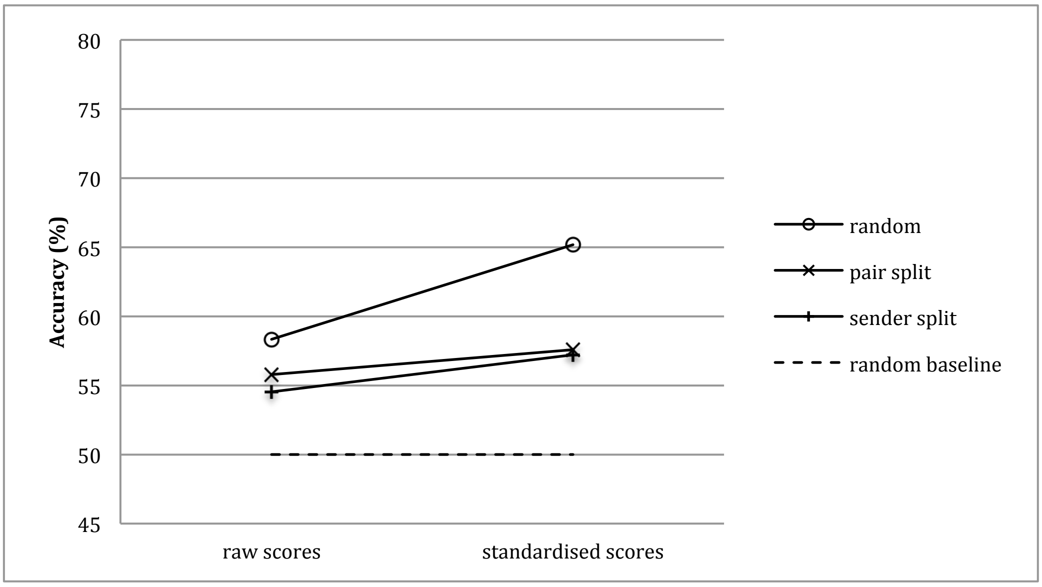

Figure 18: Enron corpus, hierarchical versus level classification, effects of standardisation under different data partitioning.

We repeat the comparison of results under the three different partitioning conditions, with and without standardisation. Again, we observe an improvement for standardised results in all cases, although the increase is most significant under random partitioning, but only a couple of percentage points for the pairwise- and sender-partitioned cases.

As before, we select the standardised and pairwise-partitioned data to examine in more detail. The resulting confusion matrix shows that we are still seeing very unbalanced results.

	hierarchical	level
hierarchical	**2876 (84.76%)**	517
level	1935	**439 (18.49%)**

Table 5: Confusion matrix for hierarchical versus level classification

In the hierarchical case, individual feature ablation results in a small decrease in performance for every feature which is removed. The top ten features from ablation can be divided into punctuation choice (semicolons; question marks; exclamation marks; hyphens; uppercase letters), length features (characters per word; number of sentences), and some parts of speech (determiners; adjectives; prepositions). Choice of punctuation and length features, in particular, are very much open to variation of individual choice, as evidenced by their value in the authorship attribution task (Grieve 2007; Stamatatos et al. 2001).

For the second stage, we use only those messages which are genuinely hierarchical in nature, and endeavour to distinguish those which are sent upwards versus downwards in the hierarchy. This gives us a smaller dataset than we have dealt with previously, consisting of 1916 upward communications and 1477 downward. The most common class baseline for this task is 56.5%, while the random baseline remains 50%.

Again, we use the same divisions of the data for ten-fold cross validation, in the three cases of random, pairwise, and sender partitioning.

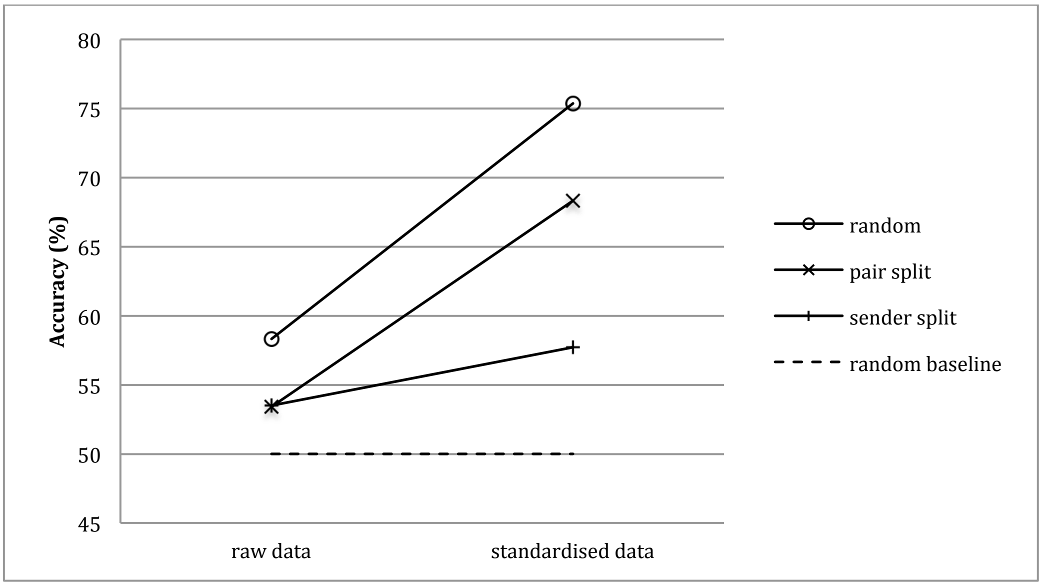

Figure 19: Enron corpus, upwards versus downwards communications, comparison of partitioning and standardisation results.

The first thing we note is that the difference between raw and standardised scores is larger in this case than in the hierarchical condition.

This is also the only segment of the task with a direct analogue in previous literature: Bramsen et al. (2011) report 78.1% accuracy on an upspeak/downspeak classifier using a mixture of word n-grams and part of speech n-grams, using a subset of the Enron corpus partitioned by sender. Note that although we match their case as far as possible, we are unable to guarantee the same subset of the Enron emails, nor the same ground truth regarding roles, as Bramsen et al. do not publish exact details of their implementation. While our results are much lower (57.71% for the sender partitioned case), we are using only 40 stylistic features, contrasted with the almost unconstrained size of an n-gram model. This difference in performance may be characterised as the cost of excluding topical information, 'what is communicated' as opposed to 'how it is conveyed'.

Examining our pairwise partitioned data in more detail, the confusion matrix shows a tendency to classify messages as 'upwards', resulting in 90% accuracy for actual upward messages, but only 40% for those messages going down the hierarchy.

	upwards	downwards
upwards	**1725 (90.03%)**	191
downwards	886	**591 (40.01%)**

Table 6: Confusion matrix for upwards versus downwards classification

Feature ablation for the up-down task shows a drop in performance for only 14 of the available features. Again, punctuation is a productive category (quotation marks; exclamation marks; semicolons; colons; ampersands; full stops), as are length features (number of sentences; words per sentence; number of characters). Lexical features directly associated with politeness and (in)formality also feature more strongly: contractions, polite phrases, alphanumeric words, and deictic expressions are all present on the list.

Equally interesting are the omissions: no part of speech tag categories make the list, nor do hedges or expletives, which are two other categories that theory suggests might be particularly indicative of hierarchy. In the case of expletives, this may be attributed to the overall scarcity of feature observations.

Having examined the ablation results for both the hierarchical and up/down classifiers separately, we are now in a position to compare the results of these analyses: in particular, we are interested in which features provide the largest contribution in each of the two cases. We are able to compare the "Top 10" lists generated under each stage (Table 7).

Only three features (marked by italics) make the list in both cases: percentage of semicolons, number of sentences, and percentage of exclamation marks. Features are shown in Table 7, along with the resulting drop in performance, in percentage points. Full ablation results are included in Appendix B.

	Stage 1: Hierarchical v Level	Stage 2: Upward vs Downward
1st	*PercentSemicolon* *-2.16149784*	PercentQuote -0.986540324
2nd	PercentQuestionmark -1.767487066	*NumSentences* *-0.723805324*
3rd	CharactersPerWord -1.49902995	*PercentExclamation* *-0.635587306*
4th	*NumSentences* *-1.466898264*	*PercentSemicolon* *-0.608054922*
5th	PercentDeterminer -1.439738716	PercentColon -0.544628664
6th	PercentAdjective -1.340865885	WordsPerSentence -0.516221101
7th	PercentHyphen -1.18488448	PercentAlphanumeric -0.469954731
8th	*PercentExclamation* *-1.158088441*	PercentAmpersand -0.392623103
9th	PercentUppercase -1.154132564	PercentLetter -0.370076188
10th	PercentPreposition -1.151816837	PercentContractions -0.362431133

Table 7: Comparison of ablation results for binary classifiers on the Enron corpus, partitioned by pair

We observed a significant difference between raw and standardised scores for our 3-way classification, but breaking the task down into two binary stages shows that the difference is small for the hierarchical-level classifier, with most of the difference in the final result being made up by the up-down classifier.

We have previously noted that the Enron corpus is inherently unbalanced, with very different levels of production from different individuals, so the personal quirks of more productive individuals may be dominating the results. We are also unsure of the level of confidence we should have in the ground truth: that is to say, we do not know in which cases there are personal relationships running in parallel to the organisational relationships with which we are concerning ourselves.

4.2. Muir-Joinson Speech Corpus

Having observed mixed results on the Enron data, we now look to a more carefully synthesised dataset. As the Muir-Joinson speech corpus was generated under experimental settings, relationship status is exclusively determined by the experimental set-up, and not complicated by factors external to the task. It is our hope that this translates into greater confidence in the veracity of the ground truth, and consequently higher classification accuracy.

For the purposes of message-level categorisation, we consider a ‘message’ in this instance to be a single turn in the dialogue, which may consist of as little as a single word. The five-minute speed networking exchanges in our dataset consist of between 5 and 79 turns, with a median of 30 and mean of 30.8.

Having established that, as expected, random data partitioning generates artificially high results, we do not repeat this kind of partitioning for any further datasets. However, we generate both pair-partitioned and sender-partitioned versions of the dataset for testing, as these represent geniunely distinct use cases to be considered in any future application development.

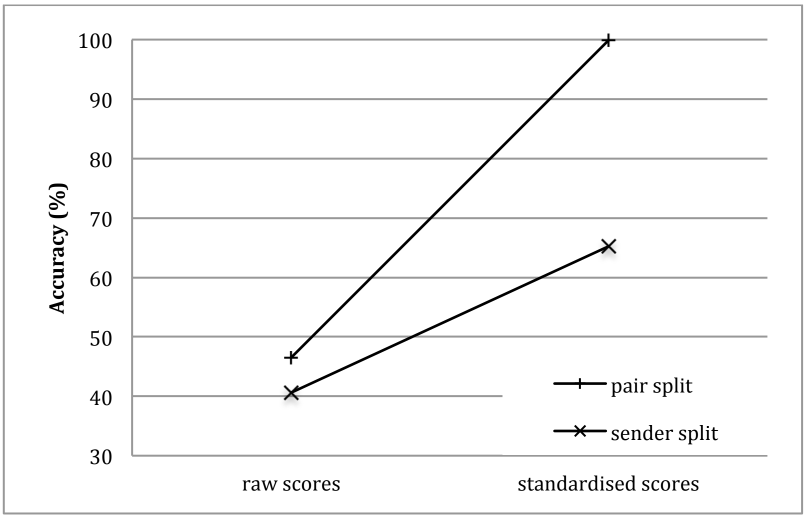

Figure 20: Effects of standardisation and partitioning on the UWE speech corpus

When the data is split using pairwise partitioning, message-level accuracy on the raw scores is 46.36%, using all features in the three-way classification task. Broken down further, this represents 45.95% accuracy for messages going up the hierarchy, 48.07% for downwards messages, and 45.22% accuracy for messages that formed part of peer-level exchanges.

The resulting confusion matrix (Table 8) does not show any clear patterns of mis-classification.

	upwards	downwards	level
upwards	**1748 (45.95%)**	910	1146
downwards	995	**2064 (48.07%)**	1235
level	1150	1282	**2305 (45.22%)**

Table 8: Confusion matrix for speech data, three way classification, partitioned by pair, without standardisation

	upwards	downwards	level
upwards	**4250 (99.93%)**	3	0
downwards	3	**4249 (99.84%)**	4
level	0	1	**4685 (99.98%)**

Table 9: Confusion matrix for speech data, three way classification, partitioned by pair, standardised by author

After standardising the data, we see an extreme improvement in the pairwise case, to 99.92%. There is little to be gleaned from a confusion matrix when so few examples are mis-classified, as we are not able to make statistically valid claims about the distributions of mistakes, but we present details here for completeness (Table 9). This improvement gives us very strong support for our second hypothesis; although we are working in a constrained scenario generated under experimental conditions, the fact that this level of accuracy can be obtained in these circumstances indicates the potential for achieving similar results in real-world data if we are able to identify the reasons for the difference in performance.

Under sender partitioning, the initial score is lower at 40.62%. We would always expect to observe comparatively lower accuracy for sender-partitioned data, as in this case we are building a model using one set of individuals, and using it to classify messages written by other authors whose language is not represented in the training set. Sender-partitioned results are therefore more dependent on commonalities and trends that hold true between individuals, whereas when messages are only partitioned by sender-author pair, we can use messages from A to B as part of the input to our model when classifying messages from A to C.

The increase after standardisation, while still significant, is less dramatic, giving a classification accuracy of 65.28%. In this case, we can see from the confusion matrix that the overall performance is being depressed by failures to correctly classify downwards messages, for which success is less than random even after standardisation. This may indicate that behaviour of superiors towards subordinates is something particularly personal to the individual; whether this is an artefact of the simulated environment under consideration here, or a more general phenomenon, is a question that can only be answered by future experiments.

	upwards	downwards	level
upwards	**3899 (91.68%)**	221	133
downwards	1353	**1217 (28.60%)**	1686
level	733	536	**3417 (72.92%)**

Table 10: Confusion matrix for speech data, three-way classification, partitioned by sender, standardised

Again, it may prove more informative to consider the two-class problems separately. Unfortunately, due to the implementation of the experimental set-up, there were no individuals who took part in exchanges under more than one role, so the dataset is somewhat artificially partitioned at this stage. We use only standardised results for the following analysis, as these gave the better performance, and we are interested in characterising the remaining errors.

For a binary classifier distinguishing hierarchical from level messages, our overall performance on the speech data, partitioned by sender, is 75.92% accuracy, with an F-measure of 0.770. Examination of the confusion matrix (Table 11) shows that level messages are more likely to be mis-classified as hierarchical.

For pairwise partitioned data, the headline statistics are more impressive: accuracy is 99.98%, with an F-measure of 0.9998.

	hierarchical	level
hierarchical	**6715 (78.92%)**	1794
level	1384	**3302 (70.47%)**

Table 11: Confusion matrix: speech data, hierarchical classification, partitioned by sender, standardised

	hierarchical	level
hierarchical	**8506 (99.97%)**	3
level	0	**4686 (100.00%)**

Table 12: Confusion matrix: speech data, hierarchical classifier, partitioned by pair, standardised

We can contrast this with the results of another binary classifier, classifying hierarchical messages as upward or downward. With data partitioned by sender, this gives an accuracy of 58.84% under 5-fold cross validation, with an F-measure of 0.585. The confusion matrix in this case is much more seriously skewed, with a strong bias towards classifying messages as upwards.

	upwards	downwards
upwards	**3811 (89.61%)**	442
downwards	3060	**1196 (28.10%)**

Table 13: Confusion matrix: speech data, up/down classifier, partitioned by sender, standardised

	upwards	downwards
upwards	**4247**	6
downwards	2	**4254**

Table 14: Confusion matrix: speech data, up/down classifier, partitioned by pair, standardised

The fact that this extreme distribution disappears when data is partitioned pairwise gives weight to our earlier hypothesis that poor results in the sender-partitioned case are down to individual differences when choosing how to address a subordinate individual, and indicates that this behaviour is consistent for the individual.

We do not learn much from the confusion matrices that we couldn't see from the three-way classification output. However, one particular advantage of splitting out the task into two consecutive hierarchical classifiers is that we are able to do feature analysis on a more fine-grained scale. For example, we can do feature ablation on each separately labelled dataset.

Starting with the hierarchical versus level classification, and the pairwise partitioned data, we observe that most individual features result in a decrease in performance. This is perhaps to be expected given the overall high performance of the classifier on this dataset. The surprising exceptions are removal of modals, emoticons, and character count, removal of which results in a small improvement in classifier performance. Table 15 shows the top 10 features for both binary classifiers, and full results are presented in Appendix B.

	Stage 1: Hierarchical v Level	Stage 2: Upward vs Downward
1st	PercentSemicolon -0.077177048	PercentPreposition -0.075331783
2nd	PercentNumWord -0.051128729	PercentRptLetterWord -0.066941229
3rd	PercentDeixis -0.038302739	PercentPeriod -0.064502573
4th	PercentContractions -0.038145048	PercentQuote -0.062475743
5th	*PercentExclamation -0.034248981*	*PercentPolite -0.061379583*
6th	*PercentPolite -0.033439983*	PercentPronoun -0.059400501
7th	PercentAdjective -0.033439983	PercentModal -0.045001668
8th	HeylighenDewaele -0.033302443	*CharactersPerWord -0.044575723*
9th	*CharactersPerWord -0.031362701*	NumParagraphs -0.038153907
10th	NumWords -0.031217085	*PercentExclamation -0.036697733*

Table 15: Comparison of ablation results for binary classifiers on the Muir-Joinson speech corpus, partitioned by pair

	Stage 1: Hierarchical v Level	Stage 2: Upward vs Downward
1st	PercentUppercase -3.644851158	PercentBracket 4.904938749
2nd	PercentQuote -2.616580655	NumParagraphs 4.988227656
3rd	PercentPronoun -2.418121512	PercentDeterminer 5.575987796
4th	PercentAdjective -1.973813703	PercentEmoticon 6.355616561
5th	PercentPolite -1.79687139	PercentInterjection 6.39301373
6th	HeylighenDewaele -1.517055488	PercentLetter 6.459406168
7th	PercentRptLetterWord -1.013506715	WordsPerSentence 6.545937893
8th	PercentHedges -0.934252051	PercentPeriod 7.279744221
9th	PercentNoun -0.279133156	PercentNoun 8.397261071
10th	PercentSemicolon -0.132629879	PercentAmpersand 8.576934892

Table 16: Comparison of ablation results for binary classifiers on the Muir-Joinson speech corpus, partitioned by sender

In the case of classifying upwards and downwards messages, similarly on pair-partitioned data, removal of character count again results in a small performance improvement, along with nouns, interjections, and a number of character-level features. However, the changes in this case are so slight (never dropping below 99.83% accuracy) that this is not a significant result.

In the sender split case, where the base accuracy is lower, ablation demonstrates some more dramatic effects (see Table 16). For the up/down classifier, removal of tag questions or quotation marks results in an improvement of sixteen percentage points; these are both comparatively rare features. It is somewhat troubling that punctuation features, and the choice of numerals over spelling numbers out in words, are at the discretion of the transcriber; it is therefore possible that this introduced inconsistencies into the data. In future data collection exercises, where speech data is used, consideration should be given to the level of guidelines provided to transcribers.

For hierarchical versus level communications, in the sender-split case, the most notable results are a decrease of 3.6 percentage points when the percentage of uppercase characters is removed, followed by 2.6 percentage points for quotation marks and 2.4 percentage points for pronouns.

4.3. Muir-Joinson CMC Corpus

As outlined in the Methodology chapter, our computer-mediated communication (CMC) data was generated under analogous circumstances to the speech data discussed above, utilising an instant message chat program. Therefore, similarities and differences observed between these two media are of particular sociolinguistic interest.

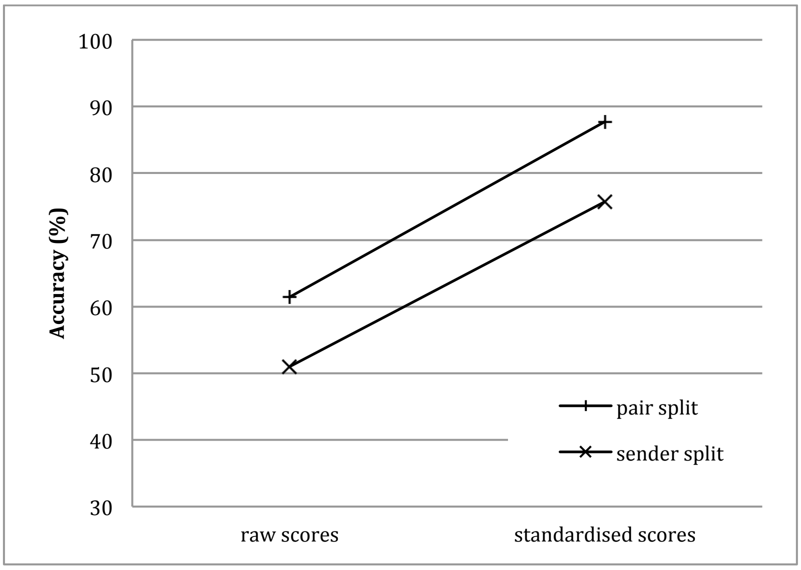

Figure 21: Effects of standardisation and partitioning on the UWE chat corpus

Accuracy for a three-way classifier trained on the raw data was 61.38% (f=0.581) in the pair-partitioned case and 50.98% (f=0.485) for the sender-partitioned case. After per-author standardisation, pairwise results increase to an accuracy of 87.66%

(f=0.876), while in the sender-partitioned case, accuracy improves to 78.73% (f=0.788). These improvements can be seen illustrated in Figure 21; comparison with Figure 20 shows that the relative increases for pair- and sender-partitioned data are more even for CMC than for speech data, as evidenced by the near-parallel lines in Figure 21.

	upwards	downwards	level
upwards	**515 (45.37%)**	607	13
downwards	581	**465 (43.79%)**	16
level	1	6	**3551 (99.80%)**

Table 17: Confusion matrix: CMC data, three-way classification, partitioned by sender, standardised

	upwards	downwards	level
upwards	**837 (73.74%)**	289	9
downwards	401	**652 (61.39%)**	9
level	1	1	**3556 (99.94%)**

Table 18: Confusion matrix: CMC data, three-way classification, partitioned by sender, standardised

From the resulting confusion matrices (Table 17 and Table 18) it is evident that level communications are the most successfully classified. This was also observed in the Enron data (the other written medium under consideration), and less so in the speech data. Unlike the Enron results, however, we do not see a bias towards mis-classifying hierarchical messages as level. If the mis-classification was due to an unbalanced training set, we would expect to see the opposite, with more bias in the CMC case, as experiemental setup resulted in almost half of all utterances occurring in a level context.

Another interesting observation is that the improvement from sender-partitioned to pair-partitioned results in the overall CMC results can be almost exclusively attributed to an improvement in classification between upwards and downwards messages, as level messages were already reliably classified. This effect is so clearly demonstrated in the output of three-way classification that there is no benefit to be gained by splitting out the binary classifiers in this instance.

However, for feature ablation, considering the problem as sequential binary classifiers is still a valuable exercise. Ablation results are not particularly informative for level messages, where classification accuracy is extremely high, but for distinguishing upwards from downwards messages there may be something of interest. For the pairwise partitioned data, the best performing features (giving the largest decrease in performance) are percentage of uppercase letters, percentage of words with affective lengthening, and words per sentence. The worst performance (where removal gives an increased accuracy) are percentage of deictic expressions, and percentage of determiners.

	Stage 1: Hierarchical v Level	Stage 2: Upward vs Downward
1st	PercentComma -0.067818426	PercentUppercase -2.121370899
2nd	PercentLetter -0.064420212	*PercentRptLetterWord* *-1.94643594*
3rd	PercentHyphen -0.019821307	WordsPerSentence -1.85723798
4th	PercentHedges -0.018420132	PercentContractions -1.812026778
5th	NumWords -0.013767956	PercentTagQ -1.777091166
6th	PercentVerb -0.011243222	PercentNumWord -1.477494125
7th	PercentExpletives -0.000261433	PercentModal -1.419236799
8th	PercentNonDictionary 0.002584095	PercentPolite -1.337869915
9th	PercentInterjection 0.010752369	NumParagraphs -1.332094323
10th	*PercentRptLetterWord* *0.012132198*	PercentBracket -1.315870399

Table 19: Comparison of ablation results for binary classifiers on the Muir-Joinson CMC corpus, partitioned by pair

	Stage 1: Hierarchical v Level	Stage 2: Upward vs Downward
1st	PercentTagQ -0.389083824	PercentUppercase 1.403049792
2nd	NumChars -0.387466749	*PercentAdjective* *2.103095786*
3rd	*PercentAdjective* *-0.367054455*	NumChars 2.379058292
4th	NumParagraphs -0.354132098	*PercentPolite* *2.382127483*
5th	*PercentPolite* *-0.352781607*	WordsPerSentence 2.386144115
6th	*PercentAlphanumeric* *-0.344271745*	*PercentQuote* *2.417578249*
7th	PercentPeriod -0.315650776	PercentLetter 2.574856522
8th	PercentHyphen -0.31307138	PercentVerb 2.616499481
9th	*PercentQuote* *-0.311075048*	*PercentAlphanumeric* *2.642746659*
10th	PercentSemicolon -0.301093108	PercentTagQ 2.646775826

Table 20: Comparison of ablation results for binary classifiers on the Muir-Joinson CMC corpus, partitioned by sender

With the data partitioned by sender, there were no features where removal gave a decrease in performance for the up/down case; the most significant increases were given by the removal of scores for contractions (as a percentage of parts of speech) and colons (as a percentage of punctuation), representing a performance improvement of 4.4 percentage points each.

The top 10 features in each case can be seen in Table 19 (partitioned by pair) and Table 20 (partitioned by sender), with the full ablation output supplied in Appendix B.

4.4. Discussion

We have constructed these experiments to test the following hypotheses:

> **H1:** Stylistic choice is a key method of expressing relationship-building, therefore relationship information can be inferred from linguistic style.
>
> **H2:** Differences from the individual's normal behaviour will be more informative than absolute feature scores, for predicting relationships.

From the results of our three-way classification experiments, using data partitioned by pair and which has not been standardised per author, we have demonstrated some success in our task, from 38.22% classification accuracy on messages from the Enron corpus, 46.36% on the Muir-Joinson speech corpus, and 61.38% on the Muir-Joinson CMC corpus. These all compare favourably to the random baseline of 33.3%, and approaching or above the most common class baselines which are 41.17% for Enron, 35.51% for Muir-Joinson speech, and 61.82% for Muir-Joinson CMC. This provides some support for H1, in general, although the unbalanced nature of the data means that the most common class baseline is a high one.

The consistent improvement when feature scores are standardised by author provides strong support to H2, across all datasets. For data partitioned by pair, once the feature scores have been standardised by author, we obtain results of 41.34% for the Enron corpus, 99.92% for the Muir-Joinson speech corpus, and 87.66% for the Muir-Joinson CMC corpus. We can also take these improved scores as additional support for H1, since we are still using stylistic features; we are simply manipulating the same scores, in an entirely predictable and repeatable manner, to produce the improvement in classification accuracy.

It is notable that we had greater success with the two experimental datasets than with the natural (Enron) data, which implies there may be some confounding factors for which we have not yet accounted, for example interference from other relationship types (friendship; emnity; romantic attachment) or social dynamics (social class; gender; age) not captured by the corporate hierarchy. Alternatively, it is possible that our approach is more effective on more informal text. To distinguish these alternatives, we would need to make use of further datasets from different genres, for example naturally occurring speech or CMC data, and experimentally-derived email data.

We have also conducted ablation experiments to look at the contribution of different features to the success of our classifiers. Perhaps the most interesting observation from these ablation tests is the level of variation between the lists generated under different conditions: there are no features that stand out as always providing a greater contribution, nor are there any that can be obviously discounted. This variation is observed not only across media, but for different conditions (e.g. binary versus three-way classifiers) on the same dataset.

5. Aggregation and Combined Approaches

In the previous chapter, we have undertaken categorisation of individual messages into three classes, according to whether the message is addressed up or down the organisational hierarchy, or whether it is a level communication between peers. We have also seen the improvements which derive from standardising feature scores by author.

We will now turn our attention to testing the intuition that we should be able to further improve our results by using more information to make fewer predictions, in other words, the effect of aggregation. We have two hypotheses to test in this space:

> **H3:** Prediction of hierarchical relationships will be more effective at the pairwise level than at the level of individual messages.
>
> **H4:** Relationship classification in a hierarchical situation will be aided by consideration of whole-network characteristics.

To test H3, we will take the outputs of message-level classification which we obtained in the previous chapter, and apply aggregation techniques (as described in section 3.5) to generate relationship-level predictions. If H3 is supported, we expect to see an improvement in classification accuracy at the relationship level.

To test H4, we take both message-level and relationship-level predictions, and use these to generate graphs representing our predictions. These graphs are then analysed using the mathematics of flows and tensions (Kochol 2004) to identify those parts of the graph in which our predictions are not self-consistent, as described in section 3.6. If H4 is supported, we expect to see higher confidence (which should translate into higher classification accuracy) for self-consistent sections of the graph, as reflected in the tension component.

5.1. Relationship-Level Categorisation

Having obtained a set of results for categorisation at the message level, we are now in a position to combine these outputs into a prediction at the level of the relationship, taking into account the sum of the evidence provided by the individual messages. If our third hypothesis is to be upheld, we expect this to generate an improvement in the overall classification accuracy.

It should be obvious that the quality of voting methods will depend on the accuracy of the underlying data points (i.e. the individual votes). As we obtained consistently better results after per-author standardisation, we use only standardised data in this chapter, to give ourselves a stronger starting point.

Our hypothesis is that aggregation should improve confidence in our results, with the consequence of higher overall classification accuracy at the relationship level.

5.1.1. Enron Emails

For the Enron data, we begin by considering aggregation of the pairwise-partitioned data, for which we observed a message-level accuracy of 41.34%.

Using simple plurality voting as our aggregation technique, we obtain a relationship-level accuracy of 38.28%, or 312 pairs correctly classified, which amounts to a drop of over three percentage points from the message-level accuracy. This is a surprising result that does not support our hypothesis that aggregation should improve accuracy (H3). Additionally, eighty pairs are not classified under a simple plurality voting system, due to having two (or occasionally three) classes with the same number of votes; this represents 8.9% of the 895 total.

Analysis of the errors shows that, as with the message-level results, level relationships are more often correctly classified, however this comes at the expense of incorrectly classifying a large number of hierarchical relationships into the level category.

85	**10.43%**	**Hierarchical, correctly labelled**
416	51.04%	Hierarchical, incorrectly labelled as level
28	3.44%	Hierarchical, labelled with incorrect polarity
227	**27.85%**	**Level, correctly labelled**
59	7.24%	Level, incorrectly labelled as hierarchical

Table 21: Analysis of errors, simple plurality voting, no threshold. Enron data, pairwise partitioned.

It is not meaningful to present a traditional confusion matrix in this context, as the difference between 'upwards, incorrectly labelled as downwards' and 'downwards, incorrectly labelled as upwards' is merely down to the random ordering of senders and recipients in the data files.

Using majority voting (a threshold of 0.5), accuracy drops further, to 38.36%, with 100 instances not classified, equating to coverage of 88.8%. Under this condition the number of pairs correctly labelled has dropped from 312 to 305.

Contrary to expectation, we find that adjusting the thresholds in general causes a (slight) decrease in accuracy: in this instance, requiring a higher level of agreement results in fewer correct instances getting through, as well as cutting out a number of incorrect instances.

Threshold:	0.5	0.6	0.7
Hierarchical, correctly labelled	**82** **10.31%**	**72** **9.61%**	**60** **8.85%**
Hierarchical, incorrectly labelled as level	409 51.45%	392 52.34%	356 52.51%
Hierarchical, labelled with incorrect polarity	26 3.27%	25 3.34%	25 3.69%
Level, correctly labelled	**223** **28.05%**	**211** **28.17%**	**193** **28.47%**
Level, incorrectly labelled as hierarchical	55 6.92%	49 6.54%	44 6.49%

Table 22: Analysis of errors under different threshold conditions. Enron data, pairwise partitioned.

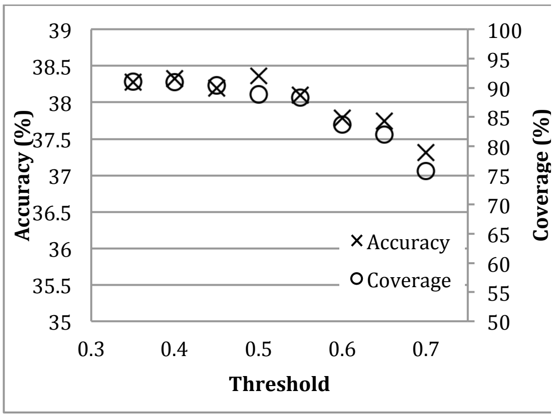

Figure 22: Effects of adjusting thresholds for the Enron data

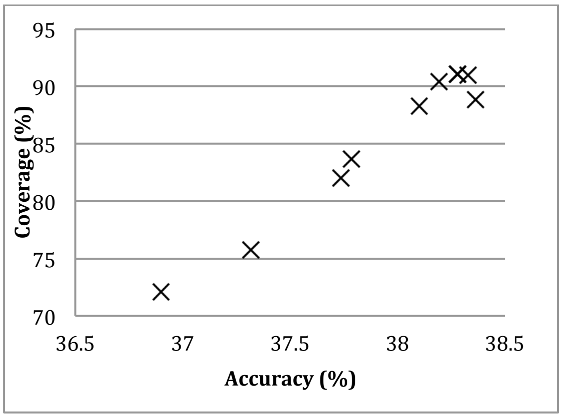

Figure 23: Accuracy against coverage for different thresholds

The Enron data is extremely unbalanced, so we should consider the possibility that distributional artefacts are affecting our results. In particular, some pairs have only exchanged one message, or have very unbalanced numbers of messages in the two different directions.

For 338 of the Enron pairs (37.8%) our dataset contains only a single message between the two interlocutors. Including these pairs in the aggregation means that we are including some message-level results in our aggregated results. We can adjust this by requiring a minimum number of messages before a pair is included in the relationship level calculations. Adjusting this minimum from one (the default) up to ten, we see a dramatic improvement at 2 and 3 (4.3 and 6.4 percentage points respectively), which then begins to level out, with only another 2.4 points' improvement when ten messages are required.

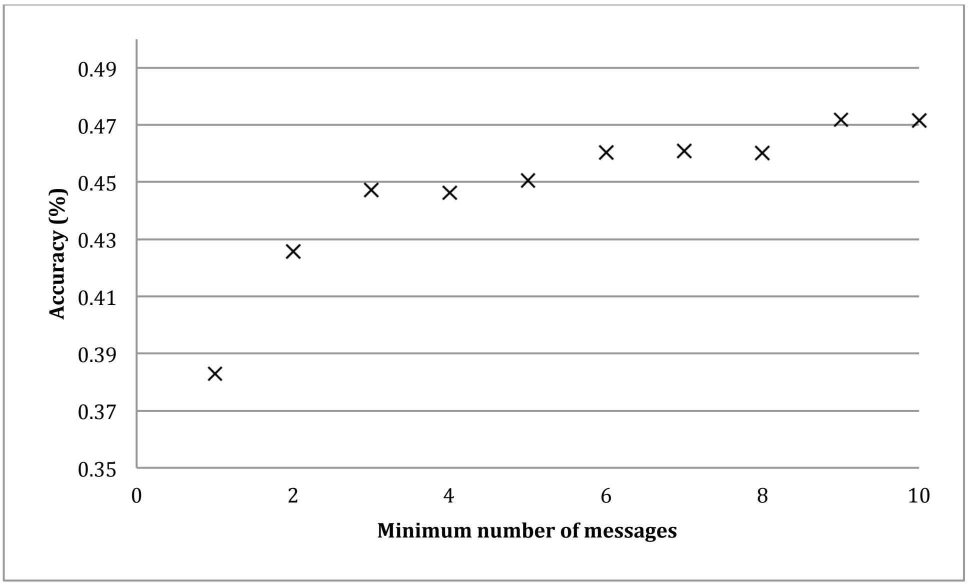

Figure 24: Effect on accuracy of requiring a minimum number of messages (Enron data)

Another way to account for imbalances in the data is to generate a prediction independently based on A's communications to B, and another based on B's communications to A, and then give these two predictions equal weight. This compensates for the weight otherwise given to highly productive individuals, whose productiveness may not necessarily be accompanied by the most typical linguistic choices.

Requiring agreement between A to B and B to A predictions, we observe an accuracy of 47.90%, which sounds like a significant improvement, however there are only 57 pairs classified under this methodology, amounting to coverage of only 6.37%.

In general, as results on this data are quite poor, aggregation appears to be compounding errors rather than compensating for them. In the next two sections we will consider the effect of aggregation on the two datasets where the individual message-level accuracy was higher.

5.1.2. Muir-Joinson Speech Corpus

We now turn our attention to the Muir-Joinson speech data. For message level classification, accuracy in the sender split case was 65.28% after standardising feature scores by sender. As we only have one five-minute interaction between each pair, we treated a single utterance as a 'message' for the earlier analysis. We are now endeavouring to characterise the relationship on the basis of all utterances making up the exchange.

Aggregation by simple plurality voting results in an improvement of 13.2 percentage points, giving an accuracy of 78.06% for those relationships which are labelled. In 18 cases, no decision can be made for the pair, as there are an equal number of data points indicating two or more classes; this equates to coverage of 89.5%.

92	**46.94%**	**Hierarchical, correctly labelled**
19	9.69%	Hierarchical, incorrectly labelled as level
16	8.16%	Hierarchical, labelled with incorrect polarity
61	**31.12%**	**Level, correctly labelled**
8	4.08%	Level, incorrectly labelled as hierarchical

Table 23: Analysis of errors, simple plurality voting, no threshold. Standardised speech data, sender split.

An analysis of the errors shows that hierarchical relationships are still the most likely to be misclassified, whether being misclassified as level, or assigned the incorrect polarity. Level relationships are less frequently misclassified as hierarchical.

By introducing a threshold to require higher levels of confidence, we can improve our accuracy, but only at the cost of significantly reduced coverage. This trade-off is entirely anticipated, and results in an approximately linear relationship (Figure 26).

Requiring majority agreement (a threshold of 0.5) results in an accuracy of 79.8%, which is an insignificant improvement, but accuracy then increases steeply, as shown in Figure 25. Coverage steeply declines at the same time. By the time the threshold reaches 0.6, we are observing accuracy of 94.30% but coverage of only 46.8%.

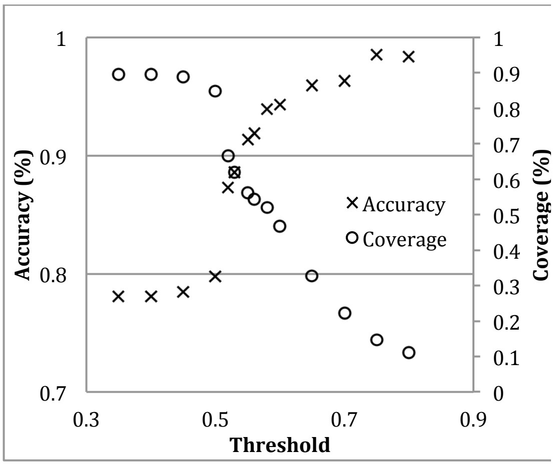

Figure 25: Effect of thresholds on accuracy and coverage

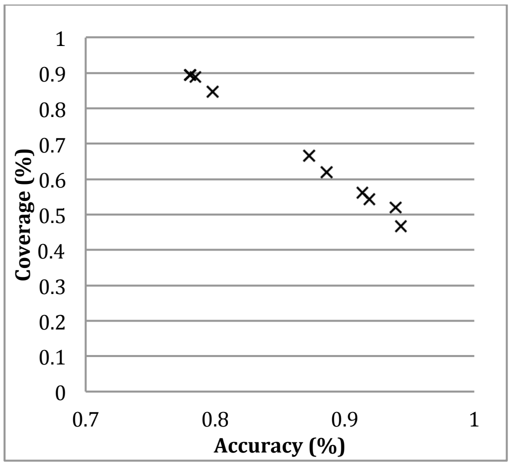

Figure 26: Accuracy versus coverage as thresholds are introduced to sender split speech data

The nature of data collection for this corpus means that the data is not unbalanced in the manner of the Enron corpus, as each interaction consists of approximately an equal number of pairs. It is therefore not useful to introduce a minimum number of messages for categorisation.

We will, however, investigate the impact of requiring independent agreement between the prediction based on the communications from A to B, and the prediction from B to A. This adjustment substantially reduces our coverage to 37.4% (from 89.5%), but increases our raw accuracy to 99.74% (from 78.06%, an increase of 21.7 percentage points). This implies that we can afford to be much more confident of those results where the two independent predictions are in agreement, even though we see fewer examples for which this is true.

We have examined the sender-partitioned case, which was the lower-scoring of our two scenarios. For the pairwise-partitioned data, there is not much space to improve from 99.92%, making the aggregation stage unnecessary for practical purposes. However in a realistic situation, classifying unknown data, we would not know the intermediate accuracy, so it is important to understand whether or not aggregation could reduce overall performance in this case. For completeness, we undertook simple aggregation and obtained 100.0%, successfully removing the small number of errors in the earlier classification without a reduction in coverage (no pairs are unlabelled in this instance).

5.1.3. Muir-Joinson CMC Corpus

Our CMC data is a very similar dataset to the speech, but due to the nature of online data collection (requiring participants to type their responses), a smaller amount of data was collected in the same amount of time.

For the sender-split data we observed message-level accuracy of 78.73%. Using simple plurality voting, this increases to 79.67%, with four pairs unlabelled (coverage of 97.3%).

Majority voting, requiring more than half of message-level predictions to agree, results in 80.34% accuracy with 95.3% coverage (eight pairs unlabelled).

Again, the threshold can be incrementally adjusted and the impact examined. Similar to the speech results, it is at thresholds above 0.5 that we start to see the most significant trade-off between coverage and accuracy, with accuracy steeply increasing as coverage declines. By 0.6 we observe an accuracy of 86.11% and coverage 75.6%; at 0.7, accuracy is up to 91.93%, covering 63.95% of pairs.

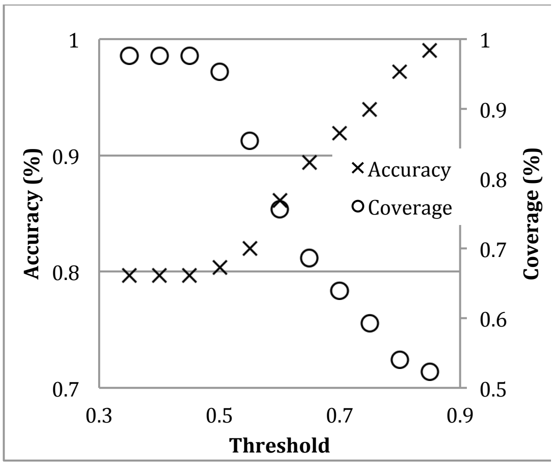

Figure 27: Impact of varying threshold, on accuracy and coverage

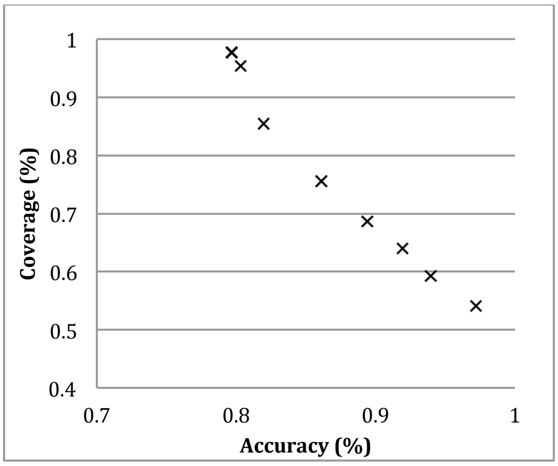

Figure 28: Accuracy against coverage for varying thresholds

Considering the requirement for agreement between the independent predictions in each direction, we achieve an accuracy of 97.66% with 40 unlabelled pairs, amounting to coverage of 78.5%. We note that for this dataset, this gives a better balance of precision and recall than applying a threshold to aggregation of all message-level scores. This may be because in CMC data, by contrast with speech, there can more easily be a significant imbalance in number of turns between the two parties to a discussion.

For CMC data we again had much higher message-level accuracy in the pairwise case, at 87.66%. Under aggregation with a simple plurality requirement, this increases to 98.92%, while maintaining 100% coverage.

5.2. Whole Network Analysis

We will now look at the extent to which performance can be improved by considering the whole network of communicants, and whether taking information from the structure of the network graph will give us any performance improvements. Our intuition is that we should be able to have more confidence in predictions which are consistent with one another, and therefore a simultaneous analysis of the whole network should give us more confidence in those predictions which reinforce one another. This gives rise to our fourth hypothesis:

> **H4:** Relationship classification in a hierarchical situation will be aided by consideration of whole-network characteristics.

We will see if this holds true in practice by the application of flows and tensions analysis (Kochol 2004), a mathematical technique which considers the composition of an entire graph in order to generate a component (the 'tension') in which every cyclical subgraph has a sum of zero. This is the graphical equivalent of ensuring that relationships between edges are transitive, so we anticipate this method should highlight any inconsistencies in our predictions. We will use this approach to analyse our predictions and identify the most self-consistent interpretation of the resulting graph.

There are two approaches to constructing the prediction graph, both of which we shall explore. One method is to use aggregate predictions (from section 5.1) as the weights for our links, with an edge weight of 0 between peers, and an edge weight of 1 for colleagues in a hierarchical relationship. The alternative approach is to use the message level predictions from Chapter 4 directly as weights on the edges of the graph, which gives a wider range of edge weights; if this were to be equally or more successful, and if it is able to match the improvements generated in section 5.1, it may be possible to skip the aggregation step entirely.

One of the major questions in applying flows and tensions analysis to this problem surrounds the interpretation of the numerical values returned in the tension component. In particular, in order to assess the impact of this additional stage on the accuracy of our predictions, we need to map these values onto our discrete categorisation of hierarchical-upwards, hierarchical-downwards, and level communication. We do this by setting an interval around the origin: messages within this interval are considered to be level. Above the threshold, messages are categorised as upspeak, and below it, downspeak. A threshold of zero translates to only edges with a score of precisely zero being classified as level, while all non-zero scores must be interpreted as hierarchical, with the polarity of the hierarchical relationship then determined by the sign of the value. We will investigate the effect on overall accuracy of varying this interval.

5.2.1. Enron Emails

We begin by constructing a power graph from the aggregated predictions from section 5.1, for the authors of the Enron corpus. We set the edge weight to 0 between two colleagues whose relationship is predicted as being level, and 1 between those in hierarchically-related pairs, with the direction of the edge indicating the direction of power in the hierarchical relationship (directionality is meaningless for an edge of weight zero). We use the pairwise-partitioned data, with feature values standardised by author, for consistency with the results presented in section 5.1. This graph is then fed into the flows and tensions algorithm, to measure the consistency between predictions. This returns the tension component of the graph, with numerical values on each edge representing the values thus calculated.

We now turn our attention to interpreting the numerical values returned by the algorithm. We can set a threshold about the origin above and below which we interpret scores to represent upward and downward relationships. A threshold of zero gives the poorest performance, with an accuracy 33.87%. This is worse than the accuracy obtained using simple plurality aggregation in section 5.1.1, which was 38.28%, and commensurately lower than the message-level accuracy of 41.34%. Accuracy improves gradually as this threshold value is increased, up to 0.5 (accuracy 48.39%, the highest we have obtained for this dataset), after which it decreases again. Note that as the threshold continues to increase, the limiting factor on the accuracy is the percentage of messages correctly assigned to the level class (which is not necessarily the most common class), as an arbitrarily high threshold value excludes any messages from being characterised as hierarchical.

As an alternative approach, we can begin with the individual message predictions between each pair (as generated in Chapter 4), and use these to determine the weights of the edges. Once we have calculated the corresponding tension component, we can vary the threshold in the same manner to map these results onto our hierarchical categories. We observe very similar results, with accuracy increasing as the threshold is increased from 0 to 0.5, although in this case the differences are somewhat less pronounced, with accuracy ranging from 40.32% at a threshold of zero, to 45.70% at 0.5.

Figure 29 demonstrates the effect on classification accuracy of varying the size of this interval, for both of these cases. The threshold value (X-axis) is half of the interval size, as the interval is symmetric about the origin (recall that positive and negative values are indicative of directionality on the edge).

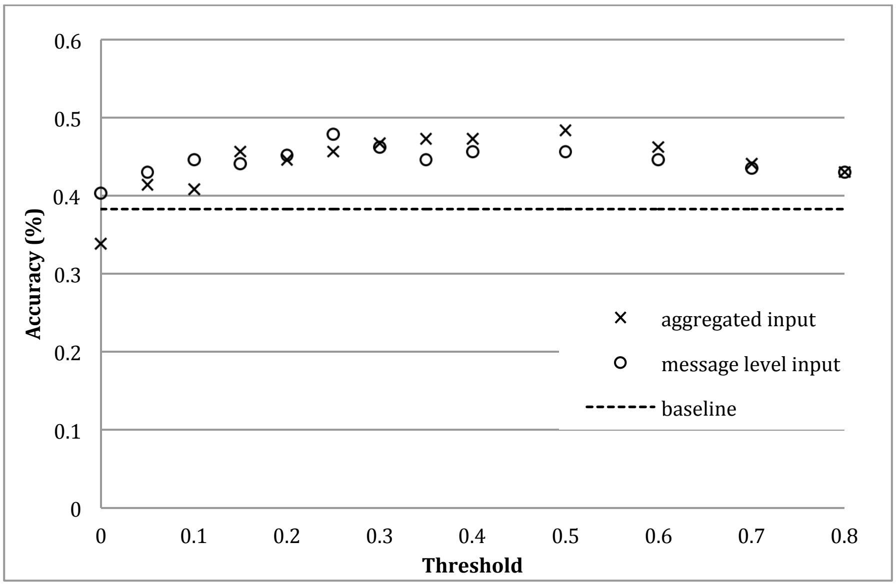

Figure 29: Effect of threshold choice on accuracy for flow-tension analysis, Enron corpus

The complete tension components, as calculated by both methods, are included in Appendix C.

5.2.2. Muir-Joinson Speech Corpus

We shall now repeat this analysis using predictions from the Muir-Joinson speech data. For speech data, partitioned by sender and with feature scores standardised on a per-sender basis, the message level accuracy reported in Chapter 4 was 65.28% (section 4.2), and simple plurality aggregation improved accuracy to 78.06% with coverage of 89.5% (section 5.1.2). We wish to see whether this can be improved by considering the shape of the resulting prediction graph.

We begin by using individual message predictions to weight the edges of our network graph, and applying flow and tension analysis. By this technique we obtain an accuracy of 59.14% for a threshold of zero, increasing to 72.04% at a threshold of 0.2. This is an improvement over the raw message level accuracy, but does not perform as well as aggregation.

When instead we use the aggregated predictions as the input from which we construct our graph, accuracy at the zero threshold is very similar at 57.53%. However, at a threshold of 0.2 the accuracy is 83.87%. Although we have obtained higher accuracy on this dataset in section 5.1.2, the notable feature of this graphical technique is that coverage remains at 100% throughout, rather than being traded off for accuracy.

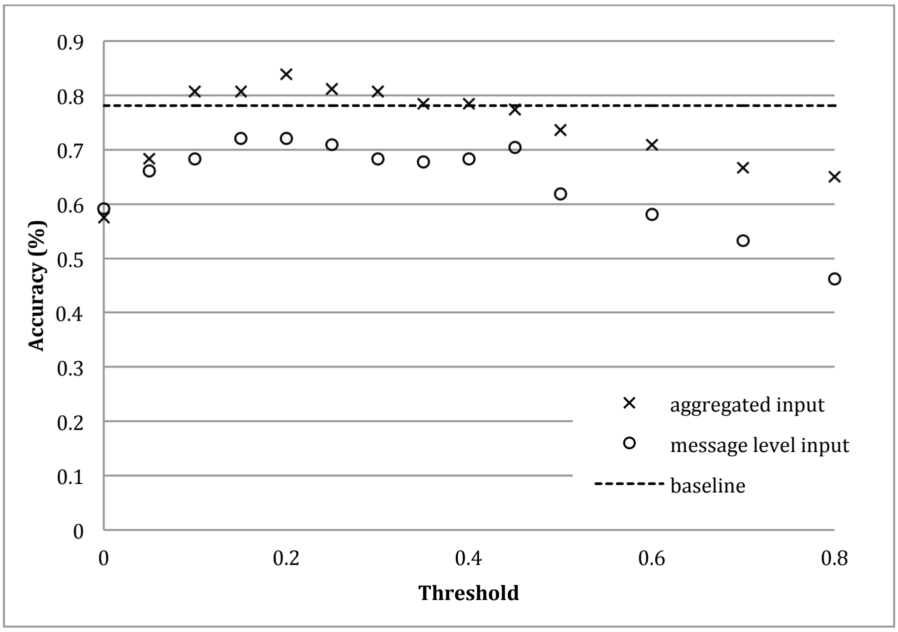

Figure 30: Effect of threshold choice on accuracy for flow-tension analysis, speech corpus

We note that the best-performing threshold in this case is much lower than the equivalent for the Enron data.

5.2.3. Muir-Joinson CMC Corpus

We next turn our attention to the Joinson-Muir CMC data. Message level accuracy was 78.73% for this dataset, when partitioned by sender and with feature scores standardised on a per-sender basis (section 4.3). This increased to 79.67% (coverage of 97.3%) under simple plurality voting (section 5.1.3).

We begin by using message level predictions to construct a weighted graph, and undertaking flow and tension analysis. With a zero threshold, analysis of the tension component gives us accuracy of only 26.34%. This increases dramatically to 77.42% at a threshold of 0.05, indicating that there are a large number of very small non-zero edge weights in this tension component. Accuracy then decreases steadily as the threshold is increased.

When the aggregated predictions from simple plurality voting are used instead, the accuracy at threshold zero is 82.26%, which then decreases as the threshold increases.

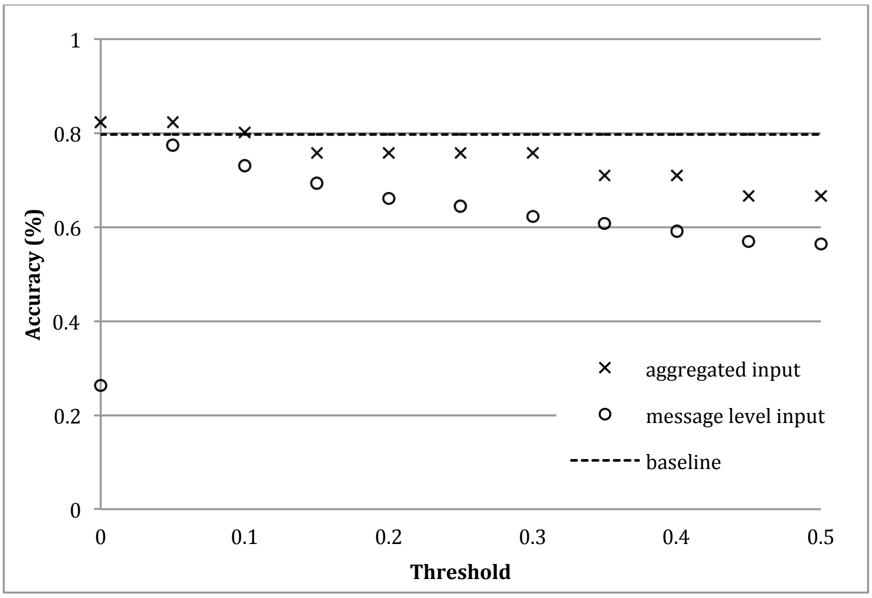

Figure 31: Effect of threshold choice on accuracy for flow-tension analysis, CMC data

Again, we observe that the pattern of accuracy as we adjust the threshold is inconsistent with the other datasets. This indicates that even though we can get some marginal improvements in accuracy by considering the whole network using graphical methods, the utility of this technique will be extremely limited in practice, as there is no straightforward method to interpret the results in the absence of existing ground truth.

5.3. Discussion

We have tested two very different methods for building upon message-level classification, with mixed results.

> **H3:** Prediction of hierarchical relationships will be more effective at the pairwise level than at the level of individual messages.

We have undertaken various methods of aggregating individual message-level predictions to provide predictions at the relationship level. We tested simple majority voting (using individual message predictions as input) as well as looking at the impact of requiring agreement between independent predictions from A to B and B to A.

Aggregating predictions across multiple messages had variable results, giving limited support for H3, the strength of which depended on the circumstances. In general, if message-level results were fairly accurate (e.g. for the Muir-Joinson speech and CMC transcripts), we obtained a further improvement by aggregation, but when message-level classification was weak (Enron), performance was degraded by aggregating these poor predictions.

Requiring agreement between individual predictions from A to B and B to A was the strongest technique for improving accuracy, resulting in an improvement in all cases, although this came at the cost of (sometimes extreme) reductions in coverage. This reduction was particularly extreme for email, as we do not have messages in both directions for all pairs in the Enron corpus, and these pairs were automatically excluded.

> **H4:** Relationship classification in a hierarchical situation will be aided by consideration of whole-network characteristics.

We then turned to a graphical approach, combining our individual message predictions and our relationship-level predictions (in two separate experiments) in the hope of improving accuracy by considering the structure of the whole graph. Although we observed improvements at certain thresholds, there was no universally best threshold value, meaning this technique is of limited use in practice. The fact that some improvements were observed at certain thresholds does provide a very limited amount of support for H4, suggesting that further work in this area might potentially be profitable. However, this is unlikely to be the first avenue of approach to further study, as other areas seem likely to produce greater benefits in the short term – for example, the use of aggregation on datasets where the base classification accuracy is sufficiently high.

Given the extremely high success rates for classification on both sets of Muir-Joinson transcripts, and the fact that aggregation improves these good results yet further, it is worthwhile designing future work to investigate methods for understanding the confidence of our results, perhaps via a bootstrapping approach. Knowing when we can have confidence in our predictions is perhaps the most powerful approach we could use to improve overall performance.

6. Linguistic Accommodation and Power[5]

Having examined the possibility of modelling social power using the distribution of linguistic style features, we now turn to the study of accommodation in linguistic style choice, in order to account for the extent to which individuals' use of stylistic features is influenced by the linguistic style of their conversational partner. We set out to test our fifth hypothesis:

> **H5:** Linguistic accommodation behaviour is partially motivated by relative social power, therefore greater accommodation of linguistic style will correlate with lower social power.

Linguistic accommodation is a well-recognised indicator of social power and social distance (Giles et al. 1991). However, different individuals will vary their language to different degrees, and only a portion of this variance will be due to accommodation. This is an analagous phenomenon to that which we have already seen, as we have observed that individuals exhibit different levels of variation in their range of scores for each feature. In our exploration of H4, we demonstrated that this individual variation has a significant impact on our ability to predict social roles correctly, if this is not accounted for in the model. We must therefore be similarly cogniscent of individual variation in the context of linguistic accommodation: some individuals will be more inclined than others to accommodate their language. We wish to systematically measure the proportion of variance which is due to linguistic accommodation, and consider the meaning of this within the social context.

This chapter presents analysis using the Zelig Quotient (Jones et al. 2014), which is a novel method of measuring how an individual orients their linguistic style to align with a particular audience, using the author's other communications as a baseline. This metric was developed in order to compensate for some of the limitations of existing techniques such as Language Style Matching (LSM; Gonzales et al. 2010). In particular, we developed the Zelig Quotient to address two key concerns with current approaches: that direction of accommodation (who is moving their style towards whom) is typically ignored in favour of absolute distances, and that an individual's tendency to accommodate is not modelled under any earlier techniques. To improve our understanding of the phenomenon of linguistic accommodation, we model linguistic variation as movement in vector space, using an individual's prior communications to calculate a baseline for each feature and their personal variance to calculate the significance of any subsequent movement.

As discussed in section 3.7, we calculate the Zelig Quotient using a set of nine function word categories from LIWC (Pennebaker et al. 2007) to enable comparison with LSM. It follows that the work described in this chapter uses a more constrained feature set than the remainder of this thesis. The nine feature categories used are

---

[5] This chapter is based on work undertaken in collaboration with Kate Muir, Simon Jones, Adam Joinson (University of the West of England), and Nigel Dewdney (University of Sheffield) which has been published in the following papers: Muir et al. 2015; Muir et al. 2016a; Muir et al. 2016b.

auxiliary verbs, articles, common adverbs, personal pronouns, indefinite pronouns, prepositions, negations, conjunctions, and quantifiers. Additionally, we perform this analysis using only the Muir-Joinson datasets (described in sections 3.1.2 and 3.1.3), where we have accompanying personality data to further inform our analysis.

6.1. Muir-Joinson Speech Corpus

We calculated linguistic accommodation between participants in conversations in the Muir-Joinson speech corpus, using the Zelig Quotient. Recall that for this dataset we have disjoint sets of participants recorded under two conditions: the hierarhical (experimental) condition, where volunteers were randomly assigned to be either judges or workers, and the non-hierarchical (control) condition, where volunteers were asked to discuss the same ideas as equal collaborators.

Figure 32 presents the pairwise speaker-to-recipient ZQs for judges (high power) versus workers (low power), as a percentage of the total number of conversations. It is clear to see from this illustration that the workers in general were more accommodating than the judges. This is backed up by the overall statistics: workers had a higher mean Zelig Quotient (μ = -0.11, σ = 0.12), while judges exhibited more divergent behaviour (μ = -0.26, σ = 0.11). A one-way ANOVA revealed social power was a significant influence on ZQ ($F(3,37)$, p=0.4, η^2=0.22).

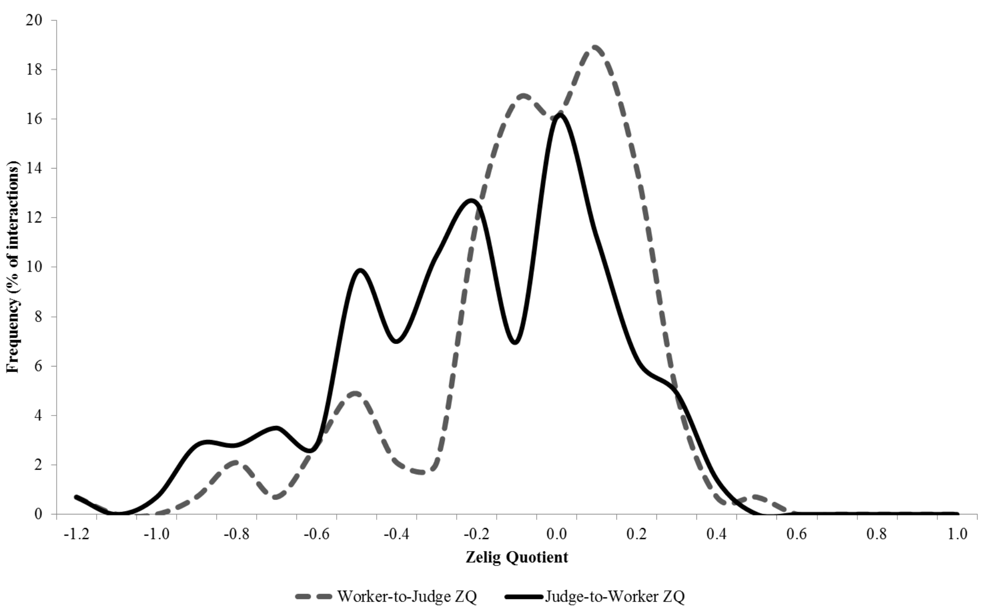

Figure 32: Pairwise speaker-to-recipient Zelig Quotient distributions for conversations between workers (low power) and judges (high power) in the Muir-Joinson speech corpus.

Compare these results to the control condition, where individuals were asked to have collaborative discussions as equal partners, rather than being placed in positions of high or low social power. In this case, we see that collaborators exhibit a Zelig Quotient average in between that of judges and workers, and somewhat closer to that of judges (μ = -0.21, σ = 0.17).

It is noteworthy that divergence was very common across the corpus. This is an unexpected result, as it is usually convergence which is associated with positive outcomes, yet divergent conversations were not ranked as negative by participants. It is possible to pick out more nuance in this distribution as it relates to social power, since judges exhibited divergence in 65% of conversations, while workers diverged in only 45% of conversations. Conversely the distribution of positive ZQ, indicating convergence, saw convergence by workers in 39% of conversations, and convergence by judges in only 24%. We contrast this to collaborators in the non-hierarhical control group, who showed similar levels of divergence across the two groups: Group A exhibited divergence in 63% of conversations and convergence in 23%, while Group B exhibited divergence in 58% of conversations and convergence in 25%, in both cases exhibiting behaviour closer to that of judges than that of workers.

By considering the self-reports that individuals gave about the quality of their interactions, we are also able to examine whether linguistic accommodation is correlated with individuals' feelings about their conversational partner. We asked participants to complete a short questionnaire after every interaction in which they reported their subjective opinions of interaction quality and impression formation. Participants were asked to score their similarity to their partner, the level of rapport or 'click', their partner's conversational ability, and also to score their partner's social attractiveness, physical attractiveness, and task attractiveness. Most of these factors did not correlate with ZQ, but there are a small number of significant exceptions. For example, when workers were more Zelig-like, judges rated them to have higher social attractiveness and gave a higher score for the level of rapport. With every standard deviation increase of a worker's ZQ above the group mean, judges' rating of rapport increased by 2.5 points and their rating of social attractiveness increased by 1 point. This correlation is not replicated for workers' assessment of judges: we found no correlation between a judge's ZQ and how their partners scored them on interaction quality or impression formation. For equal collaboration, the highest rapport ratings were associated with low to moderately divergent behaviour, while higher or lower Zelig scores both indicated lower rapport. Additionally, conversational ability scores were predicted by ZQ in all social roles: for every standard deviation increase in ZQ, conversational ability ratings increased by 1.5 points.

We also collected personality traits for all participants in the speed networking experiments, which enables a more nuanced analysis of the influences on an individual's tendency to accommodate. We used model 1 of the PROCESS macro for SPSS (Hayes 2013) to conduct a series of moderation analyses, and found that several personality traits were predictive of ZQ. With increasing agreeableness, impression management, and self-consciousness, overall ZQ decreased, whereas ZQ increased with increasing leadership, Machiavellianism, and self-monitoring. These personality traits were also shown to be significant moderators of the relationship between social power and overall ZQ. Specifically, for those in the worker role (low social power), participants who were lower in agreeableness, impression management, and self-consciousness, were more likely to accommodate their linguistic style compared to those in a high-power role. There was less effect of personality on ZQ for collaborators (equal power), which suggests that the combination of social power and personality is a potent one which leads to measurable effects.

For hierarchical interactions we have a measure of task success in the form of whether a judge awarded extra money to each worker as a result of their pitch, and how much. There is not a direct and straightforward relationship between ZQ (of either judge or worker) and the financial outcomes, however there is a weak correlation between financial reward and the judges' ratings of some of the positive social outcomes described above which were correlated with ZQ (conversational ability, similarity, and rapport) as well as one with no ZQ correlation (task attractiveness).

For comparison, we also calculated language style matching (LSM) between participants. Because this is a distance rather than a vector calculation, we do not have a measure of directionality, but we can compare individuals under our experimental and control conditions. We found that LSM was not significantly different between judge-worker pairs ($\mu = 0.73$, $\sigma = 0.12$) and pairs of collaborators ($\mu = 0.71$, $\sigma = 0.11$). Nor did a pair's LSM score correlate with any of the subjective measures of interaction quality (such as similarity, social attractiveness, or task attractiveness) or objective measures of task success (money awarded).

6.2. Muir-Joinson CMC Corpus

The Muir-Joinson CMC data was collected to be deliberately analagous to the earlier speech experiments, to allow us to repeat the above experiments in a different medium. Previous work has found that accommodation behaviours are not always replicated between speech and analagous text-based discussions (Gonzales et al. 2010), while other scholars have observed accommodation in online chat media (Huffaker et al. 2011; Riordan et al. 2013), so we are interested in the extent to which these effects are observed in our case. As above, we calculated linguistic accommodation between each pair of individuals using the Zelig Quotient. So far, we have only undertaken analysis of participants in the hierarchical condition; comparsion to a control group is intended for future work.

Figure 33 presents the pairwise speaker-to-recipient ZQs for judges (high power) versus workers (low power), as a percentage of the total number of conversations. In common with the speech results, we see that workers had a greater tendency to accommodate, while judges were more likely to diverge. For this dataset we observe divergence by judges in 63% of conversations, and divergence by workers in 57% of conversations, though this difference is not statistically significant. Convergence by judges is observed in 25% of conversations, and convergence by workers in 31%. The overall mean Zelig Quotient for workers in the CMC corpus was $\mu = -0.16$ ($\sigma = 0.07$), while judges tended to be more divergent with a lower mean ZQ, $\mu = -0.23$ ($\sigma = 0.14$). Consistent with speech results, although both judges and workers exhibit a negative mean ZQ, the judges are consistently more likely to diverge. As with the speech data, linguistic style divergence is exhibited in a large number of conversations overall.

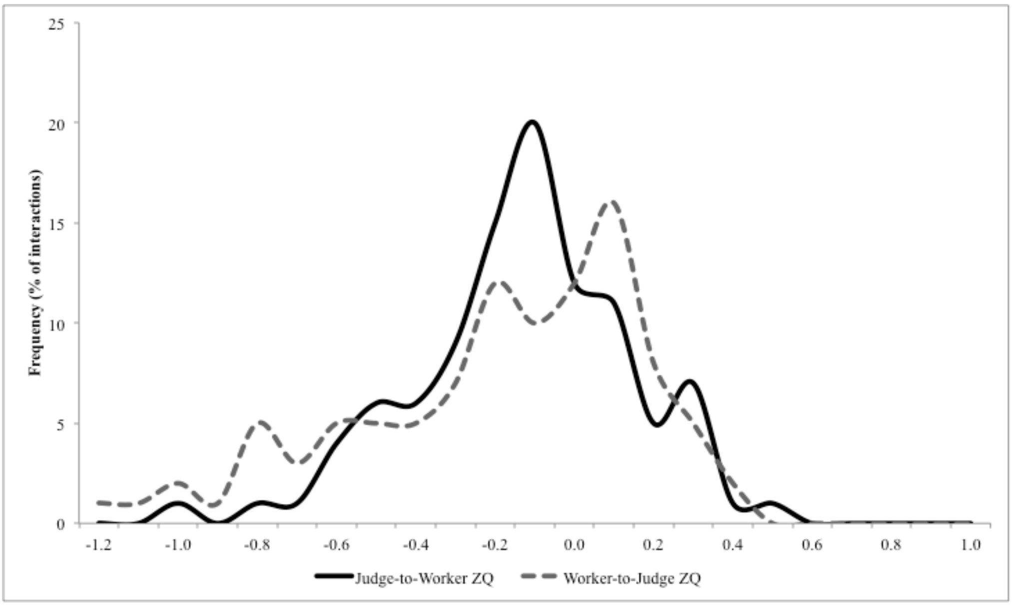

Figure 33: Pairwise speaker-to-recipient Zelig Quotient distributions for conversations between workers (low power) and judges (high power) in the Muir-Joinson CMC corpus.

We also investigated the effect of linguistic style accommodation on individuals' ratings of their partner and the quality of the interaction. We did not find any relationship between accommodation by workers, and how they were perceived by the judges. However in the inverse case, the level of linguistic style accommodation by judges was significantly and negatively correlated with workers' perception of them: as judges' ZQ increased, workers were likely to rate them lower on similarity, rapport, and social attractiveness. This is the opposite of what we observed in the face-to-face experiments, where it was workers' level of accommodation which had an effect on how they were perceived.

To complement these results, we repeated the CMC data collection using a within-subjects design where the same individuals took on both the worker and the judge role during the course of the speed networking session (the experimental setup is described more fully in Muir 2016b). In so doing, we were able to measure the extent to which individuals' accommodation behaviour was consistent between roles, and whether there was any effect of changing roles during the experiment. As with the earlier data collection, we asked participants to rate their personal power under both roles, and this manipulation check found that participants reported significantly greater power in the judge role ($\mu = 4.23$, $\sigma = 0.77$) than in the worker role ($\mu = 3.60$, $\sigma = 1.06$).

Figure 34 gives the distribution of ZQ observed in this within-subjects experiment. In common with the between-subjects experiment, judges' style diverged in a higher proportion of conversations (62%, versus 43% for workers), and judges had a lower mean Zelig Quotient ($\mu = -0.22$, $\sigma = 0.17$) than workers ($\mu = -0.09$, $\sigma = 0.11$) which also indicates greater divergence by those in a role of higher power.

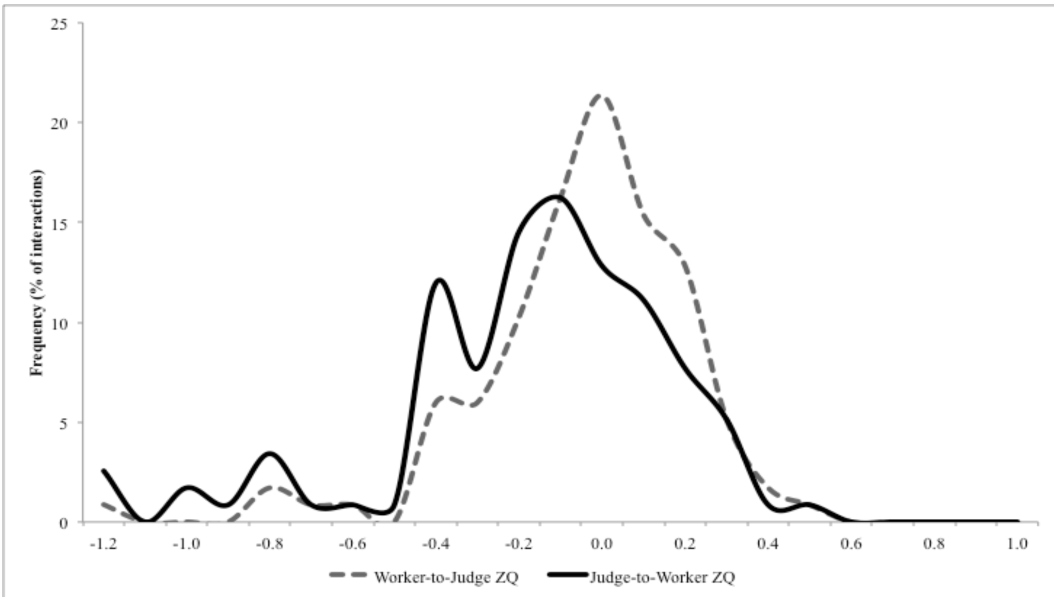

Figure 34: Pairwise speaker-to-recipient Zelig Quotient distributions for conversations between workers (low power) and judges (high power) in the within-subjects CMC experiment.

In future work, we will investigate whether personality traits mediate the effect of linguistic accommodation in CMC, in a manner consistent with that which we have observed for speech. Since other results have been broadly consistent between the two different genres, it is reasonable to expect that the same personality traits might be predictive of linguistic style accommodation. We will also examine the behaviour of the control group.

6.3. Discussion

We have examined linguistic accommodation in order to test the hypothesis that accommodative behaviour would correlate with social power:

> **H5:** Linguistic accommodation behaviour is partially motivated by relative social power, therefore greater accommodation of linguistic style will correlate with lower social power.

We have found some significant correlations between role and accomodation behaviour, in both speech and CMC experiments, which gives support to H5. In both cases, we found that individuals with more social power have less need to make use of linguistic accommodation strategies, which is in line with theoretical predictions. Additionally we have identified some personality traits which mediate these effects in our speech data, and which interact with social power in predictable ways. Having identified some specific interactions between power, personality, and language, it will be possible to define experiments to test more detailed hypotheses in this area in future work.

In Chapters 4 and 5, we set out to predict social power based on an individual's use of linguistic style features, with results that were promising but with scope for further improvement. By examining linguistic accommodation, we have identified another way in which linguistic style is influenced by social role, and by the language of others. Future work should look to combine these approaches, considering both the variation due to accommodation and the variation due to role, in order to build a more comprehensive and integrated model. Accommodative behaviour could then be incorporated as part of a predictive model of social power. Complexities in implementing this approach arise from the way personality and power interact to predict accommodation. We have conducted our studies so far on data gathered in laboratory conditions, where we were able to ask participants to complete personality questionnaires. In real-world contexts, where we would not expect to have personality inventories for individual authors, there is a research question to be addressed around whether automated personality classification from text (as reported by Oberlander & Gill 2006; Nowson & Oberlander 2007; Iacobelli et al. 2011) can be reliable enough to be used as an input to a predictive model.

We have also observed some surprising results, namely the prevalence of negative accommodation (divergence) in the whole dataset. This is unexpected as in the literature, divergence is traditionally associated with negative social outcomes such as perceptions of hostility and impoliteness (Giles & Gasiorek 2014).

However, ours is not the first experiment to observe divergence of linguistic style in the context of social power differentials. Kacewicz et al. (2013) reported that individuals who rated themselves as higher in power within a dyad conversing using an online chat-room used fewer first person singular pronouns compared to their lower power partner. This is hypothesised to reflect a greater focus on the other when in a high power role, compared to greater focus on the self in a low power role; in our study, differing use of personal pronouns by individuals in high versus low power roles would manifest as objectively measured divergence in linguistic style, but actually is in line with the individuals' social power roles. Similarly, divergence in linguistic style has been observed in messages between high and low ranking community forum members (Jones et al. 2014).

The concept of speech complementarity could account for the observed divergence in linguistic style in judges and workers (Dragojevic et al. 2015). This refers to instances where divergence is consistent with social roles. For instance, males and females have been observed to diverge from each other in their speech behaviours (such as tone), in order to remain consistent with traditional sex role stereotypes (Giles et al. 1991). Speech complementarity is an explanation that carries particular weight where there is a power differential between conversationalists, as is the case in our work. For example, in doctor-patient interactions, doctors (in the more dominant position) have been found to produce more questions and talk for longer, whereas patients were more submissive, in line with their help-seeking role (Street 1991). Speech complementarity thus reflects and reinforces social differences; if both parties expect and prefer communicative differences, speech divergence will be positively received.

The asymmetrical divergence exhibited by workers and judges is in line with predictions from communication accommodation theory, suggesting low social power triggered motivations in workers to gain the approval of the higher power partner, leading to relatively greater accommodation in linguistic style. Increasing linguistic style accommodation by low power participants also negatively correlated with their ratings of interaction quality. Hypothetically, linguistic style accommodation by low power participants could represent an attempt to repair or improve a perceived lack of rapport with their higher power interlocutor. This would be consistent with research indicating behavioural mimicry is sometimes used as a strategy to repair a failed attempt at affiliation (Lakin & Chartrand, 2003).

An alternative account is that workers were exhibiting reluctant accommodation (Soliz & Giles, 2014). Workers were dependent on positive evaluations from judges for a good outcome, in terms of being awarded extra pay. Thus, workers may have accommodated reluctantly, due to cultural norms and outside pressures, instead of responding to internal motivations to affiliate. Reluctant accommodation is usually negatively associated with evaluations of the relationship and recipient (Soliz & Giles, 2014), which is consistent with our findings.

This is very early work in the arena of relating linguistic style accommodation to social power, and has raised a number of questions which can only be answered by further work. We have demonstrated some consistent results across conversations in two different media, but we have also laid the groundwork for this methodology to be applied to other datasets in future.

7. Feature Distribution Analysis

We have conducted a variety of experiments to test our hypotheses, with varying levels of success. However, negative results are not an obstacle to scientific progress; if we can understand why our results are not entirely in line with our predictions, we can make better predictions in future. A little post hoc analysis is therefore in order. In particular, we will examine the feature distribution in our datasets in further detail, in the hope that this may shed some light on the composition of our data, and any effects this may have had on our results.

Exploratory data analysis is often taken as the first stage in a new task, but we opted not to do this as we wanted to select features on a theoretical basis, and measure the resulting performance without biasing our approach with prior knowledge of the specific feature distributions within our data. This is particularly important as we are working with three separate corpora across different genres, and we do not easily have access to additional datasets meeting our requirements for the task.

However, having obtained a range of classification results, we will now briefly turn our attention to examining the data itself in more detail. This may highlight some interesting features, or shed light on whether communicants are behaving in our data as we might expect from previous studies or from linguistic theory.

A selection of accompanying charts are presented in Appendix D, beginning on page 170.

7.1. Parts of Speech

Interjections are one part of speech that are typical of informal spoken exchanges, yet which are still observed in textual sources, particularly in the case of online dialogue.

Across all three datasets, interjections are used less often in upspeak than in downspeak, as illustrated in Table 24, although this difference is more pronounced in speech, where $\mu = 0.095$ for upwards communication and $\mu = 0.259$ for downwards communication (μ throughout this chapter represents the mean feature score for the category under discussion; in this case this is the percentage of tokens that are tagged as interjections). This indicates that those with power may feel more entitled to interrupt or speak over their subordinates, in line with theoretical predictions. In level communications, the values fall between these two extremes. In email, the one medium where an interjection can't actually be used to interrupt someone, interjections make up an extremely small proportion of text; they are more prevalent in CMC, but still comparatively rare compared to their use in spoken dialogue.

	Speech	CMC	Email
Upwards	μ =0.095	μ =0.029	μ =0.005
Level	μ =0.132	μ =0.033	μ =0.006
Downwards	μ =0.259	μ =0.035	μ =0.010

Table 24: Comparison of feature values: interjections

Another interesting feature of interjections is that they are comparatively often the only term in a single-word utterance. In fact, of 1954 single-word turns observed in the speech data, 73.5% consist of an interjection. This can be contrasted with 15.9% of single-word messages in the CMC data and 14.6% for Enron, although in both written formats, single-word turns are overall much less prevalent.

Modal verbs are often used as an indirect politeness strategy, particularly for requests (e.g. questions of the form 'can you...?' or 'would you...?'). As such, it is unsurprising that they are observed most often in upwards communication, across all three datasets, although in CMC the downwards percentage is the same (μ =0.020).

Indeed, across the parts of speech distributions for our CMC data, it is broadly the case that the upwards and downwards distributions are very close, with the level score standing out. For example, verbs and adjectives are comparatively underused in level CMC communication, as are determiners, prepositions, and pronouns. Conversely, nouns are comparatively more prevalent, occurring with about 1.4x the frequency in level communications.

By contrast, in speech data, it tends to be the downwards messages which feature the more unusually distributed parts of speech, with below-average use of verbs and nouns, conjunctions, determiners, and prepositions.

7.2. Polite Expressions, Hedges, Deixis, and Tag Questions

Unlike part of speech features, which between them must sum to 100%, there is no necessary relationship between the proportions of different lexical politeness and deference features, except insofar as they cannot sum to more than the total number of lexemes. (Percentages are expressed as a proportion of words in the utterance.)

Of those expressions we categorised as explicitly polite (such as 'please' and 'thanks'), the highest proportion was observed in email, followed by CMC, with speech trailing as the least polite medium. Contrast the mean values for upwards communication, which are typical: μ =0.018 for email, μ =0.009 for CMC, and μ =0.006 for speech.

By contrast, in our data, hedges are observed comparatively more frequently in CMC and speech, with speech having the highest frequencies, whereas email is much lower than both. As hedges are generally considered a negative politeness strategy, we would expect to see more in upspeak, but while this is true of speech (μ =0.036 upwards, versus μ =0.023 downwards and μ =0.030 level), we actually see more downward hedges in both the written media.

Deictic expressions were observed to be more frequent in email and speech, compared to CMC. In particular, level CMC had the lowest frequency of deictic expressions. In speech, by contrast, deixis was observed most often in level contexts, closely followed by upwards.

Tag questions were rare across the board, but were most common in downspeak, across all media. This is the opposite of what might have been expected from a perspective of tag questions as a politeness strategy, and implies that those in more senior positions might be using the form of tag questions to confirm that requests will be complied with.

7.3. Innovative Forms and Out of Vocabulary terms

Innovative linguistic forms are more prevalent in online contexts, so it should come as no surprise that emoticons and affective lengthening are only observed in CMC text. Indeed, the transcription methodology means that such features could never be seen in the speech data, and the Enron corpus largely predates affective linguistic innovation of this kind. Surprisingly, in the CMC data, emoticons were observed in upward and downward contexts, but not in level exchanges. Affective lengthening is observed only in CMC turns, but is seen in all power contexts, although more commonly in level communications.

By contrast, consider the distribution of alphanumeric words. We had previously imagined that these would follow a similar distribution, considering cases such as ‘l8r’ which are innovative forms belonging to textspeak and the online environment. However, we did not account for other alphanumeric strings: in the Enron context, these include flight numbers or stock market codes, which results in an unexpectedly high incidence of this feature for Enron data.

For out of vocabulary terms, overall, we note that these might be technical jargon just as easily as slang: this seems the most likely explanation for the comparative prevalence of out of vocabulary terms in the Enron data. There was also a surprisingly high rate of out of vocabulary incidence in the speech data, considering that this was professionally transcribed, and is therefore less likely to represent typographical errors. Again, this is most likely to be accounted for by slang or technical jargon.

7.4. Contractions and Expletives

Contractions are only moderately indicative of informality, and are generally accepted as standard language these days. Nevertheless, we observe a much higher rate of contraction usage in speech across all three power distributions, followed by CMC, with the lowest occurence in email. This is in line with the traditional view that writing is more formal than speech, and that synchronous online chat is “speech-like” in some respects.

At the opposite end of the spectrum, expletives remain amongst the most informal language, as even use of a taboo word requires a degree of trust. We saw the highest proportion of expletives in speech, which is spontaneous and where self-editing is not effective (even if someone apologises for their language, the original words will remain in the transcript). In email, by contrast, we saw very few examples, most of which were between peers at the same level within the organisational hierarchy.

7.5. Punctuation and Case Selection

Punctuation is a linguistic feature that can be the subject of almost unlimited variation, allowing the creative author to express a huge range of styles.

Exclamation marks are perhaps the most inherently informal of punctuation marks. They are most commonly seen in CMC, especially in downward communication. In the Enron email data, exclamations are not very common, but when they do occur it is more often in level communications, which fits with the informality of this punctuation choice. In speech transcripts, where punctuation is at the discretion of the transcriber choice, we can only assume this is a proxy for tone of voice – exclamation marks are rare in this case, and almost entirely limited to upwards speech.

The respective distributions of commas and periods are generally uninteresting, but both show one feature that is an artefact of taking scores as a proportion of a small number of instances: in this case, occurrence of a particular punctuation mark as a proportion of all symbols. This is visible as distinct spikes in the bins that encompass 1/2 and 1/3. Both commas and periods are comparatively little used in CMC, when contrasted to email and speech distributions; this is true across all power scenarios.

Case distribution shows the most variation for CMC data, as would be expected for innovative use. In particular, the spike of uppercase use in level CMC is likely to be when capitalisation is used for emphasis.

7.6. Message Length Features

It is to be expected that both character and word counts exhibit strong variation by genre, since the length of an email is effectively unlimited, while speech and CMC are both synchronous media in which short turns are the norm. There is also a variation by relationship status: in both speech and chat, the longest turns are observed going up the hierarchy, while level and downward utterances skew shorter.

Similarly for the case of words per sentence, Enron data exhibited a broader range, going up to longer sentences, while speech and CMC both tended towards shorter sentences, particularly in level and downwards communications.

Characters per word, a commonly used proxy for complexity of vocabulary, also shows a clear genre bias: the longest words are used in email, followed by CMC, and finally speech. Within genres, although variation was observed, there was no consistent pattern between different power scenarios.

7.7. Discussion

We have found that our data exhibits distributions that are largely in line with our expectations, but there are a few interesting points to draw out. Surprises included the fact that tag questions were more common in downspeak than in turns addressed upwards or level, and that emoticons were used in upwards and downwards computer-mediated communications but not in peer-to-peer exchanges. The distribution of alphanumeric strings within the Enron corpus was also unexpected, and can be attributed to cases such as stock market codes and flight numbers, which are not the kind of linguistic innovation we had been hoping to capture with this feature.

In general, emails from the Enron corpus exhibit a lower incidence of many features that would usually be associated with informality in the literature. At the lexical level there are fewer interjections, fewer contractions, and very few expletives, while at the character level there are fewer exclamation marks, and less evidence of case being used innovatively to convey emphasis. In terms of message length, both speech and CMC exhibited longer turns going upwards, while this variation was also absent in Enron emails. This points towards a possible explanation for the poorer performance of our models on Enron data, and may indicate that our approach in general will be more successful for informal, ‘chatty’ media than for longer, and potentially more self-edited, exchanges; future work may wish to explore this in more depth, with an eye to developing a model which is more effective for email and similar media.

8. Conclusions & Future Work

8.1. Summary of Findings

Considering our three datasets represent different genres and collection methodologies, there are some surprising commonalities in the results. We will now consider the level of support for each of our starting hypotheses.

> **H1:** Stylistic choice is a key method of expressing relationship-building, therefore relationship information can be inferred from linguistic style.

Our results show varying support for H1 depending on the source of the data, the partitioning, and the standardisation, but in all cases performance was above the random baseline, and above the most common class baseline once feature scores had been standardised by author. This varied from weak success for Enron emails, to extremely strong results in the speech and CMC data.

With only a single dataset from each genre, we cannot test whether the medium itself is responsible for these differences in performance. It is possible that this represents a genuinely stronger signal in speech and "speech-like" chat data, contrasted with the more highly edited medium of email – this would accord with theoretical predictions, but we cannot state this with any degree of certainty. Other possible reasons for poorer performance on the Enron data include uncertainty of the hierarchy, the unique situation of Enron as a failing company, or the diversity of topics and dialogue acts represented in the corpus. Future work should look to identify more datasets across different genres and scenarios.

Although we have attempted to separate stylistic features from topical content, this is by no means a perfect separation. To give one example that has come to light since the beginning of these experiments, we imagined that the use of alphanumeric words would be an informative stylistic choice, but in the Enron context this is also a topical feature linked, for example, to stock prices and flight numbers. This specific case could be addressed in future by some specialised entity extraction alongside the existing classifiers. However, this is a single instance of what could be a more general problem: there is a chance that our results are contaminated by some element of topic. However, the fact that we are successful in the speed-networking scenarios, where topic is controlled by the experimental setup, suggests that this is not a major problem, and we can safely interpret our classification success as supporting H1.

> **H2:** Differences from the individual's normal behaviour will be more informative than absolute feature scores, for predicting relationships.

To test H2, feature scores were standardised on a per-author basis, and all classification experiments were repeated using the author-standardised scores. These were compared to the initial classification outputs, and in all cases, relative results were improved by standardisation. It is particularly notable that this method is effective at improving results even in the experimental data conditions when we only have data from each individual acting in one particular role.

This gives strong support for H2, as differences between an individual author's choices in different scenarios are emphasised when scores are standardised in this manner. Indeed, without taking account of these improvements due to standardisation by author, we would have only very limited support for H1.

> **H3:** Prediction of hierarchical relationships will be more effective at the pairwise level than at the level of individual messages.

To test our third hypothesis, we undertook various methods of aggregating individual message-level predictions to provide a prediction at the relationship level. We then compared these results to our message level predictions, to see if we could obtain greater accuracy by using more information.

Aggregating predictions across multiple messages had variable results, giving limited support for H3, the strength of which depended on the circumstances. In general, if message-level results were fairly accurate (e.g. for the Muir-Joinson speech and CMC transcripts), we obtained a further improvement by aggregation, but when message-level classification was weak (Enron), performance was degraded by aggregating these poor predictions.

Requiring agreement between individual predictions from A to B and B to A was the strongest technique for improving accuracy, resulting in an improvement in all cases, although this came at the cost of (sometimes extreme) reductions in coverage. This reduction was particularly extreme for email, as we do not have messages in both directions for all pairs in the Enron corpus.

To make use of aggregation in an operational environment, it would be necessary to first have some confidence that the message level results were good enough. Future work could investigate more thoroughly the level of individual message accuracy required to justify undertaking aggregation.

> **H4:** Relationship classification in a hierarchical situation will be aided by consideration of whole-network characteristics.

We did not have success with the graphical approach as a method of bootstrapping our individual message predictions. Although there were improvements at certain thresholds, there was no universally best threshold value, meaning the technique is of limited use in practice. The fact that some improvements were observed at certain thresholds does provide a very limited amount of support for H4, suggesting that further work in this area might potentially be profitable. However, this is unlikely to be the first avenue of approach to further study, as other areas seem likely to produce greater benefits in the short term.

> **H5:** Linguistic accommodation behaviour is partially motivated by relative social power, therefore greater accommodation of linguistic style will correlate with lower social power.

We have seen support for H5 as we observed correlations between role and the accomodation of linguistic style, in both speech and CMC experiments. However, the overall prevalence of negative accommodation (divergence) was unexpected, as divergence is traditionally associated with negative social outcomes. It is possible that this is explicable via the concept of speech complementarity (Dragojevic et al. 2015), as there is a power differential between conversationalists, and therefore divergence could be positively received as this reinforces the . Additionally, we have identified some personality traits which mediate the effects of linguistic style accommodation.

Overall, we have demonstrated that there is some validity to the assumption that stylistic choices reflect social roles. In particular, we have seen that differences from an individual's personal norms are particularly informative. This provides empirical support for the theory of identity as a social construct which changes according to context. We have also seen that individuals' accommodation of linguistic style is mediated by not only their social role but their personality, pointing towards a rich seam of research for future investigation.

It is particularly interesting to note that participants in the Muir-Joinson experiments exhibited stylistic differences between roles, as they were asked to adopt a role for the purpose of the experiment. The fact that their unconscious use of language adjusted, even in light of a temporary role, is evidence in support of the social constructionist view of identity.

We have also demonstrated the possibility of predicting relationships based on others within the same community of practice, which could be developed into an application for analysing inboxes. For the individual user, this might provide insight into how others perceive them. For law enforcement, mapping out the hierarchy of a group could assist with identifying key individuals; stylometric analysis might be particularly useful in identifying those who wield de facto social power without being explicitly named as a leader.

8.2. Future Work

Part of the reason for limiting this study to stylometric features was to enable the creation of models with greater explanatory power. However, predictive power is also important, particularly for real-world applications.

By separating out stylistic features, we have managed to account for some of the linguistic variation observed between individuals performing different roles, thereby demonstrating that style does change in line with seniority. However, for our experiments on the Enron corpus, previous studies using large n-gram models have demonstrated a higher accuracy than we have been able to obtain using stylistic features alone. This implies that there are other content-based features which are contributing to overall classification in the n-gram studies. Additionally, our examination of feature distributions in Chapter 7 indicates that authors within the Enron corpus make less use of many informal linguistic styles. To explore this problem further, while retaining the emphasis on explanatory power, it is necessary to identify some additional classes of content features which could be combined with stylistic features in a more comprehensive model.

One natural area to explore would be the intersection of topic and style. It is likely that topics of conversation naturally differ between roles and ranks, whether that is colleagues collaborating on a project, line managers discussing day-to-day issues of performance and progression, or senior leadership setting direction; at the same time, some topics such as planning travel or making lunch plans are likely to be common at all levels of an organisation. Unsupervised techniques in topic modelling, such as Latent Dirichlet Allocation (Blei et al. 2003), have seen a lot of success in recent years. LDA generates feature vectors representing topics by looking for patterns of word co-occurrence, and thus naturally discounts function words and grammatical features which occur uniformly across topics. This is orthogonal to our approach, since we have optimised for features with low semantic content, and hence could provide a complementary approach. LDA could be used to produce topic vectors from dialogue turns, and it would then be possible to look for correlations between topic and style features, and to study the intersection of these feature spaces with regard to predicting positions in the organisational hierarchy.

The distribution of dialogue acts is another aspect of language that we might reasonably expect to vary across roles. Actions such as giving instructions or requesting information may be intrinsically more or less pertinent to different positions within an organisation, or appear in different combinations. Previous work has looked at various methods of tagging dialogue acts by form or function (Core & Allen 1997; Bennett & Carbonell 2005; Lampert et al. 2006) , and any of these approaches could be combined with our stylistic methods.

Other studies have had success at divining organisational roles from network graphs of interactions. This could be another fruitful area of intersection, combining content-based features with the metadata of a contact graph.

It would be interesting to look for correlations between stylistic feature scores and other features of interest such as topic models, dialogue acts, and metadata features. Subsequently, it would be informative to examine whether taking these features in combination might result in a more nuanced approach to classifying relationships.

At the opposite end of the spectrum, further light may be shed on the nature of stylistic variation by subdividing the feature space into logical partitions (such as ‘parts of speech’ and ‘linguistic innovation’) to investigate which feature subsets contribute most to the classification exercise.

From a practical perspective of improving the classifier in those contexts where performance is poorest, one potential approach would be to adopt a bootstrapping approach, using verified data to incrementally re-train the model. For example, we have already seen that classification performance is improved when we have some existing truth data relating to particular individual and their relationships (i.e.

pairwise partitioning outperforms sender partitioning). When a new individual A is introduced to the network, assuming they are in contact with some known individuals, we can begin by calculating their relationships with those about whom we already have information, which we can infer with higher confidence. These data points could then be used to build a model for A's other interactions. A similar process to be followed for each individual added to the network, beginning with those who have the most contacts with previously well-known individuals. We thereby incrementally extend our coverage of the communications network, adding only the most confident predictions each time. In this manner, it may be possible to improve the performance of a classifier in the real world, without materially altering the approach.

We also wish to extend our work on linguistic accommodation, both to improve our theoretical understanding of this topic, and with the ultimate goal of incorporating this into a more complete model of how linguistic style covaries with social power.

Finally, we have considered only one class of personal relationship: the case of static, hierarchical relationships in business contexts. Future work could look in more detail at a more finely-grained model of organisational roles and situational power. Alternatively, moving beyond the business world, it may also be of interest to examine relationships between friends, acquaintances, and family members.

9. Appendices

Appendix A: Implementation Details for Feature Extraction

This appendix contains details of our implementation for various features described in Chapter 3.

A.1 Regular Expressions for Emoticons

These regular expressions are written for the Java regex engine, as this was the programming language used throughout our experiments.

```
EMOTICON_REGEX1 = Pattern.compile("(:[',]?[-o]?[(|Ss])");
EMOTICON_REGEX2 = Pattern.compile( "([:;8Bb=][-o]?[)(|])");
EMOTICON_REGEX3 = Pattern.compile("([:;8=][-o]?[ODSs])");
```

A.2 Deixis Word List

I	he	that	then
me	she	here	tomorrow
you	they	there	yesterday
we	this	now	

A.3 Hedging Word List

a bit	hypothetically	perhaps	theoretically
a little	I feel	plausibly	typical
according to	I mean	possibly	typically
actually	I think	potentially	usual
allegedly	I believe	practically	usually
apparently	ish	probably	with respect
appears	it feels	provisionally	with all due
approximately	kind of	rather	respect
although	largely	reportedly	would
basically	literally	reputedly	you know
by the way	mainly	roughly	afaik
clearly	maybe	said	ianal
could	may	seems	iirc
depends	might	should	istr
depending	more or less	slightly	aiui
essentially	mostly	somewhat	imo
evidently	of course	sort of	imho
fairly	often	supposedly	
figuratively	on the whole	tentative	
generally	overall	tentatively	

A.4 Politeness Word List

thanks
thank
thankyou
please
sorry
apologies

A.5 Expletives Word List

arsehole
ass
assfuck
asshole
asswipe
bastard
berk
biatch
bitch
bloody
bollicks
bollocks
boner
boob
bugger
buggery
bunghole
cacker
christacrutchian
cock
cocksmoker
cocksucker
coon
cornhole
crap
crapper
cum
cunt
dick
dickwad
dildo
dingleberry
dink
dipshit
dong
doo-doo
faggot
fart
felch
feltch
flange
fuck
fucked
fucker
fucking
fugly
gangbang
gook
groe
grostulation
gummer
hodgie
honkey
hork
hummer
jackshit
jizz
jizzum
kike
knockers
krunk
meatrack
nigga
nigger
nookey
noonan
nooner
nudger
pecker
phungky
piss
pissed
poop
poopshoot
prick
pubes
pussy
queef
quim
reestie
shit
shitcan
shitfaced
shitfit
shithead
shithouse
shitlist
shits
shitstain
slapper
slut
sod
spooge
spunk
stiffy
telesis
threesome
tinkle
tit
tittie
tosser
turd
twat
twink
wanker
wetback
whiz
willie
wordhole
wuss
zipperhead

A.6 Contractions Word List

ain't
aren't
didn't
doesn't
don't
can't
couldn't
hadn't
hasn't
haven't
isn't
mayn't
mightn't
mustn't
needn't
oughtn't
shan't
shouldn't
wasn't
weren't
won't
wouldn't
could've
must've
would've
I'm
we're
you're
he's
she's
it's
they're
I've
we've
you've
they've
I'd
we'd
you'd
he'd
she'd
they'd
I'll
we'll
you'll
he'll
she'll
they'll
let's
that's
there's
what'll
what're
who's
who'll
who're
who’d

Appendix B: Feature Ablation Results

This appendix contains a detailed breakdown of our feature ablation experiments. We include all results for completeness. Negative scores represent a decrease in performance when the feature in question is removed, indicating a valuable contribution; by contrast, positive scores represent improved performance when the feature is removed.

B.1 Feature Ablation – Enron Email Corpus

B.1.1 3-way classifier

-1.002628778	PercentInterjection
-0.967610098	PercentAlphanumeric
-0.942892817	NumWords
-0.83748455	PercentPronoun
-0.750977306	PercentVerb
-0.709427324	PercentRptLetterWord
-0.661889212	PercentDeixis
-0.559713517	PercentExpletives
-0.545357911	NumSentences
-0.513978689	PercentPeriod
-0.490934461	CharactersPerWord
-0.474982764	PercentLetter
-0.462722864	PercentModal
-0.435519729	PercentTagQ
-0.422224876	PercentExclamation
-0.415350113	PercentQuote
-0.379844833	NumChars
-0.359545409	PercentConjunction
-0.339620261	PercentPreposition
-0.286706954	PercentNonDictionary
-0.260511509	PercentHyphen
-0.231606706	PercentDeterminer
-0.213837612	PercentQuestionmark
-0.200213255	WordsPerSentence
-0.195756446	PercentNumWord
-0.017191706	HeylighenDewaele
0.021732501	PercentAdjective
0.056209755	PercentComma
0.254938469	PercentAdverb
0.286644487	PercentBracket
0.353026145	PercentContractions
0.375160364	NumParagraphs
0.411849967	PercentNoun
0.420607632	PercentEmoticon
0.475078286	PercentPolite
0.610698387	PercentAmpersand
0.720780253	PercentHedges
0.850938919	PercentUppercase
0.977837786	PercentSemicolon
1.018941344	PercentColon

Table 25: Ablation results for the Enron emails, 3-way classifier, partitioned by pair, with standardised feature scores.

B.1.2 Hierarchical vs Level Classifier

-2.16149784	PercentSemicolon
-1.767487066	PercentQuestionmark
-1.49902995	CharactersPerWord
-1.466898264	NumSentences
-1.439738716	PercentDeterminer
-1.340865885	PercentAdjective
-1.18488448	PercentHyphen
-1.158088441	PercentExclamation
-1.154132564	PercentUppercase
-1.151816837	PercentPreposition
-1.148395071	WordsPerSentence
-1.092842262	PercentInterjection
-1.026481334	PercentNumWord
-0.960872793	PercentExpletives
-0.921422489	PercentColon
-0.916231601	PercentLetter
-0.795583115	PercentBracket
-0.704117499	PercentAlphanumeric
-0.689501724	HeylighenDewaele
-0.672620128	PercentTagQ
-0.671976906	PercentAdverb
-0.667054696	PercentConjunction
-0.659534961	PercentModal
-0.655449174	PercentPronoun
-0.642690268	NumParagraphs
-0.625612724	PercentComma
-0.610059262	PercentPeriod
-0.583006966	PercentNonDictionary
-0.553435355	PercentDeixis
-0.551909183	PercentNoun
-0.518402427	PercentQuote
-0.498683389	NumChars
-0.483115057	PercentHedges
-0.453987112	PercentEmoticon
-0.403850087	PercentAmpersand
-0.384724025	PercentContractions
-0.330046384	PercentVerb
-0.292658697	PercentRptLetterWord
-0.276326436	NumWords
-0.137912733	PercentPolite

Table 26: Ablation results for the Enron emails, hierarchical vs. level classifier, partitioned by pair, with standardised feature scores.

B.1.3 Upwards vs Downwards Classifier

-0.986540324	PercentQuote
-0.723805324	NumSentences
-0.635587306	PercentExclamation
-0.608054922	PercentSemicolon
-0.544628664	PercentColon
-0.516221101	WordsPerSentence
-0.469954731	PercentAlphanumeric
-0.392623103	PercentAmpersand
-0.370076188	PercentLetter
-0.362431133	PercentContractions
-0.302135435	NumChars
-0.205930837	PercentPolite
-0.192070266	PercentDeixis
-0.13925655	PercentPeriod
0.002507123	PercentEmoticon
0.023164517	PercentQuestionmark
0.10232386	HeylighenDewaele
0.122005319	PercentInterjection
0.213517944	PercentComma
0.222396998	PercentNumWord
0.412301566	PercentHyphen
0.451406419	PercentModal
0.472661579	NumWords
0.474608779	PercentVerb
0.506547766	PercentHedges
0.570979489	CharactersPerWord
0.668677635	PercentPreposition
0.670947327	PercentTagQ
0.74347844	PercentConjunction
0.758142792	PercentExpletives
0.774670838	PercentAdjective
0.906062014	PercentNonDictionary
0.922483247	PercentRptLetterWord
0.944555666	PercentUppercase
0.958717406	PercentNoun
1.221690358	PercentDeterminer
1.386489378	PercentAdverb
1.430880589	NumParagraphs
1.871099492	PercentBracket
2.114725748	PercentPronoun

Table 27: Ablation results for the Enron emails, upwards vs. downwards classifier, partitioned by pair, with standardised feature scores.

B.2 Feature Ablation – Muir-Joinson Speech Corpus

B.2.1 Hierarchical vs Level Classifier, Split by Pairs

-0.077177048	PercentSemicolon
-0.051128729	PercentNumWord
-0.038302739	PercentDeixis
-0.038145048	PercentContractions
-0.034248981	PercentExclamation
-0.033439983	PercentPolite
-0.033439983	PercentAdjective
-0.033302443	HeylighenDewaele
-0.031362701	CharactersPerWord
-0.031217085	NumWords
-0.031039864	NumParagraphs
-0.029621926	PercentRptLetterWord
-0.029061793	PercentAmpersand
-0.028488495	PercentNonDictionary
-0.027776305	PercentTagQ
-0.027306256	PercentHyphen
-0.027070556	PercentQuote
-0.025852676	WordsPerSentence
-0.025279378	PercentComma
-0.023629777	PercentPronoun
-0.022824817	PercentBracket
-0.019483249	PercentAlphanumeric
-0.019483249	PercentVerb
-0.018301011	PercentPreposition
-0.018301011	PercentNoun
-0.01769207	PercentUppercase
-0.017492012	PercentPeriod
-0.016687052	PercentConjunction
-0.016400949	PercentExpletives
-0.015700834	PercentColon
-0.013113822	PercentDeterminer
-0.010890924	NumSentences
-0.010890924	PercentLetter
-0.008722467	PercentHedges
-0.006695588	PercentQuestionmark
-0.003535278	PercentInterjection
-0.000326161	PercentAdverb
0.005665987	PercentModal
0.006043266	PercentEmoticon
0.013021633	NumChars

Table 28: Ablation results for the Muir-Joinson speech corpus, hierarchical vs. level classifier, partitioned by pair, with standardised feature scores.

B.2.2 Upwards vs Downwards Classifier, Split by Pairs

-0.075331783	PercentPreposition
-0.066941229	PercentRptLetterWord
-0.064502573	PercentPeriod
-0.062475743	PercentQuote
-0.061379583	PercentPolite
-0.059400501	PercentPronoun
-0.045001668	PercentModal
-0.044575723	CharactersPerWord
-0.038153907	NumParagraphs
-0.036697733	PercentExclamation
-0.036687389	PercentAdjective
-0.03362249	HeylighenDewaele
-0.031059811	PercentAdverb
-0.028060843	PercentColon
-0.028060843	PercentVerb
-0.028005257	PercentNumWord
-0.026538738	PercentContractions
-0.026538738	PercentNonDictionary
-0.024871923	PercentDeixis
-0.019904217	WordsPerSentence
-0.019223657	PercentBracket
-0.018437699	PercentComma
-0.015004506	PercentHedges
-0.013682698	PercentDeterminer
-0.002516761	PercentTagQ
-0.000984311	PercentConjunction
0.003679409	PercentAmpersand
0.003905056	PercentQuestionmark
0.004575272	NumWords
0.010628794	PercentLetter
0.017857257	PercentAlphanumeric
0.018067898	NumChars
0.019389706	PercentExpletives
0.021056522	PercentUppercase
0.028150618	PercentSemicolon
0.029683067	PercentInterjection
0.041072588	NumSentences
0.059556206	PercentEmoticon
0.061223022	PercentHyphen
0.06531815	PercentNoun

Table 29: Ablation results for the Muir-Joinson speech corpus, upwards vs. downwards classifier, partitioned by pair, with standardised feature scores.

B.2.3 Hierarchical vs Level Classifier, Split by Sender

-3.644851158	PercentUppercase
-2.616580655	PercentQuote
-2.418121512	PercentPronoun
-1.973813703	PercentAdjective
-1.79687139	PercentPolite
-1.517055488	HeylighenDewaele
-1.013506715	PercentRptLetterWord
-0.934252051	PercentHedges
-0.279133156	PercentNoun
-0.132629879	PercentSemicolon
-0.114695554	PercentAmpersand
-0.049532843	PercentHyphen
0.060052487	NumChars
0.308410633	PercentContractions
0.541388088	PercentComma
0.550628734	PercentEmoticon
0.600065742	PercentExclamation
0.731715899	PercentConjunction
0.820869342	NumWords
1.017091532	NumParagraphs
1.043080069	WordsPerSentence
1.584059258	PercentDeixis
1.664273226	PercentAlphanumeric
1.872464779	PercentPreposition
1.896397545	PercentBracket
2.307047225	PercentDeterminer
2.465059461	PercentExpletives
2.742008487	PercentInterjection
2.881071247	PercentNonDictionary
3.210403009	CharactersPerWord
3.236750434	PercentNumWord
3.282777921	PercentAdverb
3.59411557	PercentPeriod
4.042139963	PercentLetter
4.318338535	PercentModal
4.438239798	PercentVerb
4.557497558	PercentQuestionmark
4.749691102	PercentTagQ
4.982357903	NumSentences
7.055846926	PercentColon

Table 30: Ablation results for the Muir-Joinson speech corpus, hierarchical vs. level classifier, partitioned by sender, with standardised feature scores.

B.2.1 Upwards vs Downwards Classifier, Split by Sender

4.904938749	PercentBracket
4.988227656	NumParagraphs
5.575987796	PercentDeterminer
6.355616561	PercentEmoticon
6.39301373	PercentInterjection
6.459406168	PercentLetter
6.545937893	WordsPerSentence
7.279744221	PercentPeriod
8.397261071	PercentNoun
8.576934892	PercentAmpersand
8.927286763	NumWords
9.078224237	PercentRptLetterWord
9.240133035	PercentHyphen
9.476565806	PercentExpletives
9.517275138	PercentAdverb
9.537085212	PercentComma
11.07478795	PercentQuestionmark
11.18597707	PercentVerb
11.28843918	PercentExclamation
11.41816371	PercentPreposition
11.49029076	PercentHedges
11.55677551	PercentPronoun
11.80742103	PercentConjunction
12.28579564	PercentDeixis
12.31745134	CharactersPerWord
12.7908495	PercentUppercase
12.89732666	NumSentences
13.2439905	PercentNumWord
13.28501804	PercentSemicolon
13.63480079	NumChars
13.65832047	PercentModal
13.66524324	PercentContractions
14.24477092	PercentAdjective
14.81814372	PercentNonDictionary
15.09363199	PercentColon
15.19213727	HeylighenDewaele
15.61769391	PercentAlphanumeric
15.91586634	PercentPolite
16.18897923	PercentTagQ
16.37029556	PercentQuote

Table 31: Ablation results for the Muir-Joinson speech corpus, upwards vs. downwards classifier, partitioned by sender, with standardised feature scores.

B.3 Feature Ablation – Muir-Joinson CMC Corpus

B.3.1 Hierarchical vs Level Classifier, Split by Pair

-0.067818426	PercentComma
-0.064420212	PercentLetter
-0.019821307	PercentHyphen
-0.018420132	PercentHedges
-0.013767956	NumWords
-0.011243222	PercentVerb
-0.000261433	PercentExpletives
0.002584095	PercentNonDictionary
0.010752369	PercentInterjection
0.012132198	PercentRptLetterWord
0.014274645	HeylighenDewaele
0.016672171	PercentDeixis
0.018840951	NumSentences
0.023641681	PercentEmoticon
0.024051493	PercentQuestionmark
0.027392703	PercentContractions
0.040196145	PercentAdjective
0.04042489	PercentPolite
0.041394908	PercentAmpersand
0.042545991	PercentUppercase
0.0437661	PercentNumWord
0.044533705	PercentConjunction
0.04622815	PercentColon
0.047789045	PercentPreposition
0.04948349	PercentBracket
0.061381438	PercentExclamation
0.061381438	WordsPerSentence
0.063075883	CharactersPerWord
0.06334282	PercentSemicolon
0.075202577	NumChars
0.075288447	PercentDeterminer
0.075698258	PercentModal
0.076146262	PercentPeriod
0.078724853	PercentAlphanumeric
0.08191133	PercentTagQ
0.083558096	NumParagraphs
0.089290652	PercentPronoun
0.093041673	PercentQuote
0.109298529	PercentAdverb
0.161915312	PercentNoun

Table 32: Ablation results for the Muir-Joinson CMC corpus, hierarchical vs. level classifier, partitioned by pair, with standardised feature scores.

B.3.2 Upwards vs Downwards Classifier, Split by Pair

-2.121370899	PercentUppercase
-1.94643594	PercentRptLetterWord
-1.85723798	WordsPerSentence
-1.812026778	PercentContractions
-1.777091166	PercentTagQ
-1.477494125	PercentNumWord
-1.419236799	PercentModal
-1.337869915	PercentPolite
-1.332094323	NumParagraphs
-1.315870399	PercentBracket
-1.256928109	PercentAdverb
-1.251986889	CharactersPerWord
-1.236739793	PercentSemicolon
-1.214128431	PercentAmpersand
-1.050095146	PercentExclamation
-1.030878934	PercentAdjective
-0.995667041	PercentQuestionmark
-0.973430878	PercentHyphen
-0.951883248	PercentQuote
-0.799656254	PercentNoun
-0.780888694	PercentHedges
-0.726255423	PercentVerb
-0.724207271	PercentLetter
-0.714253312	PercentAlphanumeric
-0.657514764	NumChars
-0.529442548	PercentNonDictionary
-0.398507223	PercentPronoun
-0.338273105	PercentConjunction
-0.099912148	NumWords
-0.082865077	NumSentences
-0.029981946	PercentInterjection
0.069320736	PercentPreposition
0.086119695	PercentEmoticon
0.091811207	HeylighenDewaele
0.11099066	PercentExpletives
0.208940114	PercentColon
0.479154829	PercentComma
0.799827058	PercentPeriod
1.271341169	PercentDeixis
1.534233304	PercentDeterminer

Table 33: Ablation results for the Muir-Joinson CMC corpus, upwards vs. downwards classifier, partitioned by pair, with standardised feature scores.

B.3.3 Hierarchical vs Level Classifier, Split by Sender

-0.389083824	PercentTagQ
-0.387466749	NumChars
-0.367054455	PercentAdjective
-0.354132098	NumParagraphs
-0.352781607	PercentPolite
-0.344271745	PercentAlphanumeric
-0.315650776	PercentPeriod
-0.31307138	PercentHyphen
-0.311075048	PercentQuote
-0.301093108	PercentSemicolon
-0.281354215	WordsPerSentence
-0.249351819	PercentQuestionmark
-0.247760024	PercentPreposition
-0.231891135	PercentNoun
-0.196088075	PercentExpletives
-0.147005938	PercentNonDictionary
-0.031703377	PercentHedges
0.472232805	PercentComma
0.495750619	PercentContractions
0.534957722	PercentColon
0.548642743	PercentAdverb
0.598080543	PercentConjunction
0.599793925	PercentVerb
0.609064171	PercentDeterminer
0.610122799	PercentNumWord
0.612233675	PercentAmpersand
0.619730655	PercentBracket
0.63316505	PercentModal
0.64014295	NumWords
0.662424816	PercentDeixis
0.664510411	HeylighenDewaele
0.66686259	PercentPronoun
0.681695188	PercentInterjection
0.68604598	PercentUppercase
0.692403319	NumSentences
0.697369498	PercentRptLetterWord
0.698226189	PercentExclamation
0.705844757	PercentEmoticon
0.761817121	CharactersPerWord
0.776034331	PercentLetter

Table 34: Ablation results for the Muir-Joinson CMC corpus, hierarchical vs. level classifier, partitioned by sender, with standardised feature scores.

B.3.4 Upwards vs Downwards Classifier, Split by Sender

1.403049792	PercentUppercase
2.103095786	PercentAdjective
2.379058292	NumChars
2.382127483	PercentPolite
2.386144115	WordsPerSentence
2.417578249	PercentQuote
2.574856522	PercentLetter
2.616499481	PercentVerb
2.642746659	PercentAlphanumeric
2.646775826	PercentTagQ
2.76375933	PercentHyphen
2.976359532	PercentPeriod
3.026613767	NumWords
3.037138975	PercentNumWord
3.12365607	PercentQuestionmark
3.180907274	CharactersPerWord
3.185922366	PercentConjunction
3.196162966	PercentExclamation
3.288300138	PercentHedges
3.344496451	PercentInterjection
3.397006688	PercentDeterminer
3.40489592	PercentAmpersand
3.515198239	NumParagraphs
3.552828052	PercentPronoun
3.559212565	PercentDeixis
3.609308573	PercentNonDictionary
3.662570519	PercentRptLetterWord
3.671202822	HeylighenDewaele
3.776950214	PercentAdverb
3.790705861	PercentPreposition
3.861944053	PercentSemicolon
4.078857823	PercentComma
4.125798103	NumSentences
4.161396621	PercentEmoticon
4.165941968	PercentModal
4.228599551	PercentBracket
4.235205346	PercentExpletives
4.24484104	PercentNoun
4.412184084	PercentContractions
4.449654646	PercentColon

Table 35: Ablation results for the Muir-Joinson CMC corpus, upwards vs. downwards classifier, partitioned by sender, with standardised feature scores.

Appendix C: Tension Components, Full Output

For completeness, we include here the full output of the tension components, as discussed in section 5.2.

C.1 Tension Component of Enron Message-Level Prediction Graph

rod.hayslett@enron.com jeff.skilling@enron.com -0.220200
rod.hayslett@enron.com shelley.corman@enron.com -0.158560
rod.hayslett@enron.com sally.beck@enron.com -0.356440
rod.hayslett@enron.com danny.mccarty@enron.com -0.158617
rod.hayslett@enron.com stanley.horton@enron.com -0.258215
rod.hayslett@enron.com drew.fossum@enron.com -0.214024
rod.hayslett@enron.com steven.kean@enron.com -0.346046
rod.hayslett@enron.com bill.rapp@enron.com -0.108525
rod.hayslett@enron.com tana.jones@enron.com -0.233880
rod.hayslett@enron.com lindy.donoho@enron.com -0.077653
rod.hayslett@enron.com tracy.geaccone@enron.com -0.476131
jeff.skilling@enron.com greg.whalley@enron.com -0.047244
jeff.skilling@enron.com vince.kaminski@enron.com -0.289789
jeff.skilling@enron.com sally.beck@enron.com -0.136240
jeff.skilling@enron.com stanley.horton@enron.com -0.038015
jeff.skilling@enron.com steven.kean@enron.com -0.125846
jeff.skilling@enron.com jeff.dasovich@enron.com -0.154129
jeff.skilling@enron.com david.delainey@enron.com 0.154091
jeff.skilling@enron.com jeffrey.shankman@enron.com 0.054305
jeff.skilling@enron.com a..shankman@enron.com 0.005526
jeff.skilling@enron.com shankman@enron.com -1.000000
mark.taylor@enron.com craig.breslau@enron.com 0.203110
mark.taylor@enron.com kay.mann@enron.com -0.037478
mark.taylor@enron.com harry.arora@enron.com 0.079193
mark.taylor@enron.com stewart.rosman@enron.com 0.127486
mark.taylor@enron.com gerald.nemec@enron.com 0.098868
mark.taylor@enron.com stephanie.panus@enron.com 0.034050
mark.taylor@enron.com louise.kitchen@enron.com 0.051474
mark.taylor@enron.com shelley.corman@enron.com 0.117969
mark.taylor@enron.com michelle.cash@enron.com -0.061112
mark.taylor@enron.com elizabeth.sager@enron.com 0.042936
mark.taylor@enron.com andy.zipper@enron.com 0.052058
mark.taylor@enron.com peter.keavey@enron.com -0.108184
mark.taylor@enron.com phillip.allen@enron.com -0.289444
mark.taylor@enron.com marie.heard@enron.com 0.040028
mark.taylor@enron.com don.black@enron.com 0.092870
mark.taylor@enron.com susan.bailey@enron.com 0.113107
mark.taylor@enron.com tana.jones@enron.com 0.042648
mark.taylor@enron.com mark.haedicke@enron.com 0.134150
mark.taylor@enron.com patrice.mims@enron.com -0.028079
mark.taylor@enron.com richard.sanders@enron.com -0.014435
mark.taylor@enron.com stacy.dickson@enron.com 0.059534
craig.breslau@enron.com louise.kitchen@enron.com -0.151637
craig.breslau@enron.com vince.kaminski@enron.com -0.436571
craig.breslau@enron.com scott.neal@enron.com -0.054105
craig.breslau@enron.com mike.grigsby@enron.com -0.495649
craig.breslau@enron.com errol.mclaughlin@enron.com -0.257099
craig.breslau@enron.com john.arnold@enron.com -0.245690
craig.breslau@enron.com susan.bailey@enron.com -0.090004
craig.breslau@enron.com patrice.mims@enron.com -0.231189
craig.breslau@enron.com eric.bass@enron.com -0.021951
craig.breslau@enron.com dan.hyvl@enron.com -0.146328

kay.mann@enron.com chris.germany@enron.com 0.033439
kay.mann@enron.com m..presto@enron.com -0.078598
kay.mann@enron.com michelle.cash@enron.com -0.023634
kay.mann@enron.com elizabeth.sager@enron.com 0.080414
kay.mann@enron.com laura.luce@enron.com 0.294264
kay.mann@enron.com drew.fossum@enron.com 0.099982
kay.mann@enron.com marie.heard@enron.com 0.077505
kay.mann@enron.com susan.bailey@enron.com 0.150585
kay.mann@enron.com tana.jones@enron.com 0.080126
kay.mann@enron.com jake.thomas@enron.com 0.013992
kay.mann@enron.com richard.sanders@enron.com 0.023043
kay.mann@enron.com john.llodra@enron.com 0.073742
harry.arora@enron.com robert.stalford@enron.com -0.172727
harry.arora@enron.com doug.gilbert-smith@enron.com 0.219299
harry.arora@enron.com stephanie.panus@enron.com -0.045143
harry.arora@enron.com louise.kitchen@enron.com -0.027719
harry.arora@enron.com john.lavorato@enron.com 0.150285
harry.arora@enron.com vince.kaminski@enron.com -0.312654
harry.arora@enron.com m..presto@enron.com -0.195269
harry.arora@enron.com benjamin.rogers@enron.com 0.065521
harry.arora@enron.com gautam.gupta@enron.com 0.397756
harry.arora@enron.com andy.zipper@enron.com -0.027134
harry.arora@enron.com jeffrey.shankman@enron.com 0.031441
harry.arora@enron.com kevin.presto@enron.com -0.155742
harry.arora@enron.com j.kaminski@enron.com -0.265389
jane.tholt@enron.com cooper.richey@enron.com 0.150854
jane.tholt@enron.com tori.kuykendall@enron.com -0.150854
cooper.richey@enron.com jonathan.mckay@enron.com -0.066532
cooper.richey@enron.com stewart.rosman@enron.com 0.033200
cooper.richey@enron.com tim.heizenrader@enron.com 0.123301
cooper.richey@enron.com matthew.lenhart@enron.com -0.216088
cooper.richey@enron.com mike.grigsby@enron.com -0.386825
cooper.richey@enron.com john.zufferli@enron.com -0.301704
cooper.richey@enron.com phillip.allen@enron.com -0.383730
cooper.richey@enron.com jeff.richter@enron.com 0.038275
cooper.richey@enron.com randall.gay@enron.com -0.247375
vladi.pimenov@enron.com jonathan.mckay@enron.com 0.154992
vladi.pimenov@enron.com chris.germany@enron.com 0.123200
vladi.pimenov@enron.com john.griffith@enron.com 0.036301
vladi.pimenov@enron.com geoff.storey@enron.com 0.246875
vladi.pimenov@enron.com chris.gaskill@enron.com 0.438632
jonathan.mckay@enron.com chris.germany@enron.com -0.031792
jonathan.mckay@enron.com frank.vickers@enron.com 0.089440
jonathan.mckay@enron.com john.griffith@enron.com -0.118692
jonathan.mckay@enron.com f..brawner@enron.com 0.148547
jonathan.mckay@enron.com john.lavorato@enron.com 0.201724
jonathan.mckay@enron.com scott.neal@enron.com 0.121251
jonathan.mckay@enron.com john.zufferli@enron.com -0.235172
jonathan.mckay@enron.com chris.dorland@enron.com -0.346694
jonathan.mckay@enron.com kam.keiser@enron.com 0.091766
jonathan.mckay@enron.com phillip.allen@enron.com -0.317198
jonathan.mckay@enron.com barry.tycholiz@enron.com 0.091481
jonathan.mckay@enron.com john.arnold@enron.com -0.070334
jonathan.mckay@enron.com mike.maggi@enron.com -0.308301
jonathan.mckay@enron.com brad.mckay@enron.com 0.108199
jonathan.mckay@enron.com hunter.shively@enron.com 0.105738
jonathan.mckay@enron.com lavorato@enron.com 0.453176
jonathan.mckay@enron.com f..calger@enron.com -0.544657
jonathan.mckay@enron.com martin.cuilla@enron.com 0.011549
jonathan.mckay@enron.com steve.wang@enron.com -0.089532

diana.scholtes@enron.com phillip.platter@enron.com 0.642930
diana.scholtes@enron.com holden.salisbury@enron.com 0.349176
diana.scholtes@enron.com paul.choi@enron.com 0.641746
diana.scholtes@enron.com sean.crandall@enron.com 0.607411
diana.scholtes@enron.com stewart.rosman@enron.com 0.493065
diana.scholtes@enron.com tim.heizenrader@enron.com 0.583166
diana.scholtes@enron.com cara.semperger@enron.com 0.456177
diana.scholtes@enron.com mike.swerzbin@enron.com 0.481982
diana.scholtes@enron.com tim.belden@enron.com 0.364655
diana.scholtes@enron.com tom.alonso@enron.com 0.640427
diana.scholtes@enron.com jeff.richter@enron.com 0.498139
diana.scholtes@enron.com kate.symes@enron.com 0.431835
diana.scholtes@enron.com mark.fischer@enron.com 0.451261
diana.scholtes@enron.com lloyd.will@enron.com 0.420962
diana.scholtes@enron.com ken.lay@enron.com -0.259009
phillip.platter@enron.com paul.choi@enron.com -0.001184
phillip.platter@enron.com jeff.richter@enron.com -0.144791
phillip.platter@enron.com kate.symes@enron.com -0.211095
holden.salisbury@enron.com tim.heizenrader@enron.com 0.233990
holden.salisbury@enron.com cara.semperger@enron.com 0.107001
holden.salisbury@enron.com geir.solberg@enron.com 0.293571
holden.salisbury@enron.com tim.belden@enron.com 0.015479
holden.salisbury@enron.com tom.alonso@enron.com 0.291251
holden.salisbury@enron.com mark.guzman@enron.com 0.172485
holden.salisbury@enron.com craig.dean@enron.com 0.208834
holden.salisbury@enron.com jeff.richter@enron.com 0.148964
holden.salisbury@enron.com kate.symes@enron.com 0.082660
holden.salisbury@enron.com mark.fischer@enron.com 0.102085
paul.choi@enron.com stewart.rosman@enron.com -0.148681
paul.choi@enron.com kim.ward@enron.com -0.231157
paul.choi@enron.com tom.alonso@enron.com -0.001319
paul.choi@enron.com mark.guzman@enron.com -0.120086
paul.choi@enron.com chris.stokley@enron.com -0.399389
paul.choi@enron.com matt.motley@enron.com -0.125472
sean.crandall@enron.com stewart.rosman@enron.com -0.114347
sean.crandall@enron.com tim.heizenrader@enron.com -0.024246
sean.crandall@enron.com cara.semperger@enron.com -0.151235
sean.crandall@enron.com tim.belden@enron.com -0.242757
sean.crandall@enron.com kate.symes@enron.com -0.175576
sean.crandall@enron.com timothy.heizenrader@enron.com -0.121378
sean.crandall@enron.com jake.thomas@enron.com -0.265319
sean.crandall@enron.com matt.motley@enron.com -0.091138
stewart.rosman@enron.com m..forney@enron.com -0.203274
stewart.rosman@enron.com kim.ward@enron.com -0.082476
stewart.rosman@enron.com elizabeth.sager@enron.com -0.084550
stewart.rosman@enron.com mike.swerzbin@enron.com -0.011083
stewart.rosman@enron.com tim.belden@enron.com -0.128410
stewart.rosman@enron.com susan.bailey@enron.com -0.014380
stewart.rosman@enron.com chris.stokley@enron.com -0.250708
stewart.rosman@enron.com john.forney@enron.com -0.166022
stewart.rosman@enron.com kate.symes@enron.com -0.061230
tim.heizenrader@enron.com matthew.lenhart@enron.com -0.339389
tim.heizenrader@enron.com greg.whalley@enron.com -0.208503
tim.heizenrader@enron.com cara.semperger@enron.com -0.126989
tim.heizenrader@enron.com chris.dorland@enron.com -0.536528
tim.heizenrader@enron.com keith.holst@enron.com -0.329152
tim.heizenrader@enron.com mike.swerzbin@enron.com -0.101184
tim.heizenrader@enron.com phillip.allen@enron.com -0.507031
tim.heizenrader@enron.com tom.alonso@enron.com 0.057261
tim.heizenrader@enron.com hunter.shively@enron.com -0.084095

tim.heizenrader@enron.com robert.badeer@enron.com 0.092046
tim.heizenrader@enron.com matt.motley@enron.com -0.066892
larry.may@enron.com robert.stalford@enron.com -0.118422
larry.may@enron.com john.griffith@enron.com -0.115825
larry.may@enron.com andy.zipper@enron.com 0.027171
larry.may@enron.com peter.keavey@enron.com -0.133071
larry.may@enron.com john.arnold@enron.com -0.067467
larry.may@enron.com mike.maggi@enron.com -0.305434
larry.may@enron.com brad.mckay@enron.com 0.111066
larry.may@enron.com chris.gaskill@enron.com 0.286506
robert.stalford@enron.com joe.stepenovitch@enron.com 0.276284
robert.stalford@enron.com joe.parks@enron.com 0.129225
robert.stalford@enron.com m..presto@enron.com -0.022541
robert.stalford@enron.com sally.beck@enron.com 0.013623
robert.stalford@enron.com andy.zipper@enron.com 0.145593
joe.stepenovitch@enron.com don.baughman@enron.com 0.009167
joe.stepenovitch@enron.com juan.hernandez@enron.com -0.155245
joe.stepenovitch@enron.com lloyd.will@enron.com -0.127366
joe.stepenovitch@enron.com larry.campbell@enron.com 0.049729
joe.parks@enron.com chris.germany@enron.com -0.039729
joe.parks@enron.com gerald.nemec@enron.com 0.063177
joe.parks@enron.com chris.gaskill@enron.com 0.275703
joe.parks@enron.com eric.bass@enron.com 0.145469
joe.parks@enron.com dan.hyvl@enron.com 0.021092
joe.parks@enron.com kevin.ruscitti@enron.com -0.086485
matthew.lenhart@enron.com jason.wolfe@enron.com 0.431441
matthew.lenhart@enron.com mike.grigsby@enron.com -0.170737
matthew.lenhart@enron.com chris.dorland@enron.com -0.197139
matthew.lenhart@enron.com kam.keiser@enron.com 0.241321
matthew.lenhart@enron.com jay.reitmeyer@enron.com 0.099402
matthew.lenhart@enron.com frank.ermis@enron.com 0.668493
matthew.lenhart@enron.com phillip.allen@enron.com -0.167642
matthew.lenhart@enron.com jeff.richter@enron.com 0.254363
matthew.lenhart@enron.com robert.badeer@enron.com 0.431435
matthew.lenhart@enron.com eric.bass@enron.com 0.302961
matthew.lenhart@enron.com tori.kuykendall@enron.com -0.085620
matthew.lenhart@enron.com phillip.love@enron.com 0.102302
matthew.lenhart@enron.com martin.cuilla@enron.com 0.161104
matthew.lenhart@enron.com matt.motley@enron.com 0.272497
jason.wolfe@enron.com mike.grigsby@enron.com -0.602178
jason.wolfe@enron.com keith.holst@enron.com -0.421205
jason.wolfe@enron.com kam.keiser@enron.com -0.190120
jason.wolfe@enron.com frank.ermis@enron.com 0.237052
jason.wolfe@enron.com barry.tycholiz@enron.com -0.190405
jason.wolfe@enron.com h..foster@enron.com 0.222583
jason.wolfe@enron.com eric.bass@enron.com -0.128480
jason.wolfe@enron.com phillip.love@enron.com -0.329139
greg.whalley@enron.com kenneth.lay@enron.com -0.199337
greg.whalley@enron.com louise.kitchen@enron.com 0.042390
greg.whalley@enron.com f..brawner@enron.com 0.167216
greg.whalley@enron.com john.lavorato@enron.com 0.220394
greg.whalley@enron.com vince.kaminski@enron.com -0.242545
greg.whalley@enron.com tim.belden@enron.com -0.010008
greg.whalley@enron.com sally.beck@enron.com -0.088996
greg.whalley@enron.com andy.zipper@enron.com 0.042974
greg.whalley@enron.com stanley.horton@enron.com 0.009229
greg.whalley@enron.com steven.kean@enron.com -0.078602
greg.whalley@enron.com david.delainey@enron.com 0.201335
greg.whalley@enron.com jim.schwieger@enron.com -0.022366
greg.whalley@enron.com jeffrey.shankman@enron.com 0.101550

greg.whalley@enron.com john.arnold@enron.com -0.051664
greg.whalley@enron.com chris.gaskill@enron.com 0.302309
greg.whalley@enron.com mark.haedicke@enron.com 0.125066
greg.whalley@enron.com rick.buy@enron.com -0.053425
greg.whalley@enron.com a..shankman@enron.com 0.052770
greg.whalley@enron.com lavorato@enron.com 0.471846
greg.whalley@enron.com fletcher.sturm@enron.com 0.258643
greg.whalley@enron.com j..kean@enron.com -0.470073
greg.whalley@enron.com tom.donohoe@enron.com -0.023714
greg.whalley@enron.com j.kaminski@enron.com -0.195280
greg.whalley@enron.com j..sturm@enron.com 0.197856
greg.whalley@enron.com richard.sanders@enron.com -0.023519
kenneth.lay@enron.com louise.kitchen@enron.com 0.241727
kenneth.lay@enron.com john.lavorato@enron.com 0.419731
kenneth.lay@enron.com vince.kaminski@enron.com -0.043208
kenneth.lay@enron.com benjamin.rogers@enron.com 0.334967
kenneth.lay@enron.com sally.beck@enron.com 0.110342
kenneth.lay@enron.com stanley.horton@enron.com 0.208567
kenneth.lay@enron.com steven.kean@enron.com 0.120736
kenneth.lay@enron.com richard.shapiro@enron.com 0.271250
kenneth.lay@enron.com david.delainey@enron.com 0.400673
kenneth.lay@enron.com jim.schwieger@enron.com 0.176972
kenneth.lay@enron.com kevin.hyatt@enron.com 0.513022
kenneth.lay@enron.com a..shankman@enron.com 0.252107
kenneth.lay@enron.com j..kean@enron.com -0.270735
kenneth.lay@enron.com tom.donohoe@enron.com 0.175623
chris.germany@enron.com frank.vickers@enron.com 0.121233
chris.germany@enron.com scott.neal@enron.com 0.153044
chris.germany@enron.com kam.keiser@enron.com 0.123558
chris.germany@enron.com judy.townsend@enron.com 0.570751
chris.germany@enron.com chris.gaskill@enron.com 0.315432
chris.germany@enron.com kevin.hyatt@enron.com 0.326807
chris.germany@enron.com john.hodge@enron.com 0.277629
chris.germany@enron.com kevin.ruscitti@enron.com -0.046756
chris.germany@enron.com tom.donohoe@enron.com -0.010592
chris.germany@enron.com martin.cuilla@enron.com 0.043341
chris.germany@enron.com randall.gay@enron.com -0.149051
chris.germany@enron.com susan.pereira@enron.com 0.000000
frank.vickers@enron.com kim.ward@enron.com -0.072183
frank.vickers@enron.com louise.kitchen@enron.com -0.065720
frank.vickers@enron.com scott.neal@enron.com 0.031811
frank.vickers@enron.com laura.luce@enron.com 0.139592
frank.vickers@enron.com jeff.dasovich@enron.com -0.214995
frank.vickers@enron.com marie.heard@enron.com -0.077166
frank.vickers@enron.com barry.tycholiz@enron.com 0.002041
frank.vickers@enron.com john.hodge@enron.com 0.156396
m..forney@enron.com cara.semperger@enron.com 0.166386
m..forney@enron.com doug.gilbert-smith@enron.com 0.374280
m..forney@enron.com tim.belden@enron.com 0.074864
m..forney@enron.com rogers.herndon@enron.com 0.271533
m..forney@enron.com chris.stokley@enron.com -0.047434
m..forney@enron.com robert.badeer@enron.com 0.385421
m..forney@enron.com juan.hernandez@enron.com 0.103292
m..forney@enron.com eric.saibi@enron.com 0.038042
m..forney@enron.com jeff.king@enron.com 0.088061
m..forney@enron.com lloyd.will@enron.com 0.131171
cara.semperger@enron.com geir.solberg@enron.com 0.186570
cara.semperger@enron.com tim.belden@enron.com -0.091522
cara.semperger@enron.com tom.alonso@enron.com 0.184250
cara.semperger@enron.com mark.guzman@enron.com 0.065484

cara.semperger@enron.com craig.dean@enron.com 0.101833
cara.semperger@enron.com jeff.richter@enron.com 0.041963
cara.semperger@enron.com kate.symes@enron.com -0.024341
cara.semperger@enron.com mark.fischer@enron.com -0.004916
doug.gilbert-smith@enron.com m..presto@enron.com -0.414568
doug.gilbert-smith@enron.com michelle.cash@enron.com -0.359604
doug.gilbert-smith@enron.com rogers.herndon@enron.com -0.102746
doug.gilbert-smith@enron.com d..steffes@enron.com 0.001021
doug.gilbert-smith@enron.com juan.hernandez@enron.com -0.270988
doug.gilbert-smith@enron.com mike.curry@enron.com -0.442635
doug.gilbert-smith@enron.com eric.saibi@enron.com -0.336238
doug.gilbert-smith@enron.com jeff.king@enron.com -0.286218
sandra.brawner@enron.com john.griffith@enron.com -0.027312
sandra.brawner@enron.com louise.kitchen@enron.com 0.115100
sandra.brawner@enron.com scott.neal@enron.com 0.212632
sandra.brawner@enron.com brad.mckay@enron.com 0.199580
john.griffith@enron.com scott.neal@enron.com 0.239943
john.griffith@enron.com benjamin.rogers@enron.com 0.235652
john.griffith@enron.com errol.mclaughlin@enron.com 0.036950
john.griffith@enron.com andy.zipper@enron.com 0.142996
john.griffith@enron.com john.arnold@enron.com 0.048358
john.griffith@enron.com mike.maggi@enron.com -0.189609
john.griffith@enron.com brad.mckay@enron.com 0.226891
john.griffith@enron.com daren.farmer@enron.com 0.005513
gerald.nemec@enron.com kim.ward@enron.com -0.053857
gerald.nemec@enron.com stephanie.panus@enron.com -0.064818
gerald.nemec@enron.com jim.schwieger@enron.com -0.112149
gerald.nemec@enron.com mark.whitt@enron.com 0.060524
gerald.nemec@enron.com theresa.staab@enron.com 0.067287
gerald.nemec@enron.com barry.tycholiz@enron.com 0.020367
gerald.nemec@enron.com john.hodge@enron.com 0.174723
gerald.nemec@enron.com paul.lucci@enron.com -0.314194
gerald.nemec@enron.com dan.hyvl@enron.com -0.042085
gerald.nemec@enron.com richard.sanders@enron.com -0.113303
kim.ward@enron.com stephanie.panus@enron.com -0.010961
kim.ward@enron.com chris.foster@enron.com 0.104164
kim.ward@enron.com mike.grigsby@enron.com -0.337550
kim.ward@enron.com jim.schwieger@enron.com -0.058292
kim.ward@enron.com mark.guzman@enron.com 0.111071
kim.ward@enron.com barry.tycholiz@enron.com 0.074224
kim.ward@enron.com john.arnold@enron.com -0.087590
kim.ward@enron.com mike.maggi@enron.com -0.325558
kim.ward@enron.com susan.bailey@enron.com 0.068096
kim.ward@enron.com john.forney@enron.com -0.083546
kim.ward@enron.com kate.symes@enron.com 0.021246
kim.ward@enron.com paul.lucci@enron.com -0.260337
kim.ward@enron.com tori.kuykendall@enron.com -0.252433
kim.ward@enron.com dan.hyvl@enron.com 0.011772
kim.ward@enron.com m..tholt@enron.com 0.293478
stephanie.panus@enron.com elizabeth.sager@enron.com 0.008887
stephanie.panus@enron.com marie.heard@enron.com 0.005978
stephanie.panus@enron.com susan.bailey@enron.com 0.079057
stephanie.panus@enron.com tana.jones@enron.com 0.008599
louise.kitchen@enron.com f..brawner@enron.com 0.124827
louise.kitchen@enron.com john.lavorato@enron.com 0.178004
louise.kitchen@enron.com vince.kaminski@enron.com -0.284935
louise.kitchen@enron.com scott.neal@enron.com 0.097531
louise.kitchen@enron.com shelley.corman@enron.com 0.066495
louise.kitchen@enron.com chris.foster@enron.com 0.097700
louise.kitchen@enron.com mike.grigsby@enron.com -0.344013

louise.kitchen@enron.com m..presto@enron.com -0.167550
louise.kitchen@enron.com john.zufferli@enron.com -0.258892
louise.kitchen@enron.com keith.holst@enron.com -0.163039
louise.kitchen@enron.com michelle.cash@enron.com -0.112586
louise.kitchen@enron.com elizabeth.sager@enron.com -0.008538
louise.kitchen@enron.com tim.belden@enron.com -0.052398
louise.kitchen@enron.com sally.beck@enron.com -0.131386
louise.kitchen@enron.com andy.zipper@enron.com 0.000584
louise.kitchen@enron.com laura.luce@enron.com 0.205312
louise.kitchen@enron.com peter.keavey@enron.com -0.159658
louise.kitchen@enron.com d..steffes@enron.com 0.248039
louise.kitchen@enron.com richard.shapiro@enron.com 0.029523
louise.kitchen@enron.com marie.heard@enron.com -0.011446
louise.kitchen@enron.com don.black@enron.com 0.041397
louise.kitchen@enron.com jeffrey.shankman@enron.com 0.059160
louise.kitchen@enron.com barry.tycholiz@enron.com 0.067761
louise.kitchen@enron.com john.arnold@enron.com -0.094054
louise.kitchen@enron.com john.shelk@enron.com -0.140006
louise.kitchen@enron.com tana.jones@enron.com -0.008825
louise.kitchen@enron.com a..martin@enron.com 0.023489
louise.kitchen@enron.com hunter.shively@enron.com 0.082018
louise.kitchen@enron.com geoff.storey@enron.com 0.068162
louise.kitchen@enron.com chris.gaskill@enron.com 0.259919
louise.kitchen@enron.com mark.haedicke@enron.com 0.082676
louise.kitchen@enron.com h..foster@enron.com 0.480749
louise.kitchen@enron.com rick.buy@enron.com -0.095815
louise.kitchen@enron.com kevin.presto@enron.com -0.128024
louise.kitchen@enron.com a..shankman@enron.com 0.010380
louise.kitchen@enron.com lavorato@enron.com 0.429456
louise.kitchen@enron.com calger@enron.com 0.089002
louise.kitchen@enron.com fletcher.sturm@enron.com 0.216253
louise.kitchen@enron.com j..kean@enron.com -0.512463
louise.kitchen@enron.com j.kaminski@enron.com -0.237670
louise.kitchen@enron.com mike.curry@enron.com -0.195616
louise.kitchen@enron.com james.steffes@enron.com -0.020419
louise.kitchen@enron.com f..calger@enron.com -0.568377
louise.kitchen@enron.com christopher.calger@enron.com -0.090064
louise.kitchen@enron.com john.llodra@enron.com -0.015210
louise.kitchen@enron.com whalley@enron.com 0.089002
f..brawner@enron.com john.lavorato@enron.com 0.053177
f..brawner@enron.com scott.neal@enron.com -0.027295
f..brawner@enron.com rogers.herndon@enron.com 0.019445
f..brawner@enron.com john.arnold@enron.com -0.218880
f..brawner@enron.com mike.maggi@enron.com -0.456848
f..brawner@enron.com tom.donohoe@enron.com -0.190931
f..brawner@enron.com steve.wang@enron.com -0.238079
john.lavorato@enron.com vince.kaminski@enron.com -0.462939
john.lavorato@enron.com scott.neal@enron.com -0.080473
john.lavorato@enron.com mike.grigsby@enron.com -0.522017
john.lavorato@enron.com m..presto@enron.com -0.345554
john.lavorato@enron.com john.zufferli@enron.com -0.436896
john.lavorato@enron.com chris.dorland@enron.com -0.548418
john.lavorato@enron.com keith.holst@enron.com -0.341043
john.lavorato@enron.com michelle.cash@enron.com -0.290590
john.lavorato@enron.com mike.swerzbin@enron.com -0.113075
john.lavorato@enron.com tim.belden@enron.com -0.230402
john.lavorato@enron.com sally.beck@enron.com -0.309390
john.lavorato@enron.com andy.zipper@enron.com -0.177420
john.lavorato@enron.com rogers.herndon@enron.com -0.033732
john.lavorato@enron.com d..steffes@enron.com 0.070035

john.lavorato@enron.com stanley.horton@enron.com -0.211164
john.lavorato@enron.com steven.kean@enron.com -0.298995
john.lavorato@enron.com jay.reitmeyer@enron.com -0.251878
john.lavorato@enron.com richard.shapiro@enron.com -0.148481
john.lavorato@enron.com david.delainey@enron.com -0.019058
john.lavorato@enron.com phillip.allen@enron.com -0.518922
john.lavorato@enron.com jim.schwieger@enron.com -0.242759
john.lavorato@enron.com don.black@enron.com -0.136607
john.lavorato@enron.com barry.tycholiz@enron.com -0.110243
john.lavorato@enron.com john.arnold@enron.com -0.272058
john.lavorato@enron.com presto@enron.com -0.000000
john.lavorato@enron.com a..martin@enron.com -0.154515
john.lavorato@enron.com chris.gaskill@enron.com 0.081916
john.lavorato@enron.com mark.haedicke@enron.com -0.095328
john.lavorato@enron.com robert.badeer@enron.com 0.080155
john.lavorato@enron.com rick.buy@enron.com -0.273819
john.lavorato@enron.com kevin.presto@enron.com -0.306027
john.lavorato@enron.com a..shankman@enron.com -0.167624
john.lavorato@enron.com lavorato@enron.com 0.251452
john.lavorato@enron.com calger@enron.com -0.089002
john.lavorato@enron.com j..kean@enron.com -0.690466
john.lavorato@enron.com steffes@enron.com -1.000000
john.lavorato@enron.com arora@enron.com -0.000000
john.lavorato@enron.com m..tholt@enron.com 0.109011
john.lavorato@enron.com j.kaminski@enron.com -0.415674
john.lavorato@enron.com j..sturm@enron.com -0.022538
john.lavorato@enron.com f..calger@enron.com -0.746381
john.lavorato@enron.com christopher.calger@enron.com -0.268068
john.lavorato@enron.com matt.motley@enron.com -0.078782
john.lavorato@enron.com s..shively@enron.com -0.018417
john.lavorato@enron.com zufferli@enron.com -0.000000
john.lavorato@enron.com john.j.lavorato@enron.com -1.000000
john.lavorato@enron.com whalley@enron.com -0.089002
john.lavorato@enron.com belden@enron.com 0.250000
vince.kaminski@enron.com sally.beck@enron.com 0.153549
vince.kaminski@enron.com andy.zipper@enron.com 0.285519
vince.kaminski@enron.com laura.luce@enron.com 0.490247
vince.kaminski@enron.com richard.shapiro@enron.com 0.314457
vince.kaminski@enron.com jim.schwieger@enron.com 0.220179
vince.kaminski@enron.com john.arnold@enron.com 0.190881
vince.kaminski@enron.com kimberly.watson@enron.com 0.278043
vince.kaminski@enron.com rick.buy@enron.com 0.189120
vince.kaminski@enron.com tracy.geaccone@enron.com 0.033858
vince.kaminski@enron.com vince.j.kaminski@enron.com 0.222979
vince.kaminski@enron.com j.kaminski@enron.com 0.047265
vince.kaminski@enron.com lloyd.will@enron.com 0.288844
scott.neal@enron.com mike.grigsby@enron.com -0.441544
scott.neal@enron.com kam.keiser@enron.com -0.029486
scott.neal@enron.com don.black@enron.com -0.056135
scott.neal@enron.com john.arnold@enron.com -0.191585
scott.neal@enron.com brad.mckay@enron.com -0.013052
scott.neal@enron.com julia.sudduth@enron.com -0.000000
scott.neal@enron.com a..martin@enron.com -0.074042
scott.neal@enron.com hunter.shively@enron.com -0.015513
scott.neal@enron.com chris.gaskill@enron.com 0.162388
scott.neal@enron.com john.hodge@enron.com 0.124585
scott.neal@enron.com s..shively@enron.com 0.062056
shelley.corman@enron.com d..steffes@enron.com 0.181544
shelley.corman@enron.com danny.mccarty@enron.com -0.000058
shelley.corman@enron.com stanley.horton@enron.com -0.099655

shelley.corman@enron.com drew.fossum@enron.com -0.055465
shelley.corman@enron.com steven.kean@enron.com -0.187486
shelley.corman@enron.com jeff.dasovich@enron.com -0.215769
shelley.corman@enron.com lindy.donoho@enron.com 0.080906
shelley.corman@enron.com kimberly.watson@enron.com -0.073386
shelley.corman@enron.com darrell.schoolcraft@enron.com -0.374221
shelley.corman@enron.com kevin.hyatt@enron.com 0.204800
shelley.corman@enron.com teb.lokey@enron.com 0.051586
shelley.corman@enron.com tracy.geaccone@enron.com -0.317571
shelley.corman@enron.com lynn.blair@enron.com -0.083424
chris.foster@enron.com jeff.dasovich@enron.com -0.246975
chris.foster@enron.com kate.symes@enron.com -0.082917
chris.foster@enron.com lavorato@enron.com 0.331756
mike.grigsby@enron.com john.zufferli@enron.com 0.085121
mike.grigsby@enron.com chris.dorland@enron.com -0.026402
mike.grigsby@enron.com keith.holst@enron.com 0.180974
mike.grigsby@enron.com kam.keiser@enron.com 0.412058
mike.grigsby@enron.com mike.swerzbin@enron.com 0.408942
mike.grigsby@enron.com tim.belden@enron.com 0.291615
mike.grigsby@enron.com jay.reitmeyer@enron.com 0.270139
mike.grigsby@enron.com frank.ermis@enron.com 0.839230
mike.grigsby@enron.com phillip.allen@enron.com 0.003095
mike.grigsby@enron.com tom.alonso@enron.com 0.567387
mike.grigsby@enron.com don.black@enron.com 0.385409
mike.grigsby@enron.com mark.whitt@enron.com 0.451931
mike.grigsby@enron.com barry.tycholiz@enron.com 0.411774
mike.grigsby@enron.com john.arnold@enron.com 0.249959
mike.grigsby@enron.com chris.gaskill@enron.com 0.603932
mike.grigsby@enron.com h..foster@enron.com 0.824761
mike.grigsby@enron.com robert.badeer@enron.com 0.602172
mike.grigsby@enron.com eric.bass@enron.com 0.473698
mike.grigsby@enron.com tori.kuykendall@enron.com 0.085117
mike.grigsby@enron.com m..tholt@enron.com 0.631027
mike.grigsby@enron.com martin.cuilla@enron.com 0.331841
mike.grigsby@enron.com randall.gay@enron.com 0.139450
mike.grigsby@enron.com mcuilla@enron.com 0.200000
mike.grigsby@enron.com matt.motley@enron.com 0.443234
mike.grigsby@enron.com s..shively@enron.com 0.503600
m..presto@enron.com chris.dorland@enron.com -0.202865
m..presto@enron.com elizabeth.sager@enron.com 0.159012
m..presto@enron.com mike.swerzbin@enron.com 0.232479
m..presto@enron.com tim.belden@enron.com 0.115152
m..presto@enron.com rogers.herndon@enron.com 0.311822
m..presto@enron.com don.black@enron.com 0.208946
m..presto@enron.com john.arnold@enron.com 0.073496
m..presto@enron.com robert.benson@enron.com 0.109296
m..presto@enron.com mike.curry@enron.com -0.028067
m..presto@enron.com lloyd.will@enron.com 0.171459
m..presto@enron.com john.llodra@enron.com 0.152340
john.zufferli@enron.com chris.dorland@enron.com -0.111523
john.zufferli@enron.com tim.belden@enron.com 0.206494
john.zufferli@enron.com david.delainey@enron.com 0.417837
john.zufferli@enron.com a..shankman@enron.com 0.269272
john.zufferli@enron.com lavorato@enron.com 0.688348
john.zufferli@enron.com lloyd.will@enron.com 0.262801
benjamin.rogers@enron.com gautam.gupta@enron.com 0.332235
benjamin.rogers@enron.com rogers.herndon@enron.com 0.051032
benjamin.rogers@enron.com stanley.horton@enron.com -0.126400
benjamin.rogers@enron.com david.delainey@enron.com 0.065705
benjamin.rogers@enron.com robert.benson@enron.com -0.151494

benjamin.rogers@enron.com joe.quenet@enron.com -0.118271
gautam.gupta@enron.com j..sturm@enron.com -0.270009
gautam.gupta@enron.com paul.thomas@enron.com -1.000000
chris.dorland@enron.com keith.holst@enron.com 0.207375
chris.dorland@enron.com barry.tycholiz@enron.com 0.438175
chris.dorland@enron.com geoff.storey@enron.com 0.438576
chris.dorland@enron.com chris.gaskill@enron.com 0.630334
chris.dorland@enron.com robert.benson@enron.com 0.312161
chris.dorland@enron.com lavorato@enron.com 0.799871
chris.dorland@enron.com joe.quenet@enron.com 0.345384
chris.dorland@enron.com mike.carson@enron.com 0.223204
chris.dorland@enron.com j..sturm@enron.com 0.525881
chris.dorland@enron.com martin.cuilla@enron.com 0.358243
keith.holst@enron.com kam.keiser@enron.com 0.231085
keith.holst@enron.com mike.swerzbin@enron.com 0.227968
keith.holst@enron.com jay.reitmeyer@enron.com 0.089165
keith.holst@enron.com jeff.dasovich@enron.com 0.013765
keith.holst@enron.com phillip.allen@enron.com -0.177879
keith.holst@enron.com barry.tycholiz@enron.com 0.230800
keith.holst@enron.com jeff.richter@enron.com 0.244126
keith.holst@enron.com geoff.storey@enron.com 0.231201
keith.holst@enron.com m..tholt@enron.com 0.450054
keith.holst@enron.com j.kaminski@enron.com -0.074631
ryan.slinger@enron.com geir.solberg@enron.com -0.141193
ryan.slinger@enron.com mark.guzman@enron.com -0.262280
ryan.slinger@enron.com craig.dean@enron.com -0.225930
ryan.slinger@enron.com eric.linder@enron.com -0.570597
geir.solberg@enron.com tim.belden@enron.com -0.278092
geir.solberg@enron.com mark.guzman@enron.com -0.121086
geir.solberg@enron.com craig.dean@enron.com -0.084737
geir.solberg@enron.com chris.stokley@enron.com -0.400390
geir.solberg@enron.com h..foster@enron.com 0.255054
geir.solberg@enron.com kate.symes@enron.com -0.210911
geir.solberg@enron.com mark.fischer@enron.com -0.191486
geir.solberg@enron.com eric.linder@enron.com -0.429403
b..sanders@enron.com michelle.cash@enron.com -0.160327
b..sanders@enron.com elizabeth.sager@enron.com -0.056279
b..sanders@enron.com richard.shapiro@enron.com -0.018219
b..sanders@enron.com jeff.dasovich@enron.com -0.197016
b..sanders@enron.com james.steffes@enron.com -0.068160
michelle.cash@enron.com elizabeth.sager@enron.com 0.104048
michelle.cash@enron.com sally.beck@enron.com -0.018800
michelle.cash@enron.com rogers.herndon@enron.com 0.256858
michelle.cash@enron.com danny.mccarty@enron.com 0.179023
michelle.cash@enron.com steven.kean@enron.com -0.008406
michelle.cash@enron.com geoff.storey@enron.com 0.180748
michelle.cash@enron.com mark.haedicke@enron.com 0.195262
michelle.cash@enron.com j..kean@enron.com -0.399877
michelle.cash@enron.com j.kaminski@enron.com -0.125084
michelle.cash@enron.com richard.sanders@enron.com 0.046677
michelle.cash@enron.com s..shively@enron.com 0.272173
elizabeth.sager@enron.com tim.belden@enron.com -0.043860
elizabeth.sager@enron.com sally.beck@enron.com -0.122848
elizabeth.sager@enron.com rogers.herndon@enron.com 0.152810
elizabeth.sager@enron.com marie.heard@enron.com -0.002908
elizabeth.sager@enron.com tana.jones@enron.com -0.000288
elizabeth.sager@enron.com john.forney@enron.com -0.081472
elizabeth.sager@enron.com mark.haedicke@enron.com 0.091214
elizabeth.sager@enron.com richard.sanders@enron.com -0.057371
elizabeth.sager@enron.com f..calger@enron.com -0.559840

elizabeth.sager@enron.com lloyd.will@enron.com 0.012447
elizabeth.sager@enron.com john.llodra@enron.com -0.006672
errol.mclaughlin@enron.com kam.keiser@enron.com 0.173508
errol.mclaughlin@enron.com andy.zipper@enron.com 0.106046
errol.mclaughlin@enron.com john.arnold@enron.com 0.011408
kam.keiser@enron.com jay.reitmeyer@enron.com -0.141920
kam.keiser@enron.com tom.alonso@enron.com 0.155329
kam.keiser@enron.com mark.whitt@enron.com 0.039872
kam.keiser@enron.com hunter.shively@enron.com 0.013972
kam.keiser@enron.com kevin.ruscitti@enron.com -0.170314
kam.keiser@enron.com phillip.love@enron.com -0.139019
mike.swerzbin@enron.com tim.belden@enron.com -0.117327
mike.swerzbin@enron.com tana.jones@enron.com -0.073755
mike.swerzbin@enron.com matt.motley@enron.com 0.034292
mike.swerzbin@enron.com ken.lay@enron.com -0.740991
tim.belden@enron.com sally.beck@enron.com -0.078988
tim.belden@enron.com d..steffes@enron.com 0.300437
tim.belden@enron.com steven.kean@enron.com -0.068594
tim.belden@enron.com jeff.dasovich@enron.com -0.096877
tim.belden@enron.com david.delainey@enron.com 0.211343
tim.belden@enron.com phillip.allen@enron.com -0.288520
tim.belden@enron.com tom.alonso@enron.com 0.275772
tim.belden@enron.com mark.guzman@enron.com 0.157006
tim.belden@enron.com chris.stokley@enron.com -0.122298
tim.belden@enron.com jeff.richter@enron.com 0.133485
tim.belden@enron.com john.forney@enron.com -0.037611
tim.belden@enron.com mark.haedicke@enron.com 0.135074
tim.belden@enron.com h..foster@enron.com 0.533146
tim.belden@enron.com robert.badeer@enron.com 0.310557
tim.belden@enron.com kate.symes@enron.com 0.067181
tim.belden@enron.com mark.fischer@enron.com 0.086606
tim.belden@enron.com timothy.heizenrader@enron.com 0.121378
tim.belden@enron.com kevin.presto@enron.com -0.075626
tim.belden@enron.com lavorato@enron.com 0.481854
tim.belden@enron.com richard.sanders@enron.com -0.013511
tim.belden@enron.com f..calger@enron.com -0.515980
tim.belden@enron.com matt.motley@enron.com 0.151619
sally.beck@enron.com andy.zipper@enron.com 0.131970
sally.beck@enron.com laura.luce@enron.com 0.336698
sally.beck@enron.com stanley.horton@enron.com 0.098225
sally.beck@enron.com david.delainey@enron.com 0.290331
sally.beck@enron.com barry.tycholiz@enron.com 0.199147
sally.beck@enron.com rick.buy@enron.com 0.035571
sally.beck@enron.com fletcher.sturm@enron.com 0.347639
sally.beck@enron.com j..kean@enron.com -0.381077
sally.beck@enron.com j.kaminski@enron.com -0.106284
sally.beck@enron.com daren.farmer@enron.com -0.005513
andy.zipper@enron.com phillip.allen@enron.com -0.341502
andy.zipper@enron.com jim.schwieger@enron.com -0.065340
andy.zipper@enron.com jeffrey.shankman@enron.com 0.058576
andy.zipper@enron.com john.arnold@enron.com -0.094638
andy.zipper@enron.com tana.jones@enron.com -0.009410
andy.zipper@enron.com chris.gaskill@enron.com 0.259335
andy.zipper@enron.com kevin.presto@enron.com -0.128608
andy.zipper@enron.com a..shankman@enron.com 0.009796
andy.zipper@enron.com lavorato@enron.com 0.428872
andy.zipper@enron.com j..sturm@enron.com 0.154882
andy.zipper@enron.com kaminski@enron.com 0.000000
andy.zipper@enron.com steve.wang@enron.com -0.113836
laura.luce@enron.com david.delainey@enron.com -0.046367

laura.luce@enron.com barry.tycholiz@enron.com -0.137551
laura.luce@enron.com john.arnold@enron.com -0.299366
laura.luce@enron.com hunter.shively@enron.com -0.123294
laura.luce@enron.com dan.hyvl@enron.com -0.200004
laura.luce@enron.com j.kaminski@enron.com -0.442982
laura.luce@enron.com j..sturm@enron.com -0.049846
laura.luce@enron.com phillip.love@enron.com -0.276286
laura.luce@enron.com s..shively@enron.com -0.045725
peter.keavey@enron.com john.arnold@enron.com 0.065605
peter.keavey@enron.com brad.mckay@enron.com 0.244138
peter.keavey@enron.com eric.bass@enron.com 0.289344
rogers.herndon@enron.com d..steffes@enron.com 0.103767
rogers.herndon@enron.com barry.tycholiz@enron.com -0.076511
rogers.herndon@enron.com chris.stokley@enron.com -0.318968
rogers.herndon@enron.com kevin.presto@enron.com -0.272295
rogers.herndon@enron.com lavorato@enron.com 0.285184
rogers.herndon@enron.com fletcher.sturm@enron.com 0.071981
rogers.herndon@enron.com j..sturm@enron.com 0.011194
rogers.herndon@enron.com richard.ring@enron.com 0.099129
d..steffes@enron.com richard.shapiro@enron.com -0.218516
d..steffes@enron.com jeff.dasovich@enron.com -0.397313
d..steffes@enron.com david.delainey@enron.com -0.089093
d..steffes@enron.com don.black@enron.com -0.206642
d..steffes@enron.com john.shelk@enron.com -0.388045
d..steffes@enron.com richard.ring@enron.com -0.004638
danny.mccarty@enron.com stanley.horton@enron.com -0.099598
danny.mccarty@enron.com drew.fossum@enron.com -0.055407
danny.mccarty@enron.com lindy.donoho@enron.com 0.080964
danny.mccarty@enron.com kimberly.watson@enron.com -0.073329
danny.mccarty@enron.com tracy.geaccone@enron.com -0.317514
danny.mccarty@enron.com lynn.blair@enron.com -0.083366
stanley.horton@enron.com phillip.allen@enron.com -0.307757
stanley.horton@enron.com kevin.hyatt@enron.com 0.304455
stanley.horton@enron.com lavorato@enron.com 0.462616
stanley.horton@enron.com tracy.geaccone@enron.com -0.217916
stanley.horton@enron.com j..kean@enron.com -0.479302
stanley.horton@enron.com corman@enron.com 0.000000
drew.fossum@enron.com bill.rapp@enron.com 0.105500
drew.fossum@enron.com kimberly.watson@enron.com -0.017922
drew.fossum@enron.com kevin.hyatt@enron.com 0.260265
drew.fossum@enron.com teb.lokey@enron.com 0.107051
drew.fossum@enron.com tracy.geaccone@enron.com -0.262106
drew.fossum@enron.com lynn.blair@enron.com -0.027959
steven.kean@enron.com richard.shapiro@enron.com 0.150514
steven.kean@enron.com jeff.dasovich@enron.com -0.028283
steven.kean@enron.com rick.buy@enron.com 0.025177
steven.kean@enron.com lavorato@enron.com 0.550448
steven.kean@enron.com james.steffes@enron.com 0.100573
jay.reitmeyer@enron.com frank.ermis@enron.com 0.569091
jay.reitmeyer@enron.com phillip.allen@enron.com -0.267044
jay.reitmeyer@enron.com mark.whitt@enron.com 0.181792
jay.reitmeyer@enron.com theresa.staab@enron.com 0.188554
jay.reitmeyer@enron.com chris.gaskill@enron.com 0.333793
jay.reitmeyer@enron.com robert.badeer@enron.com 0.332033
jay.reitmeyer@enron.com tori.kuykendall@enron.com -0.185022
jay.reitmeyer@enron.com m..tholt@enron.com 0.360888
jay.reitmeyer@enron.com randall.gay@enron.com -0.130689
jay.reitmeyer@enron.com s..shively@enron.com 0.233461
frank.ermis@enron.com phillip.allen@enron.com -0.836135
richard.shapiro@enron.com jeff.dasovich@enron.com -0.178797

richard.shapiro@enron.com david.delainey@enron.com 0.129423
richard.shapiro@enron.com jeffrey.shankman@enron.com 0.029637
richard.shapiro@enron.com john.shelk@enron.com -0.169529
richard.shapiro@enron.com mark.haedicke@enron.com 0.053153
richard.shapiro@enron.com lavorato@enron.com 0.399934
richard.shapiro@enron.com j..kean@enron.com -0.541985
richard.shapiro@enron.com james.steffes@enron.com -0.049941
jeff.dasovich@enron.com david.delainey@enron.com 0.308220
jeff.dasovich@enron.com phillip.allen@enron.com -0.191643
jeff.dasovich@enron.com barry.tycholiz@enron.com 0.217035
jeff.dasovich@enron.com john.shelk@enron.com 0.009268
jeff.dasovich@enron.com chris.stokley@enron.com -0.025421
jeff.dasovich@enron.com jeff.richter@enron.com 0.230362
jeff.dasovich@enron.com robert.badeer@enron.com 0.407434
jeff.dasovich@enron.com kevin.hyatt@enron.com 0.420569
jeff.dasovich@enron.com j..kean@enron.com -0.363188
jeff.dasovich@enron.com m..tholt@enron.com 0.436289
jeff.dasovich@enron.com james.steffes@enron.com 0.128856
jeff.dasovich@enron.com f..calger@enron.com -0.419103
jeff.dasovich@enron.com christopher.calger@enron.com 0.059210
jeff.dasovich@enron.com lysa.akin@enron.com -0.052086
jeff.dasovich@enron.com jdasovic@enron.com 0.300000
david.delainey@enron.com barry.tycholiz@enron.com -0.091185
david.delainey@enron.com rick.buy@enron.com -0.254760
david.delainey@enron.com lavorato@enron.com 0.270511
david.delainey@enron.com j..kean@enron.com -0.671408
david.delainey@enron.com mike.curry@enron.com -0.354562
david.delainey@enron.com john.llodra@enron.com -0.174156
phillip.allen@enron.com don.black@enron.com 0.382314
phillip.allen@enron.com mark.whitt@enron.com 0.448836
phillip.allen@enron.com barry.tycholiz@enron.com 0.408679
phillip.allen@enron.com jeff.richter@enron.com 0.422005
phillip.allen@enron.com chris.gaskill@enron.com 0.600837
phillip.allen@enron.com robert.badeer@enron.com 0.599077
phillip.allen@enron.com lavorato@enron.com 0.770374
phillip.allen@enron.com tori.kuykendall@enron.com 0.082022
tom.alonso@enron.com robert.badeer@enron.com 0.034784
tom.alonso@enron.com kate.symes@enron.com -0.208592
tom.alonso@enron.com mark.fischer@enron.com -0.189166
marie.heard@enron.com susan.bailey@enron.com 0.073079
marie.heard@enron.com tana.jones@enron.com 0.002621
marie.heard@enron.com richard.sanders@enron.com -0.054463
jim.schwieger@enron.com john.arnold@enron.com -0.029298
jim.schwieger@enron.com a..martin@enron.com 0.088245
jim.schwieger@enron.com eric.bass@enron.com 0.194441
mark.guzman@enron.com john.forney@enron.com -0.194617
mark.guzman@enron.com kate.symes@enron.com -0.089825
mark.guzman@enron.com monika.causholli@enron.com 0.500000
don.black@enron.com john.arnold@enron.com -0.135450
don.black@enron.com a..martin@enron.com -0.017907
don.black@enron.com mark.haedicke@enron.com 0.041280
don.black@enron.com kevin.presto@enron.com -0.169420
don.black@enron.com eric.bass@enron.com 0.088289
jeffrey.shankman@enron.com john.arnold@enron.com -0.153214
jeffrey.shankman@enron.com rick.buy@enron.com -0.154975
mark.whitt@enron.com theresa.staab@enron.com 0.006763
mark.whitt@enron.com barry.tycholiz@enron.com -0.040157
mark.whitt@enron.com jake.thomas@enron.com -0.182877
theresa.staab@enron.com barry.tycholiz@enron.com -0.046920
barry.tycholiz@enron.com h..foster@enron.com 0.412988

barry.tycholiz@enron.com robert.benson@enron.com -0.126014
barry.tycholiz@enron.com lavorato@enron.com 0.361695
barry.tycholiz@enron.com paul.lucci@enron.com -0.334561
barry.tycholiz@enron.com dan.hyvl@enron.com -0.062453
barry.tycholiz@enron.com jake.thomas@enron.com -0.142720
barry.tycholiz@enron.com m..tholt@enron.com 0.219254
barry.tycholiz@enron.com f..calger@enron.com -0.636138
john.arnold@enron.com mike.maggi@enron.com -0.237967
john.arnold@enron.com tana.jones@enron.com 0.085228
john.arnold@enron.com brad.mckay@enron.com 0.178533
john.arnold@enron.com hunter.shively@enron.com 0.176072
john.arnold@enron.com geoff.storey@enron.com 0.162216
john.arnold@enron.com chris.gaskill@enron.com 0.353973
john.arnold@enron.com a..shankman@enron.com 0.104434
john.arnold@enron.com lavorato@enron.com 0.523510
john.arnold@enron.com fletcher.sturm@enron.com 0.310307
john.arnold@enron.com s..shively@enron.com 0.253641
bill.rapp@enron.com lindy.donoho@enron.com 0.030872
bill.rapp@enron.com kimberly.watson@enron.com -0.123421
bill.rapp@enron.com kevin.hyatt@enron.com 0.154765
bill.rapp@enron.com teb.lokey@enron.com 0.001552
bill.rapp@enron.com lynn.blair@enron.com -0.133458
susan.bailey@enron.com tana.jones@enron.com -0.070458
chris.stokley@enron.com jeff.richter@enron.com 0.255783
chris.stokley@enron.com john.forney@enron.com 0.084686
chris.stokley@enron.com h..foster@enron.com 0.655444
chris.stokley@enron.com kate.symes@enron.com 0.189479
jeff.richter@enron.com robert.badeer@enron.com 0.177072
jeff.richter@enron.com kate.symes@enron.com -0.066304
tana.jones@enron.com dan.hyvl@enron.com 0.014134
tana.jones@enron.com stacy.dickson@enron.com 0.016885
brad.mckay@enron.com fletcher.sturm@enron.com 0.131774
brad.mckay@enron.com stacy.dickson@enron.com -0.076419
lindy.donoho@enron.com kimberly.watson@enron.com -0.154293
lindy.donoho@enron.com darrell.schoolcraft@enron.com -0.455127
lindy.donoho@enron.com kevin.hyatt@enron.com 0.123893
lindy.donoho@enron.com teb.lokey@enron.com -0.029320
lindy.donoho@enron.com tracy.geaccone@enron.com -0.398478
lindy.donoho@enron.com lynn.blair@enron.com -0.164330
kimberly.watson@enron.com darrell.schoolcraft@enron.com -0.300834
kimberly.watson@enron.com kevin.hyatt@enron.com 0.278186
kimberly.watson@enron.com teb.lokey@enron.com 0.124973
kimberly.watson@enron.com tracy.geaccone@enron.com -0.244185
kimberly.watson@enron.com lynn.blair@enron.com -0.010037
kimberly.watson@enron.com j.kaminski@enron.com -0.230778
darrell.schoolcraft@enron.com kevin.hyatt@enron.com 0.579020
darrell.schoolcraft@enron.com lynn.blair@enron.com 0.290797
a..martin@enron.com judy.townsend@enron.com 0.491749
a..martin@enron.com chris.gaskill@enron.com 0.236430
a..martin@enron.com eric.bass@enron.com 0.106196
a..martin@enron.com mike.curry@enron.com -0.219106
hunter.shively@enron.com geoff.storey@enron.com -0.013856
hunter.shively@enron.com chris.gaskill@enron.com 0.177901
hunter.shively@enron.com lavorato@enron.com 0.347438
hunter.shively@enron.com kevin.ruscitti@enron.com -0.184286
hunter.shively@enron.com martin.cuilla@enron.com -0.094190
hunter.shively@enron.com lloyd.will@enron.com -0.078109
geoff.storey@enron.com chris.gaskill@enron.com 0.191758
geoff.storey@enron.com kevin.ruscitti@enron.com -0.170430
geoff.storey@enron.com j..sturm@enron.com 0.087304

geoff.storey@enron.com phillip.love@enron.com -0.139135
geoff.storey@enron.com martin.cuilla@enron.com -0.080333
geoff.storey@enron.com s..shively@enron.com 0.091425
chris.gaskill@enron.com eric.bass@enron.com -0.130234
chris.gaskill@enron.com s..shively@enron.com -0.100332
john.forney@enron.com robert.badeer@enron.com 0.348168
john.forney@enron.com don.baughman@enron.com 0.230451
john.forney@enron.com mike.carson@enron.com -0.057202
mark.haedicke@enron.com lavorato@enron.com 0.346780
patrice.mims@enron.com kevin.ruscitti@enron.com -0.022716
patrice.mims@enron.com tom.donohoe@enron.com 0.013448
h..foster@enron.com robert.badeer@enron.com -0.222590
h..foster@enron.com matt.motley@enron.com -0.381527
h..foster@enron.com richard.ring@enron.com -0.237348
robert.badeer@enron.com christopher.calger@enron.com -0.348223
robert.badeer@enron.com matt.motley@enron.com -0.158938
robert.badeer@enron.com lysa.akin@enron.com -0.459520
kate.symes@enron.com mark.fischer@enron.com 0.019425
kate.symes@enron.com jake.thomas@enron.com -0.089742
kate.symes@enron.com matt.motley@enron.com 0.084439
kevin.hyatt@enron.com tracy.geaccone@enron.com -0.522371
kevin.hyatt@enron.com lynn.blair@enron.com -0.288223
kevin.hyatt@enron.com larry.campbell@enron.com -0.090291
kevin.hyatt@enron.com michele.lokay@enron.com -0.000000
teb.lokey@enron.com tracy.geaccone@enron.com -0.369158
robert.benson@enron.com kevin.presto@enron.com -0.069770
robert.benson@enron.com j..sturm@enron.com 0.213720
robert.benson@enron.com dana.davis@enron.com -1.000000
rick.buy@enron.com a..shankman@enron.com 0.106195
rick.buy@enron.com lavorato@enron.com 0.525271
rick.buy@enron.com j..kean@enron.com -0.416648
kevin.presto@enron.com lavorato@enron.com 0.557480
kevin.presto@enron.com lloyd.will@enron.com 0.131933
kevin.presto@enron.com john.llodra@enron.com 0.112813
lavorato@enron.com christopher.calger@enron.com -0.519521
fletcher.sturm@enron.com joe.quenet@enron.com -0.241283
fletcher.sturm@enron.com jeff.king@enron.com -0.255453
don.baughman@enron.com juan.hernandez@enron.com -0.164411
don.baughman@enron.com lloyd.will@enron.com -0.136533
don.baughman@enron.com larry.campbell@enron.com 0.040562
juan.hernandez@enron.com jeff.king@enron.com -0.015231
juan.hernandez@enron.com lloyd.will@enron.com 0.027879
eric.bass@enron.com tori.kuykendall@enron.com -0.388581
eric.bass@enron.com phillip.love@enron.com -0.200659
eric.bass@enron.com martin.cuilla@enron.com -0.141857
eric.bass@enron.com charles.weldon@enron.com -0.500000
tori.kuykendall@enron.com martin.cuilla@enron.com 0.246724
tori.kuykendall@enron.com randall.gay@enron.com 0.054333
dan.hyvl@enron.com kevin.ruscitti@enron.com -0.107577
kevin.ruscitti@enron.com tom.donohoe@enron.com 0.036164
kevin.ruscitti@enron.com mike.curry@enron.com -0.093348
kevin.ruscitti@enron.com martin.cuilla@enron.com 0.090097
kevin.ruscitti@enron.com s..shively@enron.com 0.261855
vince.j.kaminski@enron.com j.kaminski@enron.com -0.175714
joe.quenet@enron.com jeff.king@enron.com -0.014170
eric.saibi@enron.com mike.carson@enron.com -0.057991
eric.saibi@enron.com lloyd.will@enron.com 0.093129
mike.carson@enron.com jeff.king@enron.com 0.108011
j..sturm@enron.com lloyd.will@enron.com -0.151557
richard.sanders@enron.com lysa.akin@enron.com -0.135452

martin.cuilla@enron.com s..shively@enron.com 0.171758
lloyd.will@enron.com f..campbell@enron.com 0.000000
s..shively@enron.com steve.wang@enron.com -0.272839

C.2 Tension Component of Enron Relationship-Level Prediction Graph

rod.hayslett@enron.com jeff.skilling@enron.com -0.032872
rod.hayslett@enron.com shelley.corman@enron.com 0.011543
rod.hayslett@enron.com sally.beck@enron.com -0.231164
rod.hayslett@enron.com danny.mccarty@enron.com -0.019692
rod.hayslett@enron.com stanley.horton@enron.com -0.234905
rod.hayslett@enron.com drew.fossum@enron.com -0.158692
rod.hayslett@enron.com steven.kean@enron.com -0.297873
rod.hayslett@enron.com bill.rapp@enron.com 0.203033
rod.hayslett@enron.com tana.jones@enron.com -0.124977
rod.hayslett@enron.com lindy.donoho@enron.com -0.003744
rod.hayslett@enron.com tracy.geaccone@enron.com -0.110658
jeff.skilling@enron.com greg.whalley@enron.com -0.158053
jeff.skilling@enron.com vince.kaminski@enron.com -0.386035
jeff.skilling@enron.com sally.beck@enron.com -0.198292
jeff.skilling@enron.com stanley.horton@enron.com -0.202033
jeff.skilling@enron.com steven.kean@enron.com -0.265000
jeff.skilling@enron.com jeff.dasovich@enron.com -0.303246
jeff.skilling@enron.com david.delainey@enron.com 0.009840
jeff.skilling@enron.com jeffrey.shankman@enron.com -0.248171
jeff.skilling@enron.com a..shankman@enron.com -0.281882
jeff.skilling@enron.com shankman@enron.com -1.000000
mark.taylor@enron.com craig.breslau@enron.com 0.206475
mark.taylor@enron.com kay.mann@enron.com -0.058520
mark.taylor@enron.com harry.arora@enron.com -0.017224
mark.taylor@enron.com stewart.rosman@enron.com 0.103095
mark.taylor@enron.com gerald.nemec@enron.com 0.305764
mark.taylor@enron.com stephanie.panus@enron.com 0.062613
mark.taylor@enron.com louise.kitchen@enron.com 0.048649
mark.taylor@enron.com shelley.corman@enron.com 0.207834
mark.taylor@enron.com michelle.cash@enron.com -0.117802
mark.taylor@enron.com elizabeth.sager@enron.com 0.014476
mark.taylor@enron.com andy.zipper@enron.com 0.017914
mark.taylor@enron.com peter.keavey@enron.com -0.003780
mark.taylor@enron.com phillip.allen@enron.com -0.402539
mark.taylor@enron.com marie.heard@enron.com 0.045148
mark.taylor@enron.com don.black@enron.com 0.221629
mark.taylor@enron.com susan.bailey@enron.com 0.056839
mark.taylor@enron.com tana.jones@enron.com 0.071314
mark.taylor@enron.com mark.haedicke@enron.com 0.092449
mark.taylor@enron.com patrice.mims@enron.com 0.062430
mark.taylor@enron.com richard.sanders@enron.com 0.001016
mark.taylor@enron.com stacy.dickson@enron.com 0.082218
craig.breslau@enron.com louise.kitchen@enron.com -0.157827
craig.breslau@enron.com vince.kaminski@enron.com -0.429092
craig.breslau@enron.com scott.neal@enron.com -0.109326
craig.breslau@enron.com mike.grigsby@enron.com -0.481370
craig.breslau@enron.com errol.mclaughlin@enron.com -0.280357
craig.breslau@enron.com john.arnold@enron.com -0.053100
craig.breslau@enron.com susan.bailey@enron.com -0.149636
craig.breslau@enron.com patrice.mims@enron.com -0.144045
craig.breslau@enron.com eric.bass@enron.com 0.049357
craig.breslau@enron.com dan.hyvl@enron.com -0.038130

kay.mann@enron.com chris.germany@enron.com 0.190645
kay.mann@enron.com m..presto@enron.com -0.116385
kay.mann@enron.com michelle.cash@enron.com -0.059282
kay.mann@enron.com elizabeth.sager@enron.com 0.072996
kay.mann@enron.com laura.luce@enron.com 0.356395
kay.mann@enron.com drew.fossum@enron.com 0.096119
kay.mann@enron.com marie.heard@enron.com 0.103668
kay.mann@enron.com susan.bailey@enron.com 0.115359
kay.mann@enron.com tana.jones@enron.com 0.129834
kay.mann@enron.com jake.thomas@enron.com -0.069756
kay.mann@enron.com richard.sanders@enron.com 0.059535
kay.mann@enron.com john.llodra@enron.com 0.062351
harry.arora@enron.com robert.stalford@enron.com 0.283396
harry.arora@enron.com doug.gilbert-smith@enron.com 0.213027
harry.arora@enron.com stephanie.panus@enron.com 0.079838
harry.arora@enron.com louise.kitchen@enron.com 0.065873
harry.arora@enron.com john.lavorato@enron.com 0.196444
harry.arora@enron.com vince.kaminski@enron.com -0.205392
harry.arora@enron.com m..presto@enron.com -0.157680
harry.arora@enron.com benjamin.rogers@enron.com 0.212783
harry.arora@enron.com gautam.gupta@enron.com 0.480194
harry.arora@enron.com andy.zipper@enron.com 0.035139
harry.arora@enron.com jeffrey.shankman@enron.com -0.067528
harry.arora@enron.com kevin.presto@enron.com 0.037254
harry.arora@enron.com j.kaminski@enron.com -0.190573
jane.tholt@enron.com cooper.richey@enron.com 0.162606
jane.tholt@enron.com tori.kuykendall@enron.com -0.162606
cooper.richey@enron.com jonathan.mckay@enron.com 0.045882
cooper.richey@enron.com stewart.rosman@enron.com -0.055173
cooper.richey@enron.com tim.heizenrader@enron.com 0.250250
cooper.richey@enron.com matthew.lenhart@enron.com -0.202711
cooper.richey@enron.com mike.grigsby@enron.com -0.433162
cooper.richey@enron.com john.zufferli@enron.com -0.479097
cooper.richey@enron.com phillip.allen@enron.com -0.560807
cooper.richey@enron.com jeff.richter@enron.com -0.189226
cooper.richey@enron.com randall.gay@enron.com -0.213350
vladi.pimenov@enron.com jonathan.mckay@enron.com 0.226426
vladi.pimenov@enron.com chris.germany@enron.com 0.154401
vladi.pimenov@enron.com john.griffith@enron.com 0.062918
vladi.pimenov@enron.com geoff.storey@enron.com 0.233572
vladi.pimenov@enron.com chris.gaskill@enron.com 0.322684
jonathan.mckay@enron.com chris.germany@enron.com -0.072025
jonathan.mckay@enron.com frank.vickers@enron.com 0.005799
jonathan.mckay@enron.com john.griffith@enron.com -0.163508
jonathan.mckay@enron.com f..brawner@enron.com -0.153513
jonathan.mckay@enron.com john.lavorato@enron.com -0.024931
jonathan.mckay@enron.com scott.neal@enron.com -0.107000
jonathan.mckay@enron.com john.zufferli@enron.com -0.524979
jonathan.mckay@enron.com chris.dorland@enron.com -0.535293
jonathan.mckay@enron.com kam.keiser@enron.com 0.008035
jonathan.mckay@enron.com phillip.allen@enron.com -0.606689
jonathan.mckay@enron.com barry.tycholiz@enron.com -0.000769
jonathan.mckay@enron.com john.arnold@enron.com -0.050774
jonathan.mckay@enron.com mike.maggi@enron.com -0.401230
jonathan.mckay@enron.com brad.mckay@enron.com -0.028810
jonathan.mckay@enron.com hunter.shively@enron.com -0.017462
jonathan.mckay@enron.com lavorato@enron.com 0.209951
jonathan.mckay@enron.com f..calger@enron.com -0.832721
jonathan.mckay@enron.com martin.cuilla@enron.com -0.081804
jonathan.mckay@enron.com steve.wang@enron.com -0.349968

diana.scholtes@enron.com phillip.platter@enron.com 0.576931
diana.scholtes@enron.com holden.salisbury@enron.com 0.073087
diana.scholtes@enron.com paul.choi@enron.com 0.678558
diana.scholtes@enron.com sean.crandall@enron.com 0.395251
diana.scholtes@enron.com stewart.rosman@enron.com 0.432577
diana.scholtes@enron.com tim.heizenrader@enron.com 0.738000
diana.scholtes@enron.com cara.semperger@enron.com 0.466581
diana.scholtes@enron.com mike.swerzbin@enron.com 0.469211
diana.scholtes@enron.com tim.belden@enron.com 0.419438
diana.scholtes@enron.com tom.alonso@enron.com 0.756204
diana.scholtes@enron.com jeff.richter@enron.com 0.298524
diana.scholtes@enron.com kate.symes@enron.com 0.330640
diana.scholtes@enron.com mark.fischer@enron.com 0.373134
diana.scholtes@enron.com lloyd.will@enron.com 0.257259
diana.scholtes@enron.com ken.lay@enron.com -0.265395
phillip.platter@enron.com paul.choi@enron.com 0.101627
phillip.platter@enron.com jeff.richter@enron.com -0.278406
phillip.platter@enron.com kate.symes@enron.com -0.246290
holden.salisbury@enron.com tim.heizenrader@enron.com 0.664914
holden.salisbury@enron.com cara.semperger@enron.com 0.393494
holden.salisbury@enron.com geir.solberg@enron.com 0.492903
holden.salisbury@enron.com tim.belden@enron.com 0.346351
holden.salisbury@enron.com tom.alonso@enron.com 0.683117
holden.salisbury@enron.com mark.guzman@enron.com 0.348483
holden.salisbury@enron.com craig.dean@enron.com 0.360787
holden.salisbury@enron.com jeff.richter@enron.com 0.225438
holden.salisbury@enron.com kate.symes@enron.com 0.257554
holden.salisbury@enron.com mark.fischer@enron.com 0.300047
paul.choi@enron.com stewart.rosman@enron.com -0.245980
paul.choi@enron.com kim.ward@enron.com -0.324486
paul.choi@enron.com tom.alonso@enron.com 0.077646
paul.choi@enron.com mark.guzman@enron.com -0.256988
paul.choi@enron.com chris.stokley@enron.com -0.352918
paul.choi@enron.com matt.motley@enron.com -0.117089
sean.crandall@enron.com stewart.rosman@enron.com 0.037326
sean.crandall@enron.com tim.heizenrader@enron.com 0.342750
sean.crandall@enron.com cara.semperger@enron.com 0.071330
sean.crandall@enron.com tim.belden@enron.com 0.024187
sean.crandall@enron.com kate.symes@enron.com -0.064610
sean.crandall@enron.com timothy.heizenrader@enron.com 0.012094
sean.crandall@enron.com jake.thomas@enron.com -0.194044
sean.crandall@enron.com matt.motley@enron.com 0.166218
stewart.rosman@enron.com m..forney@enron.com -0.172742
stewart.rosman@enron.com kim.ward@enron.com -0.078505
stewart.rosman@enron.com elizabeth.sager@enron.com -0.088618
stewart.rosman@enron.com mike.swerzbin@enron.com 0.036634
stewart.rosman@enron.com tim.belden@enron.com -0.013139
stewart.rosman@enron.com susan.bailey@enron.com -0.046255
stewart.rosman@enron.com chris.stokley@enron.com -0.106938
stewart.rosman@enron.com john.forney@enron.com -0.156654
stewart.rosman@enron.com kate.symes@enron.com -0.101937
tim.heizenrader@enron.com matthew.lenhart@enron.com -0.452961
tim.heizenrader@enron.com greg.whalley@enron.com -0.403152
tim.heizenrader@enron.com cara.semperger@enron.com -0.271420
tim.heizenrader@enron.com chris.dorland@enron.com -0.739661
tim.heizenrader@enron.com keith.holst@enron.com -0.581435
tim.heizenrader@enron.com mike.swerzbin@enron.com -0.268790
tim.heizenrader@enron.com phillip.allen@enron.com -0.811057
tim.heizenrader@enron.com tom.alonso@enron.com 0.018203
tim.heizenrader@enron.com hunter.shively@enron.com -0.221830

tim.heizenrader@enron.com robert.badeer@enron.com -0.095452
tim.heizenrader@enron.com matt.motley@enron.com -0.176532
larry.may@enron.com robert.stalford@enron.com -0.077953
larry.may@enron.com john.griffith@enron.com -0.303482
larry.may@enron.com andy.zipper@enron.com -0.326209
larry.may@enron.com peter.keavey@enron.com -0.347904
larry.may@enron.com john.arnold@enron.com -0.190748
larry.may@enron.com mike.maggi@enron.com -0.541204
larry.may@enron.com brad.mckay@enron.com -0.168784
larry.may@enron.com chris.gaskill@enron.com -0.043716
robert.stalford@enron.com joe.stepenovitch@enron.com 0.253740
robert.stalford@enron.com joe.parks@enron.com -0.057920
robert.stalford@enron.com m..presto@enron.com -0.441076
robert.stalford@enron.com sally.beck@enron.com -0.301044
robert.stalford@enron.com andy.zipper@enron.com -0.248257
joe.stepenovitch@enron.com don.baughman@enron.com -0.553288
joe.stepenovitch@enron.com juan.hernandez@enron.com -0.252597
joe.stepenovitch@enron.com lloyd.will@enron.com -0.592135
joe.stepenovitch@enron.com larry.campbell@enron.com -0.348239
joe.parks@enron.com chris.germany@enron.com -0.076126
joe.parks@enron.com gerald.nemec@enron.com 0.097513
joe.parks@enron.com chris.gaskill@enron.com 0.092157
joe.parks@enron.com eric.bass@enron.com 0.047581
joe.parks@enron.com dan.hyvl@enron.com -0.039906
joe.parks@enron.com kevin.ruscitti@enron.com -0.179139
matthew.lenhart@enron.com jason.wolfe@enron.com 0.286543
matthew.lenhart@enron.com mike.grigsby@enron.com -0.230452
matthew.lenhart@enron.com chris.dorland@enron.com -0.286700
matthew.lenhart@enron.com kam.keiser@enron.com 0.256627
matthew.lenhart@enron.com jay.reitmeyer@enron.com -0.079521
matthew.lenhart@enron.com frank.ermis@enron.com 0.523695
matthew.lenhart@enron.com phillip.allen@enron.com -0.358097
matthew.lenhart@enron.com jeff.richter@enron.com 0.013485
matthew.lenhart@enron.com robert.badeer@enron.com 0.357509
matthew.lenhart@enron.com eric.bass@enron.com 0.300275
matthew.lenhart@enron.com tori.kuykendall@enron.com -0.122501
matthew.lenhart@enron.com phillip.love@enron.com 0.240250
matthew.lenhart@enron.com martin.cuilla@enron.com 0.166788
matthew.lenhart@enron.com matt.motley@enron.com 0.276429
jason.wolfe@enron.com mike.grigsby@enron.com -0.516994
jason.wolfe@enron.com keith.holst@enron.com -0.415017
jason.wolfe@enron.com kam.keiser@enron.com -0.029915
jason.wolfe@enron.com frank.ermis@enron.com 0.237152
jason.wolfe@enron.com barry.tycholiz@enron.com -0.038719
jason.wolfe@enron.com h..foster@enron.com 0.082597
jason.wolfe@enron.com eric.bass@enron.com 0.013732
jason.wolfe@enron.com phillip.love@enron.com -0.046293
greg.whalley@enron.com kenneth.lay@enron.com -0.200239
greg.whalley@enron.com louise.kitchen@enron.com 0.043283
greg.whalley@enron.com f..brawner@enron.com 0.045271
greg.whalley@enron.com john.lavorato@enron.com 0.173854
greg.whalley@enron.com vince.kaminski@enron.com -0.227982
greg.whalley@enron.com tim.belden@enron.com 0.084590
greg.whalley@enron.com sally.beck@enron.com -0.040239
greg.whalley@enron.com andy.zipper@enron.com 0.012549
greg.whalley@enron.com stanley.horton@enron.com -0.043980
greg.whalley@enron.com steven.kean@enron.com -0.106948
greg.whalley@enron.com david.delainey@enron.com 0.167892
greg.whalley@enron.com jim.schwieger@enron.com 0.434946
greg.whalley@enron.com jeffrey.shankman@enron.com -0.090118

greg.whalley@enron.com john.arnold@enron.com 0.148010
greg.whalley@enron.com chris.gaskill@enron.com 0.295042
greg.whalley@enron.com mark.haedicke@enron.com 0.087083
greg.whalley@enron.com rick.buy@enron.com -0.123776
greg.whalley@enron.com a..shankman@enron.com -0.123829
greg.whalley@enron.com lavorato@enron.com 0.408735
greg.whalley@enron.com fletcher.sturm@enron.com 0.064372
greg.whalley@enron.com j..kean@enron.com -0.566185
greg.whalley@enron.com tom.donohoe@enron.com 0.008767
greg.whalley@enron.com j.kaminski@enron.com -0.213163
greg.whalley@enron.com j..sturm@enron.com 0.205209
greg.whalley@enron.com richard.sanders@enron.com -0.004350
kenneth.lay@enron.com louise.kitchen@enron.com 0.243522
kenneth.lay@enron.com john.lavorato@enron.com 0.374092
kenneth.lay@enron.com vince.kaminski@enron.com -0.027743
kenneth.lay@enron.com benjamin.rogers@enron.com 0.390432
kenneth.lay@enron.com sally.beck@enron.com 0.160000
kenneth.lay@enron.com stanley.horton@enron.com 0.156258
kenneth.lay@enron.com steven.kean@enron.com 0.093291
kenneth.lay@enron.com richard.shapiro@enron.com 0.263769
kenneth.lay@enron.com david.delainey@enron.com 0.368131
kenneth.lay@enron.com jim.schwieger@enron.com 0.635185
kenneth.lay@enron.com kevin.hyatt@enron.com 0.223355
kenneth.lay@enron.com a..shankman@enron.com 0.076410
kenneth.lay@enron.com j..kean@enron.com -0.365946
kenneth.lay@enron.com tom.donohoe@enron.com 0.209006
chris.germany@enron.com frank.vickers@enron.com 0.077824
chris.germany@enron.com scott.neal@enron.com -0.034975
chris.germany@enron.com kam.keiser@enron.com 0.080060
chris.germany@enron.com judy.townsend@enron.com 0.373212
chris.germany@enron.com chris.gaskill@enron.com 0.168283
chris.germany@enron.com kevin.hyatt@enron.com -0.103642
chris.germany@enron.com john.hodge@enron.com 0.054122
chris.germany@enron.com kevin.ruscitti@enron.com -0.103013
chris.germany@enron.com tom.donohoe@enron.com -0.117992
chris.germany@enron.com martin.cuilla@enron.com -0.009779
chris.germany@enron.com randall.gay@enron.com -0.187207
chris.germany@enron.com susan.pereira@enron.com -0.000000
frank.vickers@enron.com kim.ward@enron.com -0.185359
frank.vickers@enron.com louise.kitchen@enron.com -0.161300
frank.vickers@enron.com scott.neal@enron.com -0.112799
frank.vickers@enron.com laura.luce@enron.com 0.087926
frank.vickers@enron.com jeff.dasovich@enron.com -0.349776
frank.vickers@enron.com marie.heard@enron.com -0.164800
frank.vickers@enron.com barry.tycholiz@enron.com -0.006567
frank.vickers@enron.com john.hodge@enron.com -0.023702
m..forney@enron.com cara.semperger@enron.com 0.206746
m..forney@enron.com doug.gilbert-smith@enron.com 0.265450
m..forney@enron.com tim.belden@enron.com 0.159603
m..forney@enron.com rogers.herndon@enron.com 0.216876
m..forney@enron.com chris.stokley@enron.com 0.065804
m..forney@enron.com robert.badeer@enron.com 0.382714
m..forney@enron.com juan.hernandez@enron.com 0.336962
m..forney@enron.com eric.saibi@enron.com 0.062611
m..forney@enron.com jeff.king@enron.com 0.133068
m..forney@enron.com lloyd.will@enron.com -0.002575
cara.semperger@enron.com geir.solberg@enron.com 0.099410
cara.semperger@enron.com tim.belden@enron.com -0.047143
cara.semperger@enron.com tom.alonso@enron.com 0.289623
cara.semperger@enron.com mark.guzman@enron.com -0.045011

cara.semperger@enron.com craig.dean@enron.com -0.032707
cara.semperger@enron.com jeff.richter@enron.com -0.168056
cara.semperger@enron.com kate.symes@enron.com -0.135940
cara.semperger@enron.com mark.fischer@enron.com -0.093446
doug.gilbert-smith@enron.com m..presto@enron.com -0.370707
doug.gilbert-smith@enron.com michelle.cash@enron.com -0.313604
doug.gilbert-smith@enron.com rogers.herndon@enron.com -0.048574
doug.gilbert-smith@enron.com d..steffes@enron.com -0.187535
doug.gilbert-smith@enron.com juan.hernandez@enron.com 0.071512
doug.gilbert-smith@enron.com mike.curry@enron.com -0.337391
doug.gilbert-smith@enron.com eric.saibi@enron.com -0.202839
doug.gilbert-smith@enron.com jeff.king@enron.com -0.132382
sandra.brawner@enron.com john.griffith@enron.com 0.450197
sandra.brawner@enron.com louise.kitchen@enron.com 0.458204
sandra.brawner@enron.com scott.neal@enron.com 0.506705
sandra.brawner@enron.com brad.mckay@enron.com 0.584895
john.griffith@enron.com scott.neal@enron.com 0.056508
john.griffith@enron.com benjamin.rogers@enron.com 0.154917
john.griffith@enron.com errol.mclaughlin@enron.com -0.114524
john.griffith@enron.com andy.zipper@enron.com -0.022728
john.griffith@enron.com john.arnold@enron.com 0.112734
john.griffith@enron.com mike.maggi@enron.com -0.237722
john.griffith@enron.com brad.mckay@enron.com 0.134698
john.griffith@enron.com daren.farmer@enron.com -0.037758
gerald.nemec@enron.com kim.ward@enron.com -0.281175
gerald.nemec@enron.com stephanie.panus@enron.com -0.243151
gerald.nemec@enron.com jim.schwieger@enron.com 0.134548
gerald.nemec@enron.com mark.whitt@enron.com -0.158932
gerald.nemec@enron.com theresa.staab@enron.com 0.077239
gerald.nemec@enron.com barry.tycholiz@enron.com -0.102383
gerald.nemec@enron.com john.hodge@enron.com -0.119517
gerald.nemec@enron.com paul.lucci@enron.com -0.461186
gerald.nemec@enron.com dan.hyvl@enron.com -0.137419
gerald.nemec@enron.com richard.sanders@enron.com -0.304749
kim.ward@enron.com stephanie.panus@enron.com 0.038024
kim.ward@enron.com chris.foster@enron.com 0.045144
kim.ward@enron.com mike.grigsby@enron.com -0.299484
kim.ward@enron.com jim.schwieger@enron.com 0.415723
kim.ward@enron.com mark.guzman@enron.com 0.067498
kim.ward@enron.com barry.tycholiz@enron.com 0.178792
kim.ward@enron.com john.arnold@enron.com 0.128786
kim.ward@enron.com mike.maggi@enron.com -0.221670
kim.ward@enron.com susan.bailey@enron.com 0.032250
kim.ward@enron.com john.forney@enron.com -0.078148
kim.ward@enron.com kate.symes@enron.com -0.023431
kim.ward@enron.com paul.lucci@enron.com -0.180011
kim.ward@enron.com tori.kuykendall@enron.com -0.191533
kim.ward@enron.com dan.hyvl@enron.com 0.143756
kim.ward@enron.com m..tholt@enron.com 0.074780
stephanie.panus@enron.com elizabeth.sager@enron.com -0.048137
stephanie.panus@enron.com marie.heard@enron.com -0.017465
stephanie.panus@enron.com susan.bailey@enron.com -0.005774
stephanie.panus@enron.com tana.jones@enron.com 0.008700
louise.kitchen@enron.com f..brawner@enron.com 0.001988
louise.kitchen@enron.com john.lavorato@enron.com 0.130571
louise.kitchen@enron.com vince.kaminski@enron.com -0.271265
louise.kitchen@enron.com scott.neal@enron.com 0.048501
louise.kitchen@enron.com shelley.corman@enron.com 0.159185
louise.kitchen@enron.com chris.foster@enron.com 0.021085
louise.kitchen@enron.com mike.grigsby@enron.com -0.323543

louise.kitchen@enron.com m..presto@enron.com -0.223554
louise.kitchen@enron.com john.zufferli@enron.com -0.369478
louise.kitchen@enron.com keith.holst@enron.com -0.221566
louise.kitchen@enron.com michelle.cash@enron.com -0.166451
louise.kitchen@enron.com elizabeth.sager@enron.com -0.034173
louise.kitchen@enron.com tim.belden@enron.com 0.041307
louise.kitchen@enron.com sally.beck@enron.com -0.083522
louise.kitchen@enron.com andy.zipper@enron.com -0.030734
louise.kitchen@enron.com laura.luce@enron.com 0.249226
louise.kitchen@enron.com peter.keavey@enron.com -0.052429
louise.kitchen@enron.com d..steffes@enron.com -0.040382
louise.kitchen@enron.com richard.shapiro@enron.com 0.020247
louise.kitchen@enron.com marie.heard@enron.com -0.003500
louise.kitchen@enron.com don.black@enron.com 0.172981
louise.kitchen@enron.com jeffrey.shankman@enron.com -0.133401
louise.kitchen@enron.com barry.tycholiz@enron.com 0.154733
louise.kitchen@enron.com john.arnold@enron.com 0.104727
louise.kitchen@enron.com john.shelk@enron.com -0.302153
louise.kitchen@enron.com tana.jones@enron.com 0.022665
louise.kitchen@enron.com a..martin@enron.com -0.170099
louise.kitchen@enron.com hunter.shively@enron.com 0.138039
louise.kitchen@enron.com geoff.storey@enron.com 0.162647
louise.kitchen@enron.com chris.gaskill@enron.com 0.251759
louise.kitchen@enron.com mark.haedicke@enron.com 0.043800
louise.kitchen@enron.com h..foster@enron.com 0.276048
louise.kitchen@enron.com rick.buy@enron.com -0.167059
louise.kitchen@enron.com kevin.presto@enron.com -0.028619
louise.kitchen@enron.com a..shankman@enron.com -0.167112
louise.kitchen@enron.com lavorato@enron.com 0.365452
louise.kitchen@enron.com calger@enron.com 0.065285
louise.kitchen@enron.com fletcher.sturm@enron.com 0.021089
louise.kitchen@enron.com j..kean@enron.com -0.609468
louise.kitchen@enron.com j.kaminski@enron.com -0.256446
louise.kitchen@enron.com mike.curry@enron.com -0.190238
louise.kitchen@enron.com james.steffes@enron.com -0.123381
louise.kitchen@enron.com f..calger@enron.com -0.677220
louise.kitchen@enron.com christopher.calger@enron.com -0.285607
louise.kitchen@enron.com john.llodra@enron.com -0.044817
louise.kitchen@enron.com whalley@enron.com 0.065285
f..brawner@enron.com john.lavorato@enron.com 0.128582
f..brawner@enron.com scott.neal@enron.com 0.046513
f..brawner@enron.com rogers.herndon@enron.com 0.096591
f..brawner@enron.com john.arnold@enron.com 0.102738
f..brawner@enron.com mike.maggi@enron.com -0.247717
f..brawner@enron.com tom.donohoe@enron.com -0.036504
f..brawner@enron.com steve.wang@enron.com -0.196455
john.lavorato@enron.com vince.kaminski@enron.com -0.401836
john.lavorato@enron.com scott.neal@enron.com -0.082070
john.lavorato@enron.com mike.grigsby@enron.com -0.454114
john.lavorato@enron.com m..presto@enron.com -0.354124
john.lavorato@enron.com john.zufferli@enron.com -0.500049
john.lavorato@enron.com chris.dorland@enron.com -0.510362
john.lavorato@enron.com keith.holst@enron.com -0.352137
john.lavorato@enron.com michelle.cash@enron.com -0.297021
john.lavorato@enron.com mike.swerzbin@enron.com -0.039491
john.lavorato@enron.com tim.belden@enron.com -0.089264
john.lavorato@enron.com sally.beck@enron.com -0.214092
john.lavorato@enron.com andy.zipper@enron.com -0.161305
john.lavorato@enron.com rogers.herndon@enron.com -0.031991
john.lavorato@enron.com d..steffes@enron.com -0.170952

john.lavorato@enron.com stanley.horton@enron.com -0.217834
john.lavorato@enron.com steven.kean@enron.com -0.280801
john.lavorato@enron.com jay.reitmeyer@enron.com -0.303183
john.lavorato@enron.com richard.shapiro@enron.com -0.110323
john.lavorato@enron.com david.delainey@enron.com -0.005961
john.lavorato@enron.com phillip.allen@enron.com -0.581759
john.lavorato@enron.com jim.schwieger@enron.com 0.261093
john.lavorato@enron.com don.black@enron.com 0.042410
john.lavorato@enron.com barry.tycholiz@enron.com 0.024162
john.lavorato@enron.com john.arnold@enron.com -0.025844
john.lavorato@enron.com presto@enron.com -0.000000
john.lavorato@enron.com a..martin@enron.com -0.300670
john.lavorato@enron.com chris.gaskill@enron.com 0.121188
john.lavorato@enron.com mark.haedicke@enron.com -0.086770
john.lavorato@enron.com robert.badeer@enron.com 0.133847
john.lavorato@enron.com rick.buy@enron.com -0.297630
john.lavorato@enron.com kevin.presto@enron.com -0.159189
john.lavorato@enron.com a..shankman@enron.com -0.297683
john.lavorato@enron.com lavorato@enron.com 0.234881
john.lavorato@enron.com calger@enron.com -0.065285
john.lavorato@enron.com j..kean@enron.com -0.740039
john.lavorato@enron.com steffes@enron.com -1.000000
john.lavorato@enron.com arora@enron.com 0.000000
john.lavorato@enron.com m..tholt@enron.com -0.079850
john.lavorato@enron.com j.kaminski@enron.com -0.387017
john.lavorato@enron.com j..sturm@enron.com 0.031356
john.lavorato@enron.com f..calger@enron.com -0.807790
john.lavorato@enron.com christopher.calger@enron.com -0.416178
john.lavorato@enron.com matt.motley@enron.com 0.052767
john.lavorato@enron.com s..shively@enron.com -0.035194
john.lavorato@enron.com zufferli@enron.com -0.000000
john.lavorato@enron.com john.j.lavorato@enron.com -1.000000
john.lavorato@enron.com whalley@enron.com -0.065285
john.lavorato@enron.com belden@enron.com 1.000000
vince.kaminski@enron.com sally.beck@enron.com 0.187743
vince.kaminski@enron.com andy.zipper@enron.com 0.240531
vince.kaminski@enron.com laura.luce@enron.com 0.520491
vince.kaminski@enron.com richard.shapiro@enron.com 0.291512
vince.kaminski@enron.com jim.schwieger@enron.com 0.662929
vince.kaminski@enron.com john.arnold@enron.com 0.375992
vince.kaminski@enron.com kimberly.watson@enron.com 0.186378
vince.kaminski@enron.com rick.buy@enron.com 0.104206
vince.kaminski@enron.com tracy.geaccone@enron.com 0.308250
vince.kaminski@enron.com vince.j.kaminski@enron.com 0.007410
vince.kaminski@enron.com j.kaminski@enron.com 0.014819
vince.kaminski@enron.com lloyd.will@enron.com 0.150393
scott.neal@enron.com mike.grigsby@enron.com -0.372044
scott.neal@enron.com kam.keiser@enron.com 0.115035
scott.neal@enron.com don.black@enron.com 0.124480
scott.neal@enron.com john.arnold@enron.com 0.056226
scott.neal@enron.com brad.mckay@enron.com 0.078191
scott.neal@enron.com julia.sudduth@enron.com 0.000000
scott.neal@enron.com a..martin@enron.com -0.218600
scott.neal@enron.com hunter.shively@enron.com 0.089538
scott.neal@enron.com chris.gaskill@enron.com 0.203258
scott.neal@enron.com john.hodge@enron.com 0.089097
scott.neal@enron.com s..shively@enron.com 0.046876
shelley.corman@enron.com d..steffes@enron.com -0.199567
shelley.corman@enron.com danny.mccarty@enron.com -0.031235
shelley.corman@enron.com stanley.horton@enron.com -0.246448

shelley.corman@enron.com drew.fossum@enron.com -0.170235
shelley.corman@enron.com steven.kean@enron.com -0.309416
shelley.corman@enron.com jeff.dasovich@enron.com -0.347662
shelley.corman@enron.com lindy.donoho@enron.com -0.015288
shelley.corman@enron.com kimberly.watson@enron.com -0.244072
shelley.corman@enron.com darrell.schoolcraft@enron.com -0.716882
shelley.corman@enron.com kevin.hyatt@enron.com -0.179352
shelley.corman@enron.com teb.lokey@enron.com 0.106616
shelley.corman@enron.com tracy.geaccone@enron.com -0.122201
shelley.corman@enron.com lynn.blair@enron.com -0.145697
chris.foster@enron.com jeff.dasovich@enron.com -0.209561
chris.foster@enron.com kate.symes@enron.com -0.068576
chris.foster@enron.com lavorato@enron.com 0.344367
mike.grigsby@enron.com john.zufferli@enron.com -0.045935
mike.grigsby@enron.com chris.dorland@enron.com -0.056249
mike.grigsby@enron.com keith.holst@enron.com 0.101977
mike.grigsby@enron.com kam.keiser@enron.com 0.487079
mike.grigsby@enron.com mike.swerzbin@enron.com 0.414623
mike.grigsby@enron.com tim.belden@enron.com 0.364850
mike.grigsby@enron.com jay.reitmeyer@enron.com 0.150930
mike.grigsby@enron.com frank.ermis@enron.com 0.754146
mike.grigsby@enron.com phillip.allen@enron.com -0.127645
mike.grigsby@enron.com tom.alonso@enron.com 0.701616
mike.grigsby@enron.com don.black@enron.com 0.496524
mike.grigsby@enron.com mark.whitt@enron.com 0.421727
mike.grigsby@enron.com barry.tycholiz@enron.com 0.478276
mike.grigsby@enron.com john.arnold@enron.com 0.428270
mike.grigsby@enron.com chris.gaskill@enron.com 0.575302
mike.grigsby@enron.com h..foster@enron.com 0.599592
mike.grigsby@enron.com robert.badeer@enron.com 0.587960
mike.grigsby@enron.com eric.bass@enron.com 0.530727
mike.grigsby@enron.com tori.kuykendall@enron.com 0.107951
mike.grigsby@enron.com m..tholt@enron.com 0.374264
mike.grigsby@enron.com martin.cuilla@enron.com 0.397240
mike.grigsby@enron.com randall.gay@enron.com 0.219813
mike.grigsby@enron.com mcuilla@enron.com 1.000000
mike.grigsby@enron.com matt.motley@enron.com 0.506881
mike.grigsby@enron.com s..shively@enron.com 0.418920
m..presto@enron.com chris.dorland@enron.com -0.156238
m..presto@enron.com elizabeth.sager@enron.com 0.189381
m..presto@enron.com mike.swerzbin@enron.com 0.314633
m..presto@enron.com tim.belden@enron.com 0.264860
m..presto@enron.com rogers.herndon@enron.com 0.322133
m..presto@enron.com don.black@enron.com 0.396534
m..presto@enron.com john.arnold@enron.com 0.328280
m..presto@enron.com robert.benson@enron.com 0.362154
m..presto@enron.com mike.curry@enron.com 0.033316
m..presto@enron.com lloyd.will@enron.com 0.102682
m..presto@enron.com john.llodra@enron.com 0.178736
john.zufferli@enron.com chris.dorland@enron.com -0.010314
john.zufferli@enron.com tim.belden@enron.com 0.410785
john.zufferli@enron.com david.delainey@enron.com 0.494088
john.zufferli@enron.com a..shankman@enron.com 0.202366
john.zufferli@enron.com lavorato@enron.com 0.734930
john.zufferli@enron.com lloyd.will@enron.com 0.248606
benjamin.rogers@enron.com gautam.gupta@enron.com 0.267411
benjamin.rogers@enron.com rogers.herndon@enron.com -0.048331
benjamin.rogers@enron.com stanley.horton@enron.com -0.234173
benjamin.rogers@enron.com david.delainey@enron.com -0.022301
benjamin.rogers@enron.com robert.benson@enron.com -0.008309

benjamin.rogers@enron.com joe.quenet@enron.com -0.196166
gautam.gupta@enron.com j..sturm@enron.com -0.252395
gautam.gupta@enron.com paul.thomas@enron.com -1.000000
chris.dorland@enron.com keith.holst@enron.com 0.158226
chris.dorland@enron.com barry.tycholiz@enron.com 0.534524
chris.dorland@enron.com geoff.storey@enron.com 0.542439
chris.dorland@enron.com chris.gaskill@enron.com 0.631551
chris.dorland@enron.com robert.benson@enron.com 0.518393
chris.dorland@enron.com lavorato@enron.com 0.745244
chris.dorland@enron.com joe.quenet@enron.com 0.330536
chris.dorland@enron.com mike.carson@enron.com 0.249063
chris.dorland@enron.com j..sturm@enron.com 0.541718
chris.dorland@enron.com martin.cuilla@enron.com 0.453489
keith.holst@enron.com kam.keiser@enron.com 0.385102
keith.holst@enron.com mike.swerzbin@enron.com 0.312646
keith.holst@enron.com jay.reitmeyer@enron.com 0.048953
keith.holst@enron.com jeff.dasovich@enron.com 0.033090
keith.holst@enron.com phillip.allen@enron.com -0.229622
keith.holst@enron.com barry.tycholiz@enron.com 0.376299
keith.holst@enron.com jeff.richter@enron.com 0.141959
keith.holst@enron.com geoff.storey@enron.com 0.384213
keith.holst@enron.com m..tholt@enron.com 0.272287
keith.holst@enron.com j.kaminski@enron.com -0.034880
ryan.slinger@enron.com geir.solberg@enron.com -0.063846
ryan.slinger@enron.com mark.guzman@enron.com -0.208267
ryan.slinger@enron.com craig.dean@enron.com -0.195963
ryan.slinger@enron.com eric.linder@enron.com -0.531923
geir.solberg@enron.com tim.belden@enron.com -0.146552
geir.solberg@enron.com mark.guzman@enron.com -0.144421
geir.solberg@enron.com craig.dean@enron.com -0.132117
geir.solberg@enron.com chris.stokley@enron.com -0.240351
geir.solberg@enron.com h..foster@enron.com 0.088190
geir.solberg@enron.com kate.symes@enron.com -0.235350
geir.solberg@enron.com mark.fischer@enron.com -0.192856
geir.solberg@enron.com eric.linder@enron.com -0.468077
b..sanders@enron.com michelle.cash@enron.com 0.131996
b..sanders@enron.com elizabeth.sager@enron.com 0.264274
b..sanders@enron.com richard.shapiro@enron.com 0.318694
b..sanders@enron.com jeff.dasovich@enron.com 0.109970
b..sanders@enron.com james.steffes@enron.com 0.175065
michelle.cash@enron.com elizabeth.sager@enron.com 0.132278
michelle.cash@enron.com sally.beck@enron.com 0.082929
michelle.cash@enron.com rogers.herndon@enron.com 0.265030
michelle.cash@enron.com danny.mccarty@enron.com 0.294400
michelle.cash@enron.com steven.kean@enron.com 0.016220
michelle.cash@enron.com geoff.storey@enron.com 0.329098
michelle.cash@enron.com mark.haedicke@enron.com 0.210251
michelle.cash@enron.com j..kean@enron.com -0.443018
michelle.cash@enron.com j.kaminski@enron.com -0.089995
michelle.cash@enron.com richard.sanders@enron.com 0.118817
michelle.cash@enron.com s..shively@enron.com 0.261827
elizabeth.sager@enron.com tim.belden@enron.com 0.075479
elizabeth.sager@enron.com sally.beck@enron.com -0.049349
elizabeth.sager@enron.com rogers.herndon@enron.com 0.132752
elizabeth.sager@enron.com marie.heard@enron.com 0.030672
elizabeth.sager@enron.com tana.jones@enron.com 0.056838
elizabeth.sager@enron.com john.forney@enron.com -0.068035
elizabeth.sager@enron.com mark.haedicke@enron.com 0.077973
elizabeth.sager@enron.com richard.sanders@enron.com -0.013461
elizabeth.sager@enron.com f..calger@enron.com -0.643047

elizabeth.sager@enron.com lloyd.will@enron.com -0.086699
elizabeth.sager@enron.com john.llodra@enron.com -0.010645
errol.mclaughlin@enron.com kam.keiser@enron.com 0.286066
errol.mclaughlin@enron.com andy.zipper@enron.com 0.091796
errol.mclaughlin@enron.com john.arnold@enron.com 0.227257
kam.keiser@enron.com jay.reitmeyer@enron.com -0.336149
kam.keiser@enron.com tom.alonso@enron.com 0.214537
kam.keiser@enron.com mark.whitt@enron.com -0.065352
kam.keiser@enron.com hunter.shively@enron.com -0.025497
kam.keiser@enron.com kevin.ruscitti@enron.com -0.183072
kam.keiser@enron.com phillip.love@enron.com -0.016377
mike.swerzbin@enron.com tim.belden@enron.com -0.049773
mike.swerzbin@enron.com tana.jones@enron.com -0.068415
mike.swerzbin@enron.com matt.motley@enron.com 0.092258
mike.swerzbin@enron.com ken.lay@enron.com -0.734605
tim.belden@enron.com sally.beck@enron.com -0.124829
tim.belden@enron.com d..steffes@enron.com -0.081688
tim.belden@enron.com steven.kean@enron.com -0.191537
tim.belden@enron.com jeff.dasovich@enron.com -0.229783
tim.belden@enron.com david.delainey@enron.com 0.083303
tim.belden@enron.com phillip.allen@enron.com -0.492495
tim.belden@enron.com tom.alonso@enron.com 0.336766
tim.belden@enron.com mark.guzman@enron.com 0.002132
tim.belden@enron.com chris.stokley@enron.com -0.093799
tim.belden@enron.com jeff.richter@enron.com -0.120913
tim.belden@enron.com john.forney@enron.com -0.143514
tim.belden@enron.com mark.haedicke@enron.com 0.002494
tim.belden@enron.com h..foster@enron.com 0.234742
tim.belden@enron.com robert.badeer@enron.com 0.223111
tim.belden@enron.com kate.symes@enron.com -0.088797
tim.belden@enron.com mark.fischer@enron.com -0.046304
tim.belden@enron.com timothy.heizenrader@enron.com -0.012094
tim.belden@enron.com kevin.presto@enron.com -0.069925
tim.belden@enron.com lavorato@enron.com 0.324145
tim.belden@enron.com richard.sanders@enron.com -0.088940
tim.belden@enron.com f..calger@enron.com -0.718526
tim.belden@enron.com matt.motley@enron.com 0.142031
sally.beck@enron.com andy.zipper@enron.com 0.052788
sally.beck@enron.com laura.luce@enron.com 0.332748
sally.beck@enron.com stanley.horton@enron.com -0.003741
sally.beck@enron.com david.delainey@enron.com 0.208131
sally.beck@enron.com barry.tycholiz@enron.com 0.238255
sally.beck@enron.com rick.buy@enron.com -0.083537
sally.beck@enron.com fletcher.sturm@enron.com 0.104610
sally.beck@enron.com j..kean@enron.com -0.525946
sally.beck@enron.com j.kaminski@enron.com -0.172924
sally.beck@enron.com daren.farmer@enron.com 0.037758
andy.zipper@enron.com phillip.allen@enron.com -0.420454
andy.zipper@enron.com jim.schwieger@enron.com 0.422398
andy.zipper@enron.com jeffrey.shankman@enron.com -0.102666
andy.zipper@enron.com john.arnold@enron.com 0.135461
andy.zipper@enron.com tana.jones@enron.com 0.053399
andy.zipper@enron.com chris.gaskill@enron.com 0.282493
andy.zipper@enron.com kevin.presto@enron.com 0.002116
andy.zipper@enron.com a..shankman@enron.com -0.136378
andy.zipper@enron.com lavorato@enron.com 0.396186
andy.zipper@enron.com j..sturm@enron.com 0.192661
andy.zipper@enron.com kaminski@enron.com 0.000000
andy.zipper@enron.com steve.wang@enron.com -0.163733
laura.luce@enron.com david.delainey@enron.com -0.124617

laura.luce@enron.com barry.tycholiz@enron.com -0.094493
laura.luce@enron.com john.arnold@enron.com -0.144499
laura.luce@enron.com hunter.shively@enron.com -0.111187
laura.luce@enron.com dan.hyvl@enron.com -0.129530
laura.luce@enron.com j.kaminski@enron.com -0.505672
laura.luce@enron.com j..sturm@enron.com -0.087300
laura.luce@enron.com phillip.love@enron.com -0.102067
laura.luce@enron.com s..shively@enron.com -0.153850
peter.keavey@enron.com john.arnold@enron.com 0.157156
peter.keavey@enron.com brad.mckay@enron.com 0.179120
peter.keavey@enron.com eric.bass@enron.com 0.259612
rogers.herndon@enron.com d..steffes@enron.com -0.138961
rogers.herndon@enron.com barry.tycholiz@enron.com 0.056153
rogers.herndon@enron.com chris.stokley@enron.com -0.151072
rogers.herndon@enron.com kevin.presto@enron.com -0.127198
rogers.herndon@enron.com lavorato@enron.com 0.266872
rogers.herndon@enron.com fletcher.sturm@enron.com -0.077491
rogers.herndon@enron.com j..sturm@enron.com 0.063347
rogers.herndon@enron.com richard.ring@enron.com 0.012836
d..steffes@enron.com richard.shapiro@enron.com 0.060629
d..steffes@enron.com jeff.dasovich@enron.com -0.148095
d..steffes@enron.com david.delainey@enron.com 0.164991
d..steffes@enron.com don.black@enron.com 0.213362
d..steffes@enron.com john.shelk@enron.com -0.261771
d..steffes@enron.com richard.ring@enron.com 0.151797
danny.mccarty@enron.com stanley.horton@enron.com -0.215213
danny.mccarty@enron.com drew.fossum@enron.com -0.138999
danny.mccarty@enron.com lindy.donoho@enron.com 0.015948
danny.mccarty@enron.com kimberly.watson@enron.com -0.212837
danny.mccarty@enron.com tracy.geaccone@enron.com -0.090965
danny.mccarty@enron.com lynn.blair@enron.com -0.114461
stanley.horton@enron.com phillip.allen@enron.com -0.363925
stanley.horton@enron.com kevin.hyatt@enron.com 0.067097
stanley.horton@enron.com lavorato@enron.com 0.452715
stanley.horton@enron.com tracy.geaccone@enron.com 0.124248
stanley.horton@enron.com j..kean@enron.com -0.522205
stanley.horton@enron.com corman@enron.com 0.000000
drew.fossum@enron.com bill.rapp@enron.com 0.361725
drew.fossum@enron.com kimberly.watson@enron.com -0.073837
drew.fossum@enron.com kevin.hyatt@enron.com -0.009117
drew.fossum@enron.com teb.lokey@enron.com 0.276851
drew.fossum@enron.com tracy.geaccone@enron.com 0.048034
drew.fossum@enron.com lynn.blair@enron.com 0.024538
steven.kean@enron.com richard.shapiro@enron.com 0.170478
steven.kean@enron.com jeff.dasovich@enron.com -0.038246
steven.kean@enron.com rick.buy@enron.com -0.016828
steven.kean@enron.com lavorato@enron.com 0.515683
steven.kean@enron.com james.steffes@enron.com 0.026849
jay.reitmeyer@enron.com frank.ermis@enron.com 0.603216
jay.reitmeyer@enron.com phillip.allen@enron.com -0.278576
jay.reitmeyer@enron.com mark.whitt@enron.com 0.270796
jay.reitmeyer@enron.com theresa.staab@enron.com 0.506967
jay.reitmeyer@enron.com chris.gaskill@enron.com 0.424372
jay.reitmeyer@enron.com robert.badeer@enron.com 0.437030
jay.reitmeyer@enron.com tori.kuykendall@enron.com -0.042980
jay.reitmeyer@enron.com m..tholt@enron.com 0.223333
jay.reitmeyer@enron.com randall.gay@enron.com 0.068882
jay.reitmeyer@enron.com s..shively@enron.com 0.267989
frank.ermis@enron.com phillip.allen@enron.com -0.881791
richard.shapiro@enron.com jeff.dasovich@enron.com -0.208723

richard.shapiro@enron.com david.delainey@enron.com 0.104362
richard.shapiro@enron.com jeffrey.shankman@enron.com -0.153648
richard.shapiro@enron.com john.shelk@enron.com -0.322400
richard.shapiro@enron.com mark.haedicke@enron.com 0.023553
richard.shapiro@enron.com lavorato@enron.com 0.345205
richard.shapiro@enron.com j..kean@enron.com -0.629715
richard.shapiro@enron.com james.steffes@enron.com -0.143628
jeff.dasovich@enron.com david.delainey@enron.com 0.313086
jeff.dasovich@enron.com phillip.allen@enron.com -0.262712
jeff.dasovich@enron.com barry.tycholiz@enron.com 0.343209
jeff.dasovich@enron.com john.shelk@enron.com -0.113676
jeff.dasovich@enron.com chris.stokley@enron.com 0.135984
jeff.dasovich@enron.com jeff.richter@enron.com 0.108869
jeff.dasovich@enron.com robert.badeer@enron.com 0.452894
jeff.dasovich@enron.com kevin.hyatt@enron.com 0.168310
jeff.dasovich@enron.com j..kean@enron.com -0.420992
jeff.dasovich@enron.com m..tholt@enron.com 0.239197
jeff.dasovich@enron.com james.steffes@enron.com 0.065095
jeff.dasovich@enron.com f..calger@enron.com -0.488743
jeff.dasovich@enron.com christopher.calger@enron.com -0.097131
jeff.dasovich@enron.com lysa.akin@enron.com -0.135421
jeff.dasovich@enron.com jdasovic@enron.com 0.000000
david.delainey@enron.com barry.tycholiz@enron.com 0.030123
david.delainey@enron.com rick.buy@enron.com -0.291668
david.delainey@enron.com lavorato@enron.com 0.240843
david.delainey@enron.com j..kean@enron.com -0.734077
david.delainey@enron.com mike.curry@enron.com -0.314847
david.delainey@enron.com john.llodra@enron.com -0.169427
phillip.allen@enron.com don.black@enron.com 0.624169
phillip.allen@enron.com mark.whitt@enron.com 0.549372
phillip.allen@enron.com barry.tycholiz@enron.com 0.605921
phillip.allen@enron.com jeff.richter@enron.com 0.371581
phillip.allen@enron.com chris.gaskill@enron.com 0.702947
phillip.allen@enron.com robert.badeer@enron.com 0.715606
phillip.allen@enron.com lavorato@enron.com 0.816640
phillip.allen@enron.com tori.kuykendall@enron.com 0.235596
tom.alonso@enron.com robert.badeer@enron.com -0.113655
tom.alonso@enron.com kate.symes@enron.com -0.425563
tom.alonso@enron.com mark.fischer@enron.com -0.383070
marie.heard@enron.com susan.bailey@enron.com 0.011691
marie.heard@enron.com tana.jones@enron.com 0.026165
marie.heard@enron.com richard.sanders@enron.com -0.044133
jim.schwieger@enron.com john.arnold@enron.com -0.286937
jim.schwieger@enron.com a..martin@enron.com -0.561762
jim.schwieger@enron.com eric.bass@enron.com -0.184480
mark.guzman@enron.com john.forney@enron.com -0.145646
mark.guzman@enron.com kate.symes@enron.com -0.090929
mark.guzman@enron.com monika.causholli@enron.com -1.000000
don.black@enron.com john.arnold@enron.com -0.068254
don.black@enron.com a..martin@enron.com -0.343080
don.black@enron.com mark.haedicke@enron.com -0.129180
don.black@enron.com kevin.presto@enron.com -0.201599
don.black@enron.com eric.bass@enron.com 0.034203
jeffrey.shankman@enron.com john.arnold@enron.com 0.238128
jeffrey.shankman@enron.com rick.buy@enron.com -0.033658
mark.whitt@enron.com theresa.staab@enron.com 0.236171
mark.whitt@enron.com barry.tycholiz@enron.com 0.056549
mark.whitt@enron.com jake.thomas@enron.com -0.275108
theresa.staab@enron.com barry.tycholiz@enron.com -0.179622
barry.tycholiz@enron.com h..foster@enron.com 0.121316

barry.tycholiz@enron.com robert.benson@enron.com -0.016132
barry.tycholiz@enron.com lavorato@enron.com 0.210719
barry.tycholiz@enron.com paul.lucci@enron.com -0.358803
barry.tycholiz@enron.com dan.hyvl@enron.com -0.035036
barry.tycholiz@enron.com jake.thomas@enron.com -0.331657
barry.tycholiz@enron.com m..tholt@enron.com -0.104012
barry.tycholiz@enron.com f..calger@enron.com -0.831952
john.arnold@enron.com mike.maggi@enron.com -0.350456
john.arnold@enron.com tana.jones@enron.com -0.082062
john.arnold@enron.com brad.mckay@enron.com 0.021965
john.arnold@enron.com hunter.shively@enron.com 0.033312
john.arnold@enron.com geoff.storey@enron.com 0.057920
john.arnold@enron.com chris.gaskill@enron.com 0.147032
john.arnold@enron.com a..shankman@enron.com -0.271839
john.arnold@enron.com lavorato@enron.com 0.260725
john.arnold@enron.com fletcher.sturm@enron.com -0.083638
john.arnold@enron.com s..shively@enron.com -0.009350
bill.rapp@enron.com lindy.donoho@enron.com -0.206778
bill.rapp@enron.com kimberly.watson@enron.com -0.435562
bill.rapp@enron.com kevin.hyatt@enron.com -0.370842
bill.rapp@enron.com teb.lokey@enron.com -0.084874
bill.rapp@enron.com lynn.blair@enron.com -0.337187
susan.bailey@enron.com tana.jones@enron.com 0.014474
chris.stokley@enron.com jeff.richter@enron.com -0.027115
chris.stokley@enron.com john.forney@enron.com -0.049716
chris.stokley@enron.com h..foster@enron.com 0.328540
chris.stokley@enron.com kate.symes@enron.com 0.005001
jeff.richter@enron.com robert.badeer@enron.com 0.344024
jeff.richter@enron.com kate.symes@enron.com 0.032116
tana.jones@enron.com dan.hyvl@enron.com 0.097032
tana.jones@enron.com stacy.dickson@enron.com 0.010904
brad.mckay@enron.com fletcher.sturm@enron.com -0.105603
brad.mckay@enron.com stacy.dickson@enron.com -0.093122
lindy.donoho@enron.com kimberly.watson@enron.com -0.228784
lindy.donoho@enron.com darrell.schoolcraft@enron.com -0.701594
lindy.donoho@enron.com kevin.hyatt@enron.com -0.164064
lindy.donoho@enron.com teb.lokey@enron.com 0.121903
lindy.donoho@enron.com tracy.geaccone@enron.com -0.106913
lindy.donoho@enron.com lynn.blair@enron.com -0.130409
kimberly.watson@enron.com darrell.schoolcraft@enron.com -0.472810
kimberly.watson@enron.com kevin.hyatt@enron.com 0.064720
kimberly.watson@enron.com teb.lokey@enron.com 0.350688
kimberly.watson@enron.com tracy.geaccone@enron.com 0.121871
kimberly.watson@enron.com lynn.blair@enron.com 0.098375
kimberly.watson@enron.com j.kaminski@enron.com -0.171559
darrell.schoolcraft@enron.com kevin.hyatt@enron.com 0.537530
darrell.schoolcraft@enron.com lynn.blair@enron.com 0.571185
a..martin@enron.com judy.townsend@enron.com 0.626788
a..martin@enron.com chris.gaskill@enron.com 0.421858
a..martin@enron.com eric.bass@enron.com 0.377282
a..martin@enron.com mike.curry@enron.com -0.020139
hunter.shively@enron.com geoff.storey@enron.com 0.024608
hunter.shively@enron.com chris.gaskill@enron.com 0.113720
hunter.shively@enron.com lavorato@enron.com 0.227413
hunter.shively@enron.com kevin.ruscitti@enron.com -0.157576
hunter.shively@enron.com martin.cuilla@enron.com -0.064342
hunter.shively@enron.com lloyd.will@enron.com -0.258911
geoff.storey@enron.com chris.gaskill@enron.com 0.089112
geoff.storey@enron.com kevin.ruscitti@enron.com -0.182184
geoff.storey@enron.com j..sturm@enron.com -0.000721

geoff.storey@enron.com phillip.love@enron.com -0.015488
geoff.storey@enron.com martin.cuilla@enron.com -0.088950
geoff.storey@enron.com s..shively@enron.com -0.067271
chris.gaskill@enron.com eric.bass@enron.com -0.044576
chris.gaskill@enron.com s..shively@enron.com -0.156383
john.forney@enron.com robert.badeer@enron.com 0.366625
john.forney@enron.com don.baughman@enron.com 0.020182
john.forney@enron.com mike.carson@enron.com -0.028521
mark.haedicke@enron.com lavorato@enron.com 0.321652
patrice.mims@enron.com kevin.ruscitti@enron.com -0.033318
patrice.mims@enron.com tom.donohoe@enron.com -0.048297
h..foster@enron.com robert.badeer@enron.com -0.011631
h..foster@enron.com matt.motley@enron.com -0.092711
h..foster@enron.com richard.ring@enron.com -0.164633
robert.badeer@enron.com christopher.calger@enron.com -0.550025
robert.badeer@enron.com matt.motley@enron.com -0.081080
robert.badeer@enron.com lysa.akin@enron.com -0.588315
kate.symes@enron.com mark.fischer@enron.com 0.042494
kate.symes@enron.com jake.thomas@enron.com -0.129434
kate.symes@enron.com matt.motley@enron.com 0.230828
kevin.hyatt@enron.com tracy.geaccone@enron.com 0.057151
kevin.hyatt@enron.com lynn.blair@enron.com 0.033655
kevin.hyatt@enron.com larry.campbell@enron.com 0.143190
kevin.hyatt@enron.com michele.lokay@enron.com 0.000000
teb.lokey@enron.com tracy.geaccone@enron.com -0.228817
robert.benson@enron.com kevin.presto@enron.com -0.167220
robert.benson@enron.com j..sturm@enron.com 0.023326
robert.benson@enron.com dana.davis@enron.com -1.000000
rick.buy@enron.com a..shankman@enron.com -0.000053
rick.buy@enron.com lavorato@enron.com 0.532511
rick.buy@enron.com j..kean@enron.com -0.442409
kevin.presto@enron.com lavorato@enron.com 0.394071
kevin.presto@enron.com lloyd.will@enron.com -0.092253
kevin.presto@enron.com john.llodra@enron.com -0.016199
lavorato@enron.com christopher.calger@enron.com -0.651059
fletcher.sturm@enron.com joe.quenet@enron.com -0.070344
fletcher.sturm@enron.com jeff.king@enron.com -0.006317
don.baughman@enron.com juan.hernandez@enron.com 0.300691
don.baughman@enron.com lloyd.will@enron.com -0.038846
don.baughman@enron.com larry.campbell@enron.com 0.205049
juan.hernandez@enron.com jeff.king@enron.com -0.203894
juan.hernandez@enron.com lloyd.will@enron.com -0.339538
eric.bass@enron.com tori.kuykendall@enron.com -0.422776
eric.bass@enron.com phillip.love@enron.com -0.060025
eric.bass@enron.com martin.cuilla@enron.com -0.133486
eric.bass@enron.com charles.weldon@enron.com -1.000000
tori.kuykendall@enron.com martin.cuilla@enron.com 0.289289
tori.kuykendall@enron.com randall.gay@enron.com 0.111862
dan.hyvl@enron.com kevin.ruscitti@enron.com -0.139233
kevin.ruscitti@enron.com tom.donohoe@enron.com -0.014979
kevin.ruscitti@enron.com mike.curry@enron.com -0.170701
kevin.ruscitti@enron.com martin.cuilla@enron.com 0.093234
kevin.ruscitti@enron.com s..shively@enron.com 0.114913
vince.j.kaminski@enron.com j.kaminski@enron.com 0.007410
joe.quenet@enron.com jeff.king@enron.com 0.064027
eric.saibi@enron.com mike.carson@enron.com -0.075043
eric.saibi@enron.com lloyd.will@enron.com -0.065186
mike.carson@enron.com jeff.king@enron.com 0.145500
j..sturm@enron.com lloyd.will@enron.com -0.282798
richard.sanders@enron.com lysa.akin@enron.com -0.276264

martin.cuilla@enron.com s..shively@enron.com 0.021679
lloyd.will@enron.com f..campbell@enron.com -0.000000
s..shively@enron.com steve.wang@enron.com -0.289844

C.3 Tension Component of Muir-Joinson Speech Message-Level Prediction Graph

buqb40 herd40 0.097105
buqb40 hoql53 0.159300
buqb40 maeg89 -0.337777
buqb40 raus83 -0.003287
buqb40 sabe64 0.014814
buqb40 shbs25 0.440204
buqb40 smtt74 0.008165
buqb40 stjn43 0.020130
buqb40 wadn66 -0.167416
herd40 chip27 -0.072945
herd40 greg44 -0.086135
herd40 guah24 -0.545734
herd40 hahg14 -0.059641
herd40 hous14 -0.086885
herd40 mual23 -0.113812
herd40 roen52 0.015037
hoql53 chip27 -0.135139
hoql53 greg44 -0.148329
hoql53 guah24 -0.607928
hoql53 hahg14 -0.121835
hoql53 hous14 -0.149079
hoql53 mual23 -0.176007
hoql53 roen52 -0.047158
maeg89 chip27 0.361938
maeg89 greg44 0.348748
maeg89 guah24 -0.110851
maeg89 hahg14 0.375242
maeg89 hous14 0.347998
maeg89 mual23 0.321071
maeg89 roen52 0.449920
raus83 chip27 0.027447
raus83 greg44 0.014257
raus83 guah24 -0.445342
raus83 hahg14 0.040751
raus83 hous14 0.013507
raus83 mual23 -0.013420
raus83 roen52 0.115429
sabe64 chip27 0.009347
sabe64 greg44 -0.003843
sabe64 guah24 -0.463442
sabe64 hahg14 0.022650
sabe64 hous14 -0.004593
sabe64 mual23 -0.031521
sabe64 roen52 0.097328
shbs25 chip27 -0.416043
shbs25 greg44 -0.429234
shbs25 guah24 -0.888832
shbs25 hahg14 -0.402740
shbs25 hous14 -0.429983
shbs25 mual23 -0.456911
shbs25 roen52 -0.328062
smtt74 chip27 0.015996
smtt74 greg44 0.002806
smtt74 guah24 -0.456793
smtt74 hahg14 0.029299
smtt74 hous14 0.002056
smtt74 mual23 -0.024872
smtt74 roen52 0.103977
stjn43 chip27 0.004031
stjn43 greg44 -0.009159
stjn43 guah24 -0.468758
stjn43 hahg14 0.017335
stjn43 hous14 -0.009909
stjn43 mual23 -0.036837
stjn43 roen52 0.092012
wadn66 chip27 0.191577
wadn66 greg44 0.178387
wadn66 guah24 -0.281212
wadn66 hahg14 0.204881
wadn66 hous14 0.177637
wadn66 mual23 0.150709
wadn66 roen52 0.279558
barp60 axht13 0.723125
barp60 brbz64 0.603835
barp60 brel45 0.464613
barp60 deeu72 -0.030745
barp60 hazh36 0.900399
barp60 hobs94 0.479459
barp60 hozg91 0.916558
barp60 jadg52 0.893299
barp60 leln17 0.012634
barp60 reeg74 0.267491
barp60 rodh93 0.170850
barp60 whqy96 0.456975
axht13 hojd22 -0.601679
axht13 moub86 -0.720981
axht13 piey32 -0.729284
axht13 sizl24 -0.714360
axht13 arus37 -0.773664
axht13 barp74 -0.753654
axht13 benz72 -0.728548
axht13 bras17 -0.742067
axht13 ches11 -0.777154
axht13 gehw13 -0.766757
axht13 maqh17 -0.314916
brbz64 hojd22 -0.482390
brbz64 moub86 -0.601691
brbz64 piey32 -0.609994
brbz64 sizl24 -0.595070
brbz64 arus37 -0.654374

brbz64 barp74 -0.634364
brbz64 benz72 -0.609258
brbz64 bras17 -0.622777
brbz64 ches11 -0.657864
brbz64 gehw13 -0.647467
brbz64 maqh17 -0.195627
brel45 hojd22 -0.343167
brel45 moub86 -0.462469
brel45 piey32 -0.470772
brel45 sizl24 -0.455848
brel45 arus37 -0.515152
brel45 barp74 -0.495142
brel45 benz72 -0.470036
brel45 bras17 -0.483555
brel45 ches11 -0.518642
brel45 gehw13 -0.508245
brel45 maqh17 -0.056404
deeu72 hojd22 0.152191
deeu72 moub86 0.032889
deeu72 piey32 0.024586
deeu72 sizl24 0.039510
deeu72 arus37 -0.019794
deeu72 barp74 0.000216
deeu72 benz72 0.025322
deeu72 bras17 0.011803
deeu72 ches11 -0.023284
deeu72 gehw13 -0.012887
deeu72 maqh17 0.438954
hazh36 hojd22 -0.778954
hazh36 moub86 -0.898255
hazh36 piey32 -0.906558
hazh36 sizl24 -0.891634
hazh36 arus37 -0.950938
hazh36 barp74 -0.930928
hazh36 benz72 -0.905822
hazh36 bras17 -0.919341
hazh36 ches11 -0.954428
hazh36 gehw13 -0.944031
hazh36 maqh17 -0.492190
hobs94 hojd22 -0.358014
hobs94 moub86 -0.477315
hobs94 piey32 -0.485618
hobs94 sizl24 -0.470694
hobs94 arus37 -0.529998
hobs94 barp74 -0.509988
hobs94 benz72 -0.484882
hobs94 bras17 -0.498401
hobs94 ches11 -0.533488
hobs94 gehw13 -0.523091
hobs94 maqh17 -0.071250
hozg91 hojd22 -0.795113
hozg91 moub86 -0.914414
hozg91 piey32 -0.922717
hozg91 sizl24 -0.907793
hozg91 arus37 -0.967097
hozg91 barp74 -0.947087
hozg91 ches11 -0.970587
hozg91 gehw13 -0.960190
hozg91 maqh17 -0.508349
jadg52 hojd22 -0.771854
jadg52 moub86 -0.891156
jadg52 piey32 -0.899458
jadg52 sizl24 -0.884534
jadg52 arus37 -0.943838
jadg52 barp74 -0.923829
jadg52 benz72 -0.898722
jadg52 bras17 -0.912241
jadg52 ches11 -0.947328
jadg52 gehw13 -0.936931
jadg52 maqh17 -0.485091
leln17 hojd22 0.108811
leln17 moub86 -0.010491
leln17 piey32 -0.018793
leln17 sizl24 -0.003869
leln17 arus37 -0.063174
leln17 barp74 -0.043164
leln17 benz72 -0.018057
leln17 bras17 -0.031577
leln17 ches11 -0.066663
leln17 gehw13 -0.056267
leln17 maqh17 0.395574
reeg74 hojd22 -0.146046
reeg74 moub86 -0.265348
reeg74 piey32 -0.273650
reeg74 sizl24 -0.258726
reeg74 arus37 -0.318031
reeg74 barp74 -0.298021
reeg74 benz72 -0.272914
reeg74 bras17 -0.286434
reeg74 ches11 -0.321520
reeg74 gehw13 -0.311124
reeg74 maqh17 0.140717
rodh93 hojd22 -0.049405
rodh93 moub86 -0.168707
rodh93 piey32 -0.177009
rodh93 sizl24 -0.162085
rodh93 arus37 -0.221389
rodh93 barp74 -0.201380
rodh93 benz72 -0.176273
rodh93 bras17 -0.189793
rodh93 ches11 -0.224879
rodh93 gehw13 -0.214482
rodh93 maqh17 0.237358
whqy96 hojd22 -0.335530
whqy96 moub86 -0.454831
whqy96 piey32 -0.463134
whqy96 sizl24 -0.448210
whqy96 arus37 -0.507514
whqy96 barp74 -0.487504
whqy96 benz72 -0.462398
whqy96 bras17 -0.475917
whqy96 ches11 -0.511004
whqy96 gehw13 -0.500607
whqy96 maqh17 -0.048767

C.4 Tension Component of Muir-Joinson Speech Relationship-Level Prediction Graph

buqb40 herd40 -0.152778
buqb40 hoql53 -0.027778
buqb40 maeg89 -0.027778
buqb40 raus83 -0.027778
buqb40 sabe64 -0.027778
buqb40 shbs25 0.472222
buqb40 smtt74 -0.152778
buqb40 stjn43 -0.027778
buqb40 wadn66 -0.027778
herd40 chip27 0.041667
herd40 greg44 0.152778
herd40 guah24 -0.625000
herd40 hahg14 0.152778
herd40 hous14 0.041667
herd40 mual23 0.041667
herd40 roen52 0.041667
hoql53 chip27 -0.083333
hoql53 greg44 0.027778
hoql53 guah24 -0.750000
hoql53 hahg14 0.027778
hoql53 hous14 -0.083333
hoql53 mual23 -0.083333
hoql53 roen52 -0.083333
maeg89 chip27 -0.083333
maeg89 greg44 0.027778
maeg89 guah24 -0.750000
maeg89 hahg14 0.027778
maeg89 hous14 -0.083333
maeg89 mual23 -0.083333
maeg89 roen52 -0.083333
raus83 chip27 -0.083333
raus83 greg44 0.027778
raus83 guah24 -0.750000
raus83 hahg14 0.027778
raus83 hous14 -0.083333
raus83 mual23 -0.083333
raus83 roen52 -0.083333
sabe64 chip27 -0.083333
sabe64 greg44 0.027778
sabe64 guah24 -0.750000
sabe64 hahg14 0.027778
sabe64 hous14 -0.083333
sabe64 mual23 -0.083333
sabe64 roen52 -0.083333
shbs25 chip27 -0.583333
shbs25 greg44 -0.472222
shbs25 guah24 -1.250000
shbs25 hahg14 -0.472222
shbs25 hous14 -0.583333
shbs25 mual23 -0.583333
shbs25 roen52 -0.583333
smtt74 chip27 0.041667
smtt74 greg44 0.152778
smtt74 guah24 -0.625000
smtt74 hahg14 0.152778
smtt74 hous14 0.041667
smtt74 mual23 0.041667
smtt74 roen52 0.041667
stjn43 chip27 -0.083333
stjn43 greg44 0.027778
stjn43 guah24 -0.750000
stjn43 hahg14 0.027778
stjn43 hous14 -0.083333
stjn43 mual23 -0.083333
stjn43 roen52 -0.083333
wadn66 chip27 -0.083333
wadn66 greg44 0.027778
wadn66 guah24 -0.750000
wadn66 hahg14 0.027778
wadn66 hous14 -0.083333
wadn66 mual23 -0.083333
wadn66 roen52 -0.083333
barp60 axht13 0.906061
barp60 brbz64 0.656061
barp60 brel45 0.322727
barp60 deeu72 -0.427273
barp60 hazh36 0.906061
barp60 hobs94 0.322727
barp60 hozg91 0.950000
barp60 jadg52 0.906061
barp60 leln17 0.572727
barp60 reeg74 0.739394
barp60 rodh93 0.572727
barp60 whqy96 0.572727
axht13 hojd22 -0.822727
axht13 moub86 -1.072727
axht13 piey32 -0.989394
axht13 sizl24 -0.906061
axht13 arus37 -1.072727
axht13 barp74 -1.239394
axht13 benz72 -1.265152
axht13 bras17 -1.174242
axht13 ches11 -1.072727
axht13 gehw13 -1.239394
axht13 maqh17 -0.239394
brbz64 hojd22 -0.572727
brbz64 moub86 -0.822727
brbz64 piey32 -0.739394
brbz64 sizl24 -0.656061
brbz64 arus37 -0.822727
brbz64 barp74 -0.989394
brbz64 benz72 -1.015152
brbz64 bras17 -0.924242
brbz64 ches11 -0.822727
brbz64 gehw13 -0.989394
brbz64 maqh17 0.010606
brel45 hojd22 -0.239394
brel45 moub86 -0.489394
brel45 piey32 -0.406061
brel45 sizl24 -0.322727
brel45 arus37 -0.489394
brel45 barp74 -0.656061
brel45 benz72 -0.681818
brel45 bras17 -0.590909

brel45 ches11 -0.489394
brel45 gehw13 -0.656061
brel45 maqh17 0.343939
deeu72 hojd22 0.510606
deeu72 moub86 0.260606
deeu72 piey32 0.343939
deeu72 sizl24 0.427273
deeu72 arus37 0.260606
deeu72 barp74 0.093939
deeu72 benz72 0.068182
deeu72 bras17 0.159091
deeu72 ches11 0.260606
deeu72 gehw13 0.093939
deeu72 maqh17 1.093939
hazh36 hojd22 -0.822727
hazh36 moub86 -1.072727
hazh36 piey32 -0.989394
hazh36 sizl24 -0.906061
hazh36 arus37 -1.072727
hazh36 barp74 -1.239394
hazh36 benz72 -1.265152
hazh36 bras17 -1.174242
hazh36 ches11 -1.072727
hazh36 gehw13 -1.239394
hazh36 maqh17 -0.239394
hobs94 hojd22 -0.239394
hobs94 moub86 -0.489394
hobs94 piey32 -0.406061
hobs94 sizl24 -0.322727
hobs94 arus37 -0.489394
hobs94 barp74 -0.656061
hobs94 benz72 -0.681818
hobs94 bras17 -0.590909
hobs94 ches11 -0.489394
hobs94 gehw13 -0.656061
hobs94 maqh17 0.343939
hozg91 hojd22 -0.866667
hozg91 moub86 -1.116667
hozg91 piey32 -1.033333
hozg91 sizl24 -0.950000
hozg91 arus37 -1.116667
hozg91 barp74 -1.283333
hozg91 ches11 -1.116667
hozg91 gehw13 -1.283333
hozg91 maqh17 -0.283333
jadg52 hojd22 -0.822727
jadg52 moub86 -1.072727
jadg52 piey32 -0.989394
jadg52 sizl24 -0.906061
jadg52 arus37 -1.072727
jadg52 barp74 -1.239394
jadg52 benz72 -1.265152
jadg52 bras17 -1.174242
jadg52 ches11 -1.072727
jadg52 gehw13 -1.239394
jadg52 maqh17 -0.239394
leln17 hojd22 -0.489394
leln17 moub86 -0.739394
leln17 piey32 -0.656061
leln17 sizl24 -0.572727
leln17 arus37 -0.739394
leln17 barp74 -0.906061
leln17 benz72 -0.931818
leln17 bras17 -0.840909
leln17 ches11 -0.739394
leln17 gehw13 -0.906061
leln17 maqh17 0.093939
reeg74 hojd22 -0.656061
reeg74 moub86 -0.906061
reeg74 piey32 -0.822727
reeg74 sizl24 -0.739394
reeg74 arus37 -0.906061
reeg74 barp74 -1.072727
reeg74 benz72 -1.098485
reeg74 bras17 -1.007576
reeg74 ches11 -0.906061
reeg74 gehw13 -1.072727
reeg74 maqh17 -0.072727
rodh93 hojd22 -0.489394
rodh93 moub86 -0.739394
rodh93 piey32 -0.656061
rodh93 sizl24 -0.572727
rodh93 arus37 -0.739394
rodh93 barp74 -0.906061
rodh93 benz72 -0.931818
rodh93 bras17 -0.840909
rodh93 ches11 -0.739394
rodh93 gehw13 -0.906061
rodh93 maqh17 0.093939
whqy96 hojd22 -0.489394
whqy96 moub86 -0.739394
whqy96 piey32 -0.656061
whqy96 sizl24 -0.572727
whqy96 arus37 -0.739394
whqy96 barp74 -0.906061
whqy96 benz72 -0.931818
whqy96 bras17 -0.840909
whqy96 ches11 -0.739394
whqy96 gehw13 -0.906061
whqy96 maqh17 0.093939

C.5 Tension Component of Muir-Joinson CMC Message-Level Prediction Graph

ixclde kzmern 0.004664
ixclde qilped -0.000143
ixclde ieogpe 0.004664
ixclde teepln 0.004664
ixclde rtoxwf 0.023628
ixclde smwext 0.005400
ixclde unblpe 0.011282
ixclde wrplod 0.004664
kzmern abofef -0.003890
kzmern nazwle 0.002689
kzmern hblqra -0.000883
kzmern brodsx -0.000690
kzmern cletym 0.002689
kzmern eglprn 0.007496
kzmern jroscd -0.002746
oslebm sprake -0.003019
oslebm twqisp -0.003019
oslebm ewtpen -0.003019
oslebm trwlia -0.003019
oslebm baolin -0.006643
oslebm dqarlt -0.003019
sprake zeplea -0.000604
sprake yaspeh -0.000604
sprake wrelcu -0.000604
sprake eiorln -0.000604
sprake pefsla -0.000604
abofef qilped -0.000917
abofef ieogpe 0.003890
abofef teepln 0.003890
abofef rtoxwf 0.022854
abofef smwext 0.004626
abofef unblpe 0.010508
abofef wrplod 0.003890
nazwle qilped -0.007496
nazwle ieogpe -0.002689
nazwle teepln -0.002689
nazwle rtoxwf 0.016275
nazwle smwext -0.001953
nazwle unblpe 0.003929
nazwle wrplod -0.002689
hblqra qilped -0.003925
hblqra ieogpe 0.000883
hblqra teepln 0.000883
hblqra rtoxwf 0.019846
hblqra smwext 0.001618
hblqra unblpe 0.007500
hblqra wrplod 0.000883
qilped brodsx 0.004118
qilped cletym 0.007496
qilped eglprn 0.012304
qilped jroscd 0.002062
brodsx ieogpe 0.000690
brodsx teepln 0.000690
brodsx rtoxwf 0.019653
brodsx smwext 0.001425
brodsx unblpe 0.007307
brodsx wrplod 0.000690
cletym ieogpe -0.002689
cletym teepln -0.002689
cletym rtoxwf 0.016275
cletym smwext -0.001953
cletym unblpe 0.003929
cletym wrplod -0.002689
eglprn ieogpe -0.007496
eglprn teepln -0.007496
eglprn rtoxwf 0.011467
eglprn smwext -0.006761
eglprn unblpe -0.000879
eglprn wrplod -0.007496
ieogpe jroscd -0.002746
jroscd teepln 0.002746
jroscd rtoxwf 0.021710
jroscd smwext 0.003481
jroscd unblpe 0.009364
jroscd wrplod 0.002746
smwext unblpe 0.005882
zeplea twqisp 0.000604
zeplea ewtpen 0.000604
zeplea trwlia 0.000604
zeplea baolin -0.003019
zeplea dqarlt 0.000604
twqisp yaspeh -0.000604
twqisp wrelcu -0.000604
twqisp eiorln -0.000604
twqisp pefsla -0.000604
ewtpen yaspeh -0.000604
ewtpen wrelcu -0.000604
ewtpen eiorln -0.000604
ewtpen pefsla -0.000604
trwlia yaspeh -0.000604
trwlia wrelcu -0.000604
trwlia eiorln -0.000604
trwlia pefsla -0.000604
baolin yaspeh 0.003019
baolin wrelcu 0.003019
baolin eiorln 0.003019
baolin pefsla 0.003019
yaspeh dqarlt 0.000604
dqarlt wrelcu -0.000604
dqarlt eiorln -0.000604
dqarlt pefsla -0.000604
awpmnm uqeuiv -0.068621
awpmnm gleneo 0.068907
awpmnm lqwenm -0.128206
awpmnm pofcer 0.117929
awpmnm ffepxs 0.196986
awpmnm pckelf -0.216995
uqeuiv kfeopd 0.007375
uqeuiv limobr -0.156585
uqeuiv lkupib 0.161342
uqeuiv twbeif -0.114121
uqeuiv trzaqt 0.015900
jwexnt bcdelf -0.365837
jwexnt indpew -0.095938
jwexnt huiegp -0.439347
jwexnt ubxcle -0.035262
jwexnt nieftl 0.295442
jwexnt trwmet 0.013085
jwexnt vdlpes 0.070397
bcdelf mneaep -0.202969

bcdelf plyinw 0.474206
bcdelf diojgs 0.445446
bcdelf bveaqf -0.021427
bcdelf nidlse 0.597390
bcdelf psfief 0.354237
mneaep indpew 0.472869
mneaep huiegp 0.129460
mneaep ubxcle 0.533545
mneaep nieftl 0.864249
mneaep trwmet 0.581892
mneaep vdlpes 0.639204
plyinw indpew -0.204307
plyinw huiegp -0.547716
plyinw ubxcle -0.143630
plyinw nieftl 0.187073
plyinw trwmet -0.095284
plyinw vdlpes -0.037972
indpew diojgs 0.175547
indpew bveaqf -0.291326
indpew nidlse 0.327491
indpew psfief 0.084337
diojgs huiegp -0.518956
diojgs ubxcle -0.114870
diojgs nieftl 0.215834
diojgs trwmet -0.066524
diojgs vdlpes -0.009211
huiegp bveaqf 0.052083
huiegp nidlse 0.670900
huiegp psfief 0.427746
gleneo kfeopd -0.130153
gleneo limobr -0.294113
gleneo lkupib 0.023814
gleneo twbeif -0.251649
gleneo trzaqt -0.121628
kfeopd lqwenm -0.066960
kfeopd pofcer 0.179175
kfeopd ffepxs 0.258232
kfeopd pckelf -0.155749
lqwenm limobr -0.096999
lqwenm lkupib 0.220927
lqwenm twbeif -0.054535
lqwenm trzaqt 0.075486
pofcer limobr -0.343134
pofcer lkupib -0.025208
pofcer twbeif -0.300670
pofcer trzaqt -0.170649
ffepxs limobr -0.422192
ffepxs lkupib -0.104265
ffepxs twbeif -0.379728
ffepxs trzaqt -0.249707
limobr pckelf 0.008211
lkupib pckelf -0.309716
ubxcle bveaqf -0.352002
ubxcle nidlse 0.266815
ubxcle psfief 0.023661
nieftl bveaqf -0.682706
nieftl nidlse -0.063889
nieftl psfief -0.307043
bveaqf trwmet 0.400349
bveaqf vdlpes 0.457661
nidlse trwmet -0.218468
nidlse vdlpes -0.161156
psfief trwmet 0.024686
psfief vdlpes 0.081998
pckelf twbeif 0.034253
pckelf trzaqt 0.164274

C.6 Tension Component of Muir-Joinson CMC Relationship-Level Prediction Graph

ixclde kzmern 0.000000
ixclde qilped 0.000000
ixclde ieogpe 0.000000
ixclde teepln 0.000000
ixclde rtoxwf 0.000000
ixclde smwext 0.000000
ixclde unblpe 0.000000
ixclde wrplod 0.000000
kzmern abofef 0.000000
kzmern nazwle 0.000000
kzmern hblqra 0.000000
kzmern brodsx 0.000000
kzmern cletym 0.000000
kzmern eglprn 0.000000
kzmern jroscd 0.000000
oslebm sprake 0.000000
oslebm twqisp 0.000000
oslebm ewtpen 0.000000
oslebm trwlia 0.000000
oslebm baolin 0.000000
oslebm dqarlt 0.000000
sprake zeplea 0.000000
sprake yaspeh 0.000000
sprake wrelcu 0.000000
sprake eiorln 0.000000
sprake pefsla 0.000000
abofef qilped 0.000000
abofef ieogpe 0.000000
abofef teepln 0.000000
abofef rtoxwf 0.000000
abofef smwext 0.000000
abofef unblpe 0.000000
abofef wrplod 0.000000
nazwle qilped 0.000000
nazwle ieogpe 0.000000
nazwle teepln 0.000000
nazwle rtoxwf 0.000000
nazwle smwext 0.000000
nazwle unblpe 0.000000
nazwle wrplod 0.000000
hblqra qilped 0.000000
hblqra ieogpe 0.000000

hblqra teepln 0.000000
hblqra rtoxwf 0.000000
hblqra smwext 0.000000
hblqra unblpe 0.000000
hblqra wrplod 0.000000
qilped brodsx 0.000000
qilped cletym 0.000000
qilped eglprn 0.000000
qilped jroscd 0.000000
brodsx ieogpe 0.000000
brodsx teepln 0.000000
brodsx rtoxwf 0.000000
brodsx smwext 0.000000
brodsx unblpe 0.000000
brodsx wrplod 0.000000
cletym ieogpe 0.000000
cletym teepln 0.000000
cletym rtoxwf 0.000000
cletym smwext 0.000000
cletym unblpe 0.000000
cletym wrplod 0.000000
eglprn ieogpe 0.000000
eglprn teepln 0.000000
eglprn rtoxwf 0.000000
eglprn smwext 0.000000
eglprn unblpe 0.000000
eglprn wrplod 0.000000
ieogpe jroscd 0.000000
jroscd teepln 0.000000
jroscd rtoxwf 0.000000
jroscd smwext 0.000000
jroscd unblpe 0.000000
jroscd wrplod 0.000000
smwext unblpe 0.000000
zeplea twqisp 0.000000
zeplea ewtpen 0.000000
zeplea trwlia 0.000000
zeplea baolin 0.000000
zeplea dqarlt 0.000000
twqisp yaspeh 0.000000
twqisp wrelcu 0.000000
twqisp eiorln 0.000000
twqisp pefsla 0.000000
ewtpen yaspeh 0.000000
ewtpen wrelcu 0.000000
ewtpen eiorln 0.000000
ewtpen pefsla 0.000000
trwlia yaspeh 0.000000
trwlia wrelcu 0.000000
trwlia eiorln 0.000000
trwlia pefsla 0.000000
baolin yaspeh 0.000000
baolin wrelcu 0.000000
baolin eiorln 0.000000
baolin pefsla 0.000000
yaspeh dqarlt 0.000000
dqarlt wrelcu 0.000000
dqarlt eiorln 0.000000
dqarlt pefsla 0.000000
awpmnm uqeuiv -0.555556
awpmnm gleneo 0.444444
awpmnm lqwenm -0.222222
awpmnm pofcer 0.444444
awpmnm ffepxs 0.444444
awpmnm pckelf -0.555556
uqeuiv kfeopd 0.222222
uqeuiv limobr -0.111111
uqeuiv lkupib 0.888889
uqeuiv twbeif -0.111111
uqeuiv trzaqt 0.555556
jwexnt bcdelf -0.918367
jwexnt indpew -0.632653
jwexnt huiegp -0.918367
jwexnt ubxcle -0.346939
jwexnt nieftl 0.510204
jwexnt trwmet -0.346939
jwexnt vdlpes -0.346939
bcdelf mneaep -0.510204
bcdelf plyinw 0.632653
bcdelf diojgs 0.918367
bcdelf bveaqf -0.510204
bcdelf nidlse 1.204082
bcdelf psfief 0.346939
mneaep indpew 0.795918
mneaep huiegp 0.510204
mneaep ubxcle 1.081633
mneaep nieftl 1.938776
mneaep trwmet 1.081633
mneaep vdlpes 1.081633
plyinw indpew -0.346939
plyinw huiegp -0.632653
plyinw ubxcle -0.061224
plyinw nieftl 0.795918
plyinw trwmet -0.061224
plyinw vdlpes -0.061224
indpew diojgs 0.632653
indpew bveaqf -0.795918
indpew nidlse 0.918367
indpew psfief 0.061224
diojgs huiegp -0.918367
diojgs ubxcle -0.346939
diojgs nieftl 0.510204
diojgs trwmet -0.346939
diojgs vdlpes -0.346939
huiegp bveaqf -0.510204
huiegp nidlse 1.204082
huiegp psfief 0.346939
gleneo kfeopd -0.777778
gleneo limobr -1.111111
gleneo lkupib -0.111111
gleneo twbeif -1.111111
gleneo trzaqt -0.444444
kfeopd lqwenm 0.111111
kfeopd pofcer 0.777778
kfeopd ffepxs 0.777778
kfeopd pckelf -0.222222
lqwenm limobr -0.444444
lqwenm lkupib 0.555556
lqwenm twbeif -0.444444
lqwenm trzaqt 0.222222

pofcer limobr -1.111111
pofcer lkupib -0.111111
pofcer twbeif -1.111111
pofcer trzaqt -0.444444
ffepxs limobr -1.111111
ffepxs lkupib -0.111111
ffepxs twbeif -1.111111
ffepxs trzaqt -0.444444
limobr pckelf 0.111111
lkupib pckelf -0.888889
ubxcle bveaqf -1.081633
ubxcle nidlse 0.632653
ubxcle psfief -0.224490
nieftl bveaqf -1.938776
nieftl nidlse -0.224490
nieftl psfief -1.081633
bveaqf trwmet 1.081633
bveaqf vdlpes 1.081633
nidlse trwmet -0.632653
nidlse vdlpes -0.632653
psfief trwmet 0.224490
psfief vdlpes 0.224490
pckelf twbeif -0.111111
pckelf trzaqt 0.555556

Appendix D: Feature Distribution Charts

This appendix contains charts illustrating the distributions of features, as discussed in Chapter 7. A note on the charts: most of these features are very sparse, so for many features, the spike at zero on a frequency chart would overwhelm the other scores. Zero is therefore omitted in all cases, and axes are adjusted to accommodate the non-zero frequency distributions.

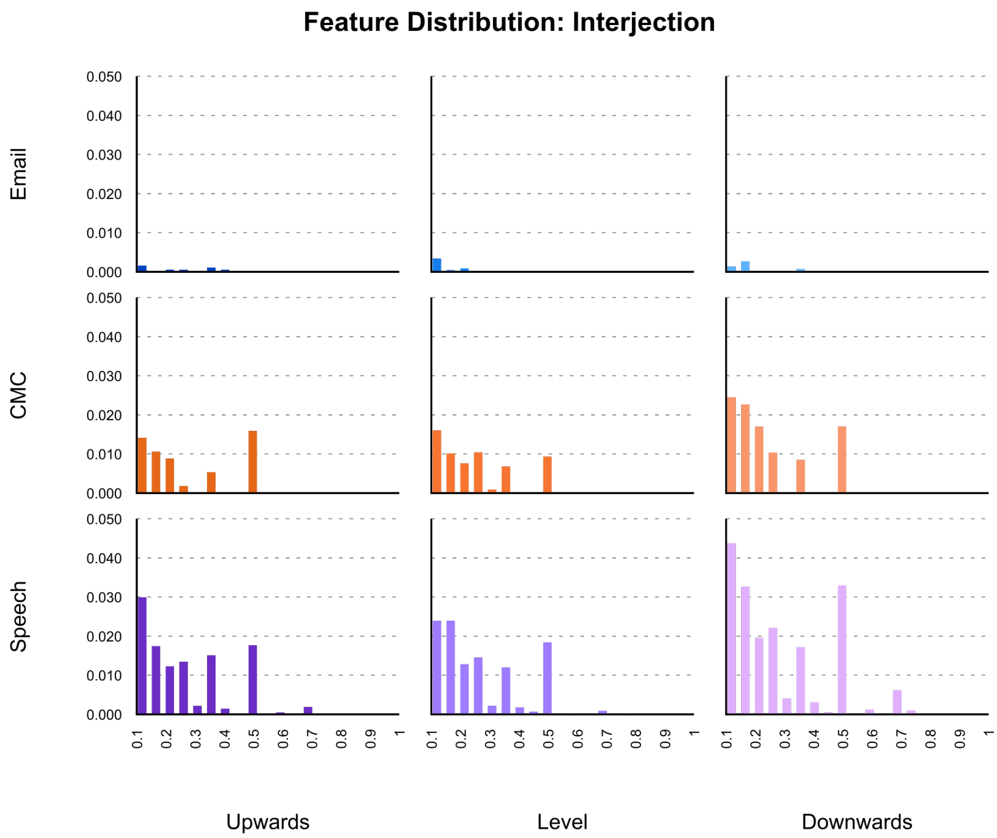

Figure 35: Frequency distributions for interjections

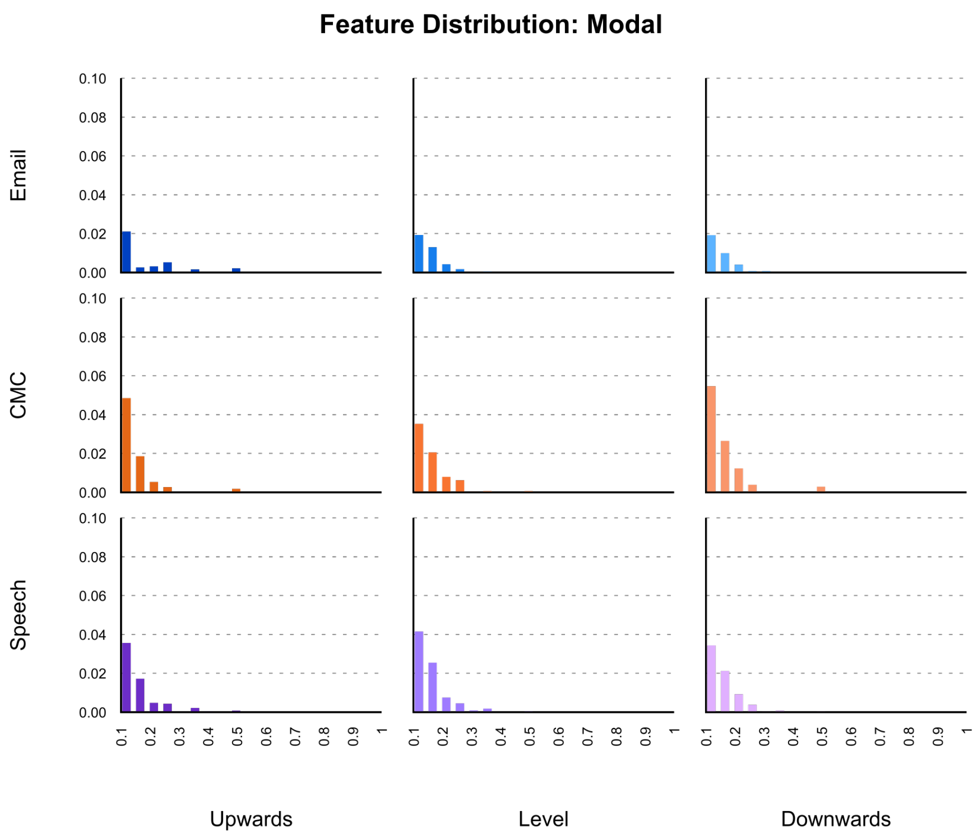

Figure 36: Frequency distributions for modal verbs

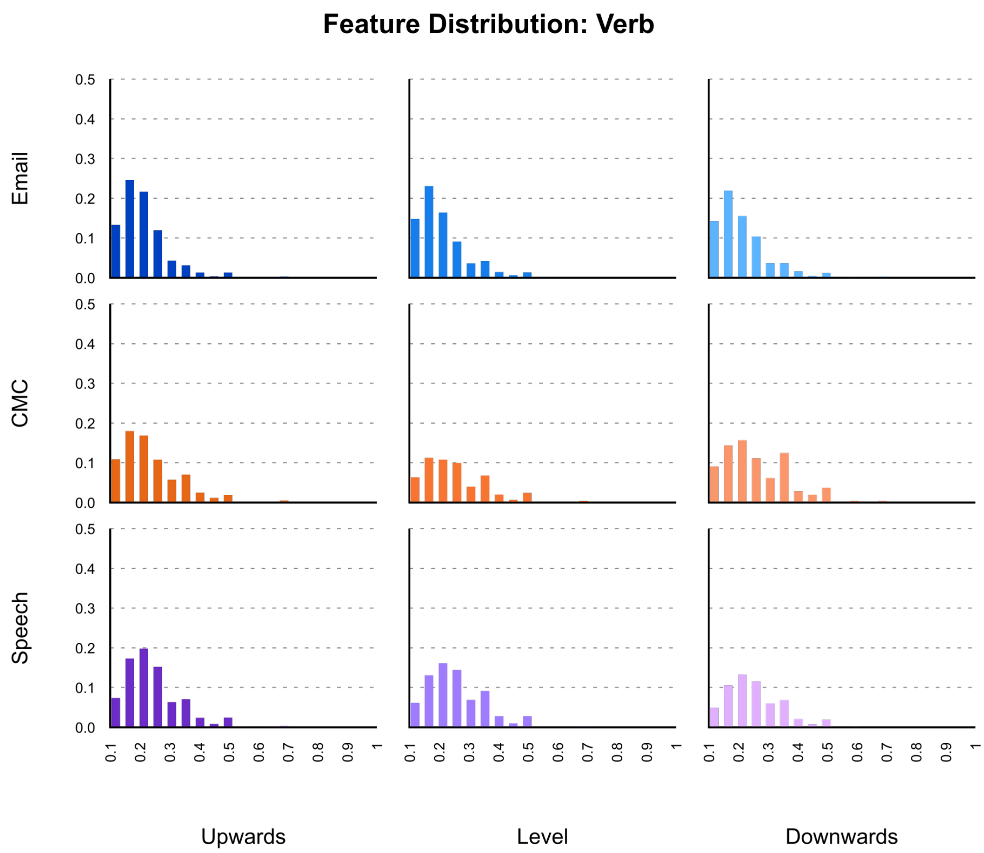

Figure 37: Frequency distributions for verbs

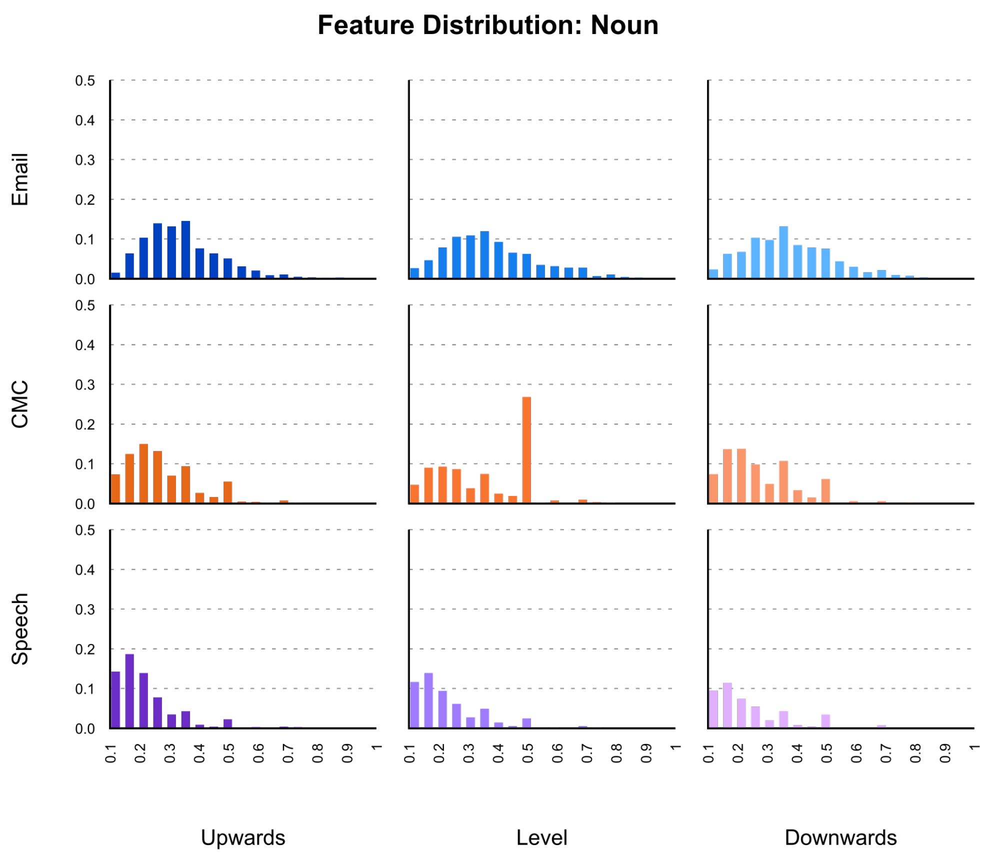

Figure 38: Frequency distributions for nouns

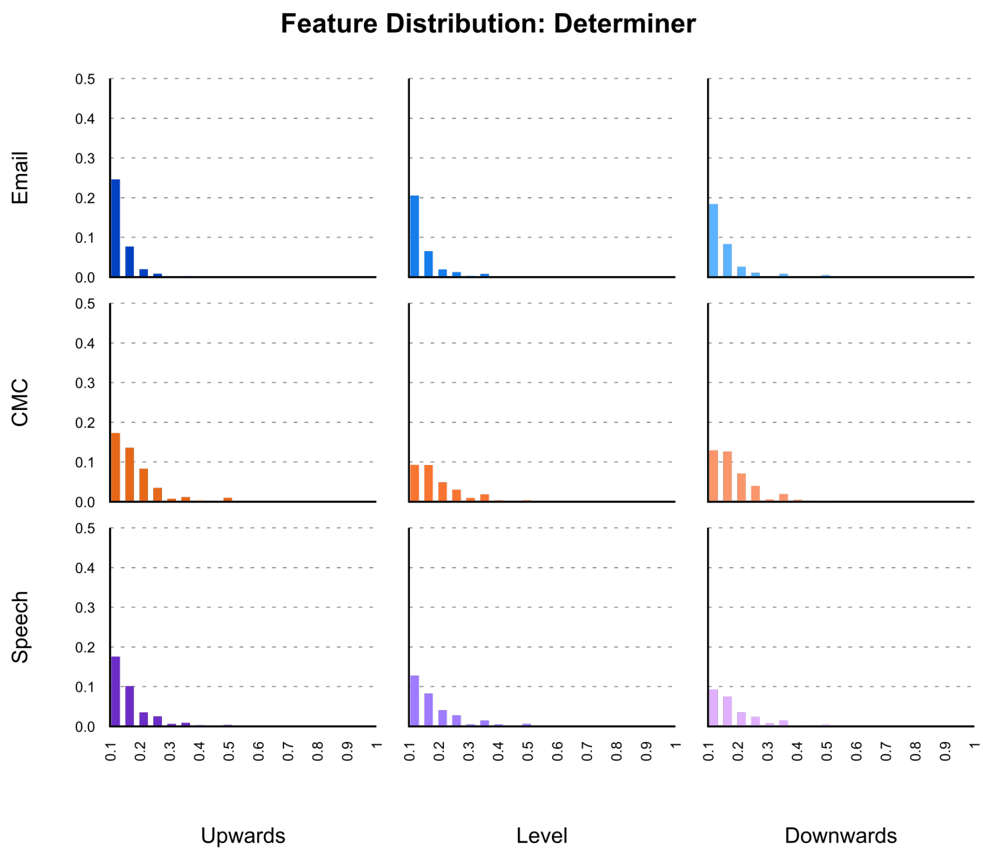

Figure 39: Frequency distributions for determiners

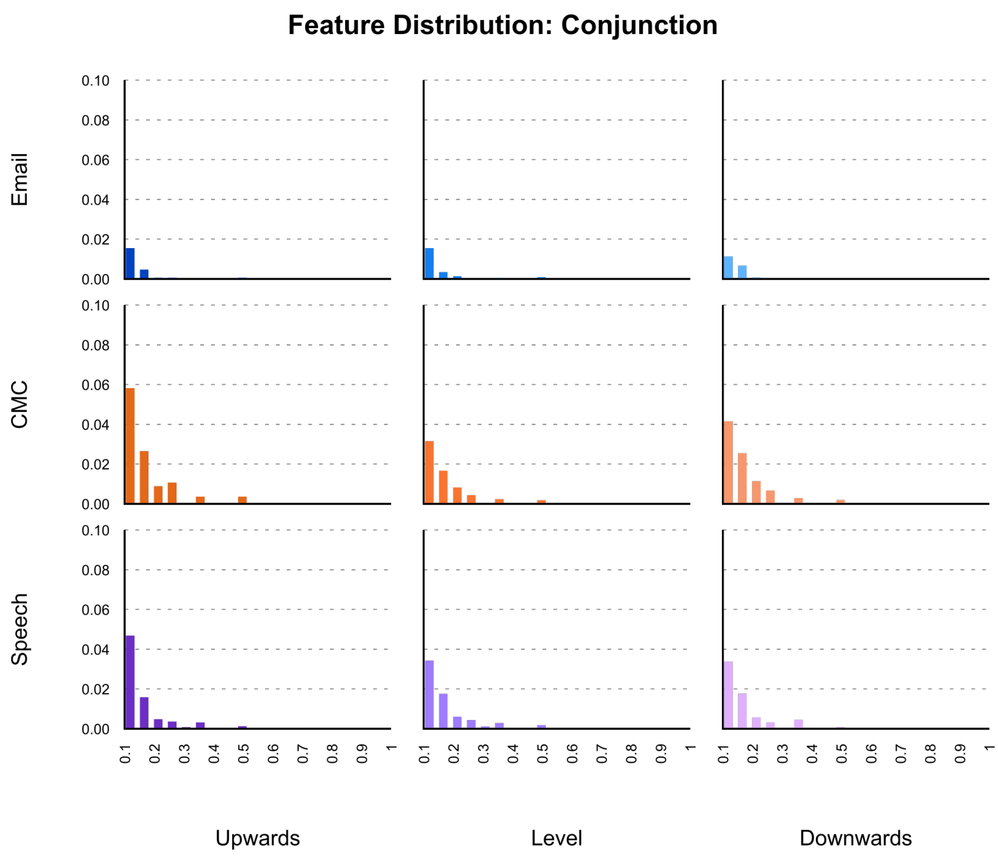

Figure 40: Frequency distributions for conjunctions

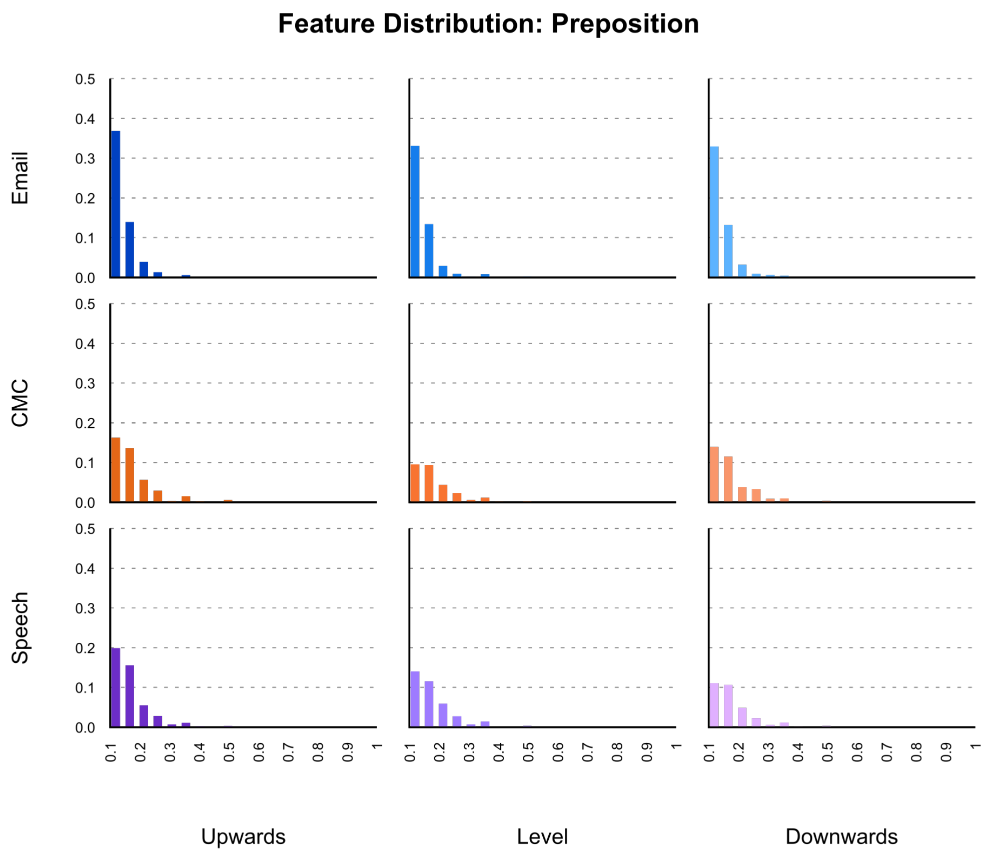

Figure 41: Frequency distributions for prepositions

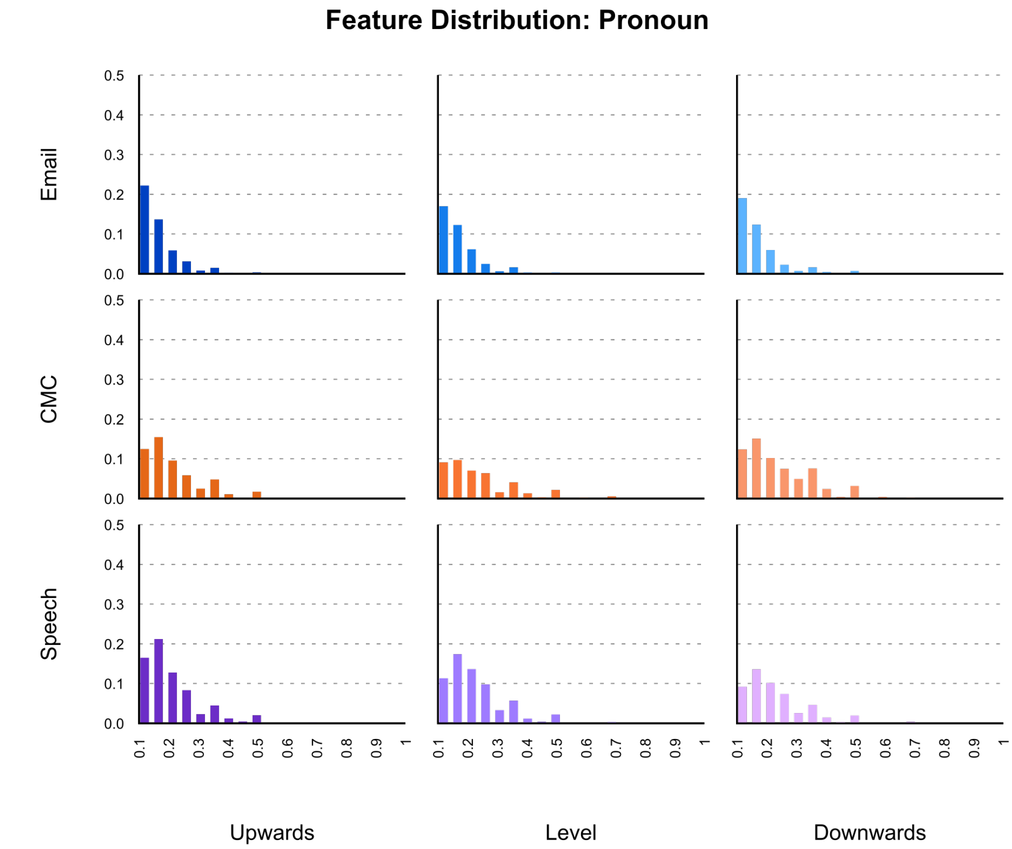

Figure 42: Frequency distributions for pronouns

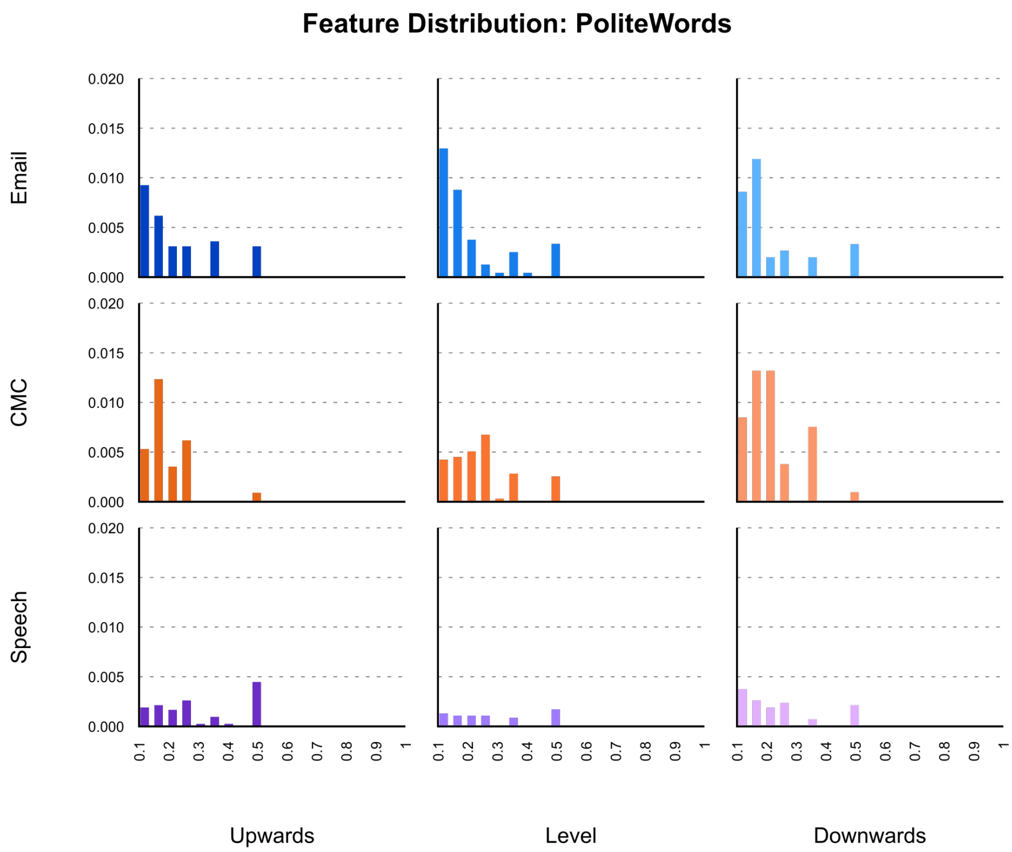

Figure 43: Frequency distributions for polite words

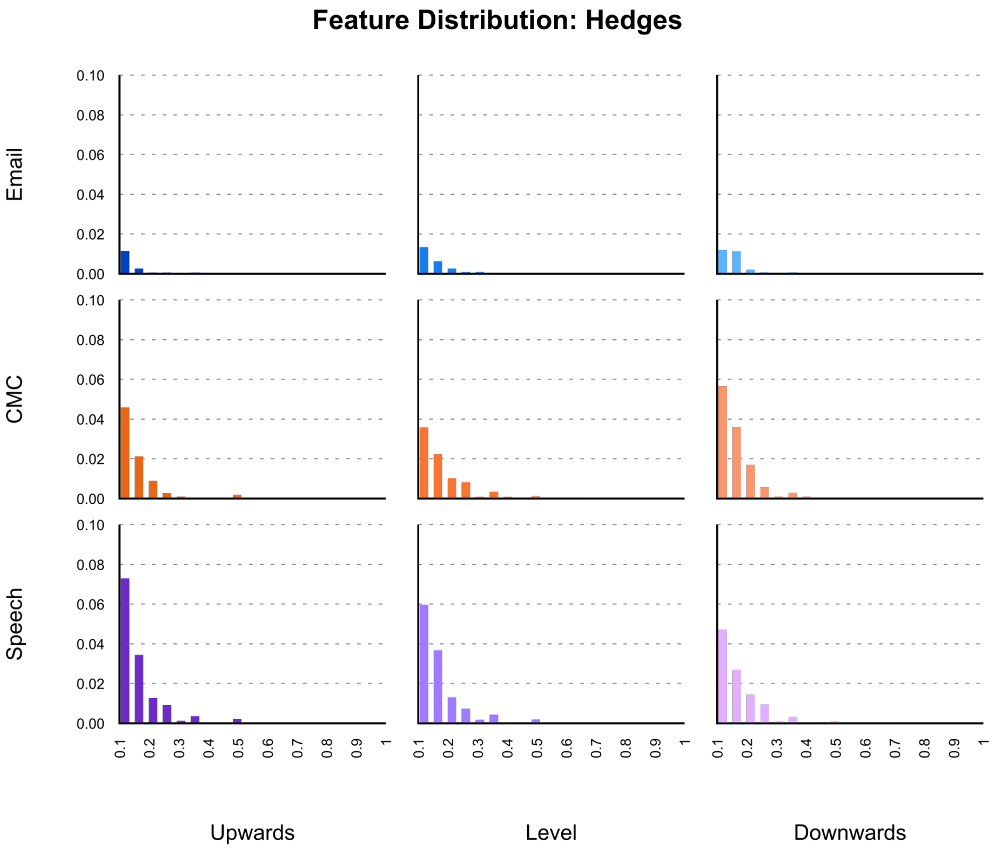

Figure 44: Frequency distributions for hedges

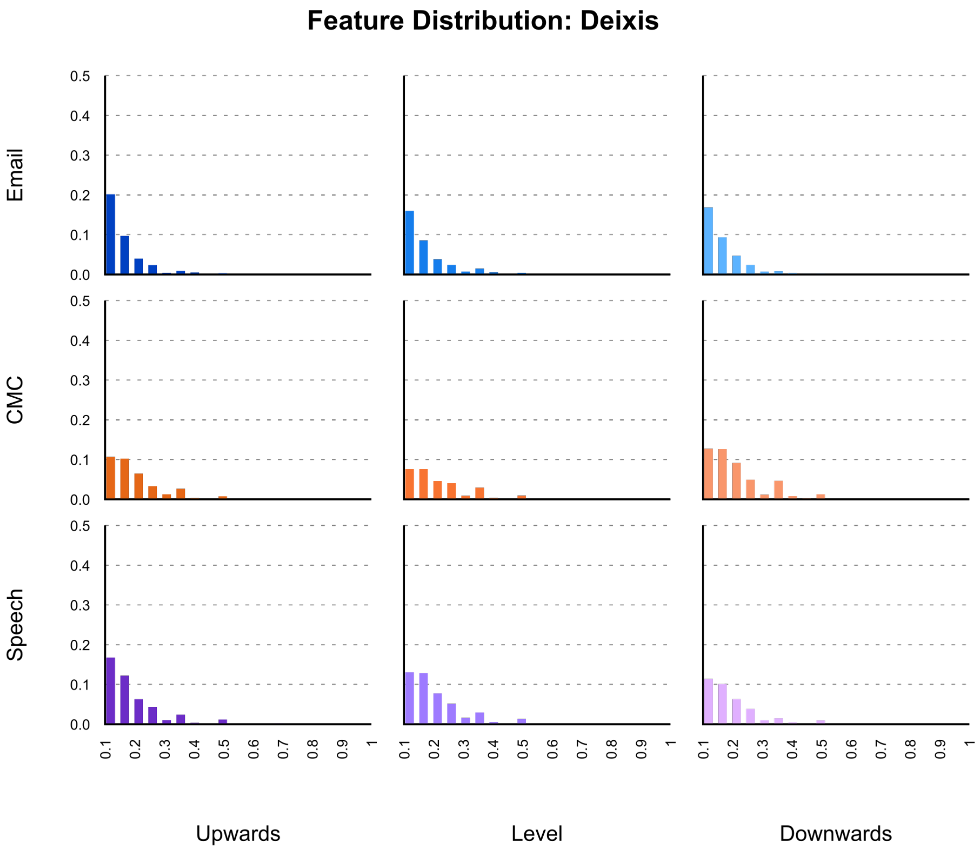

Figure 45: Frequency distributions for deixis

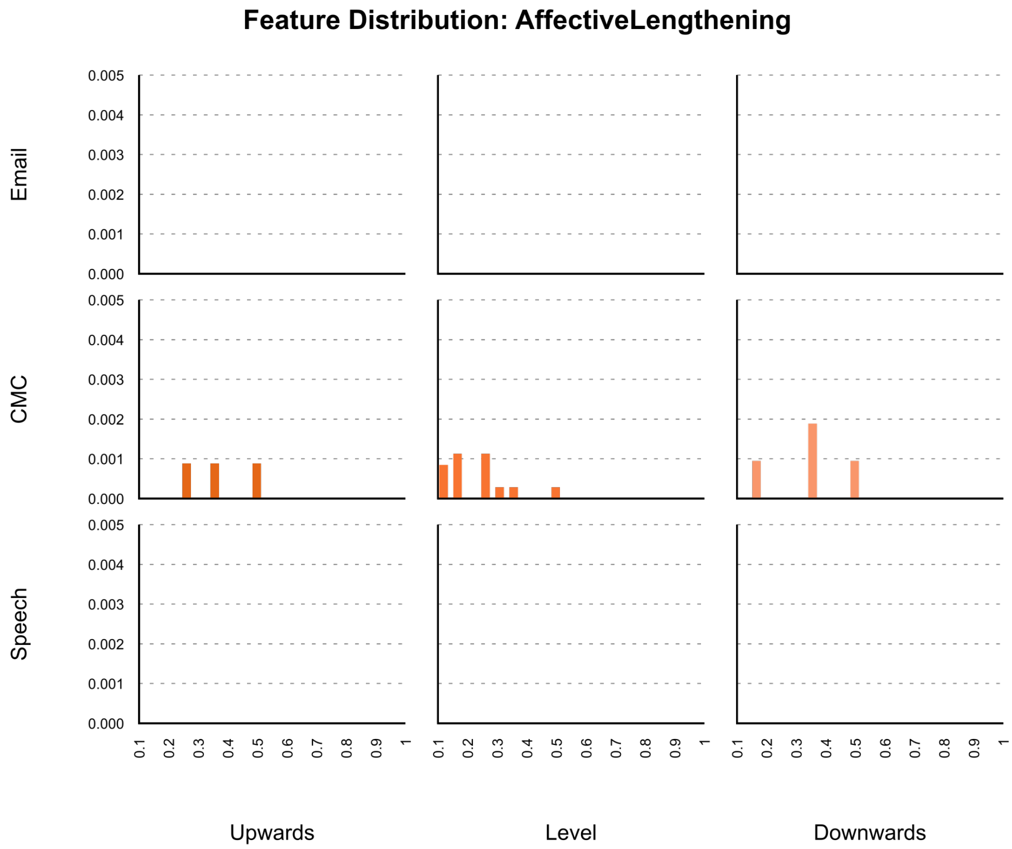

Figure 46: Frequency distributions for affective lengthening

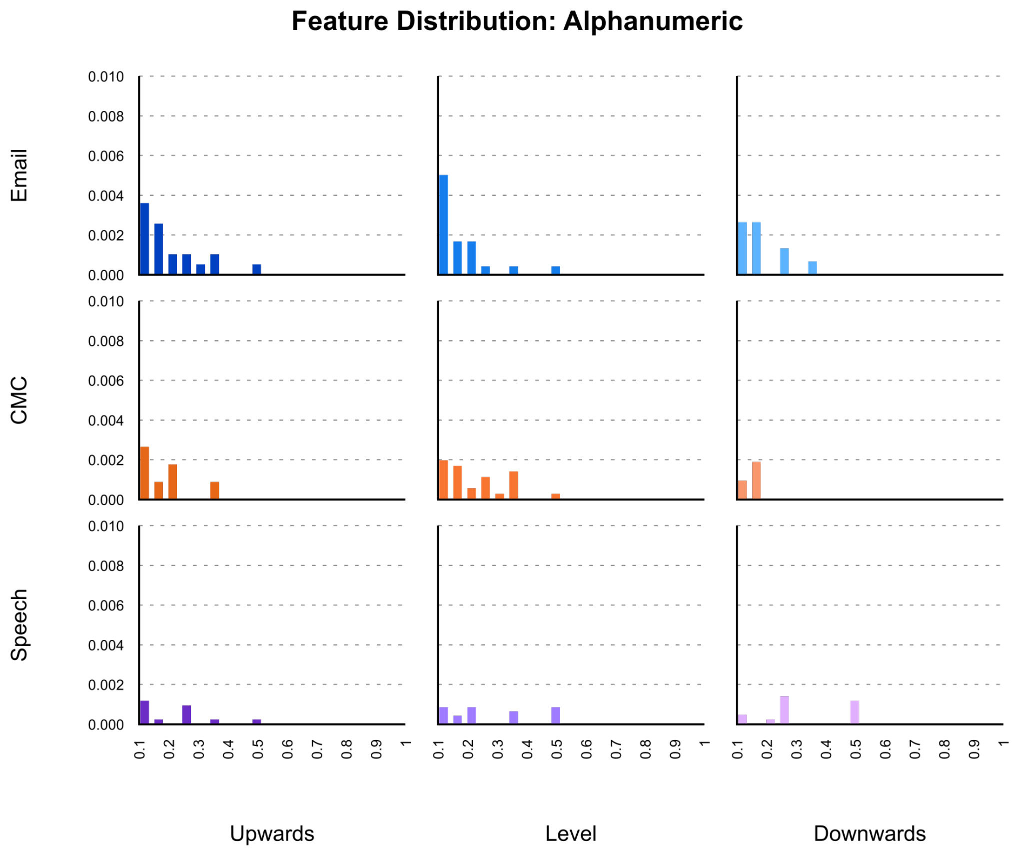

Figure 47: Frequency distributions for alphanumeric words

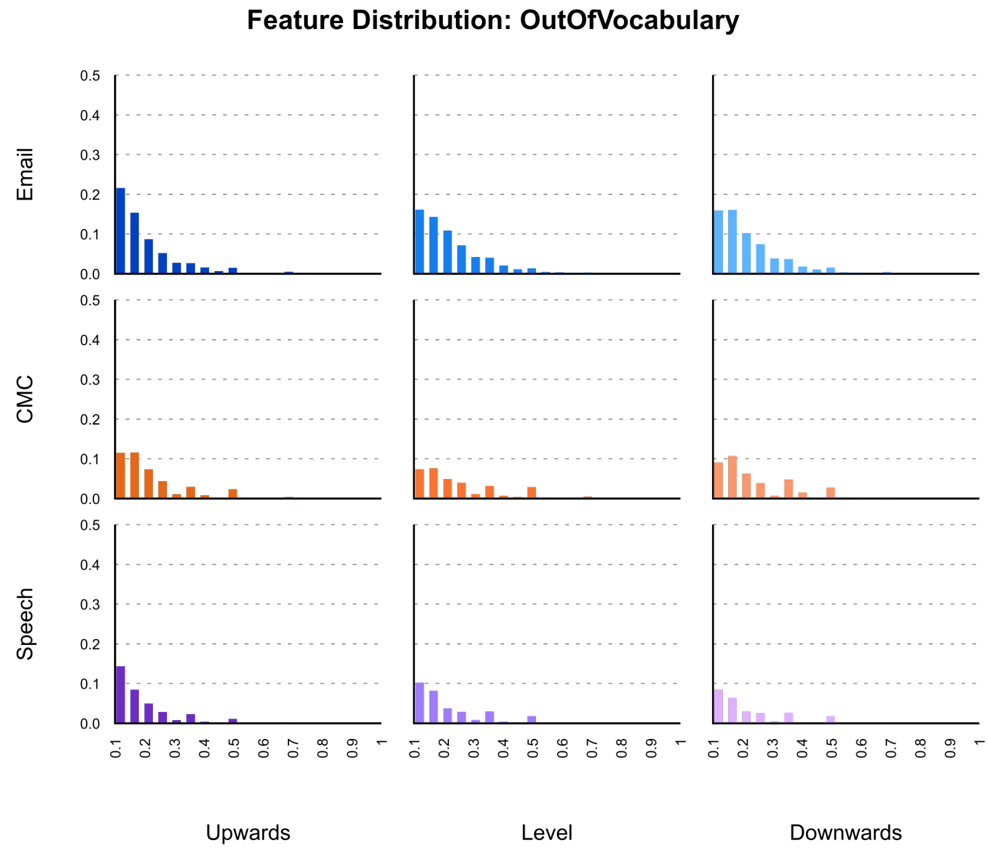

Figure 48: Frequency distributions for out of vocabulary terms

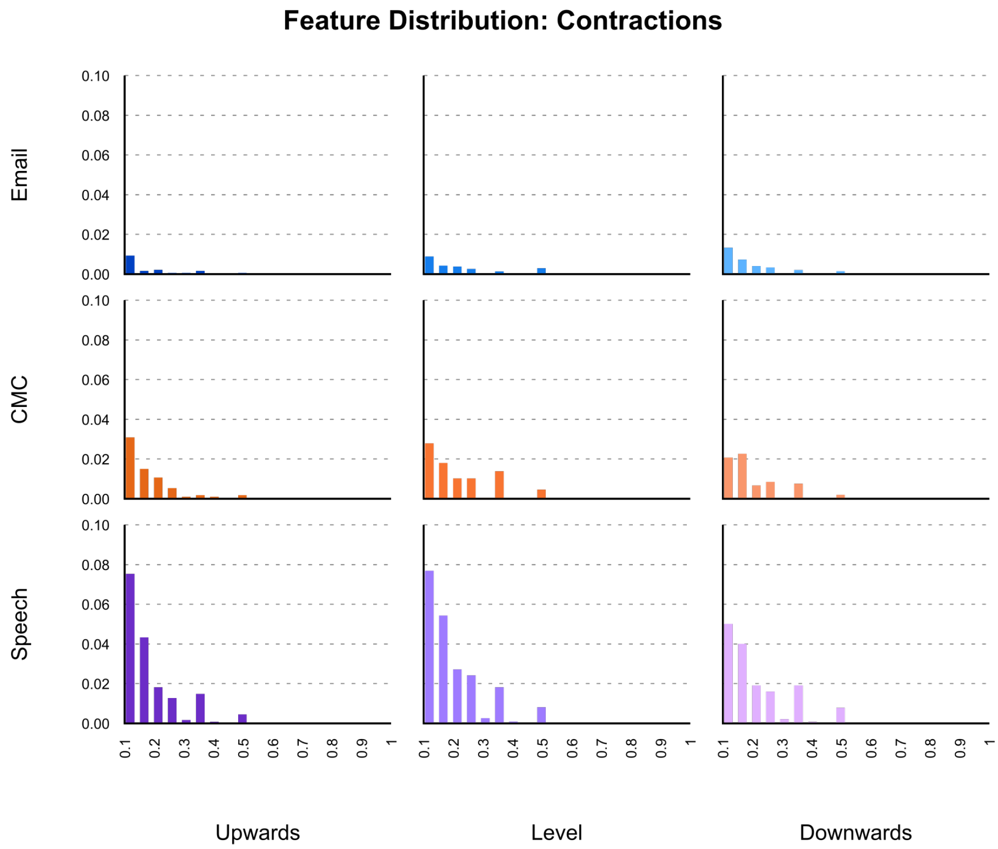

Figure 49: Frequency distributions for contractions

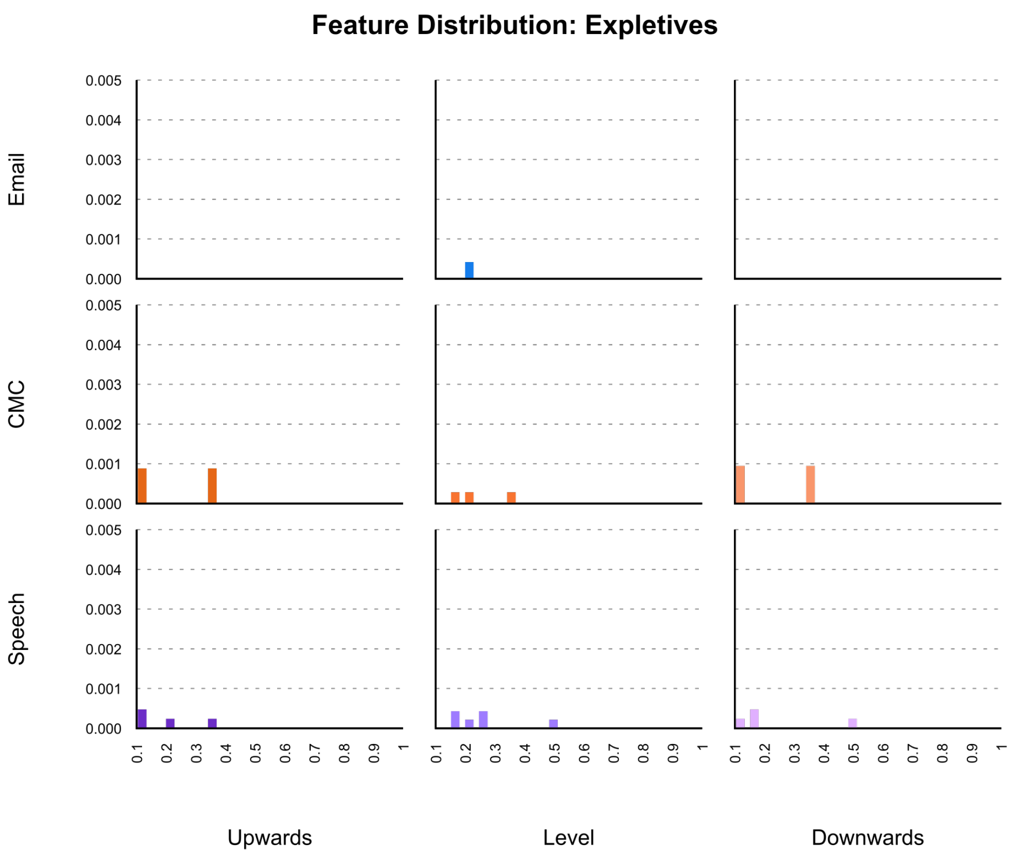

Figure 50: Frequency distributions for expletives

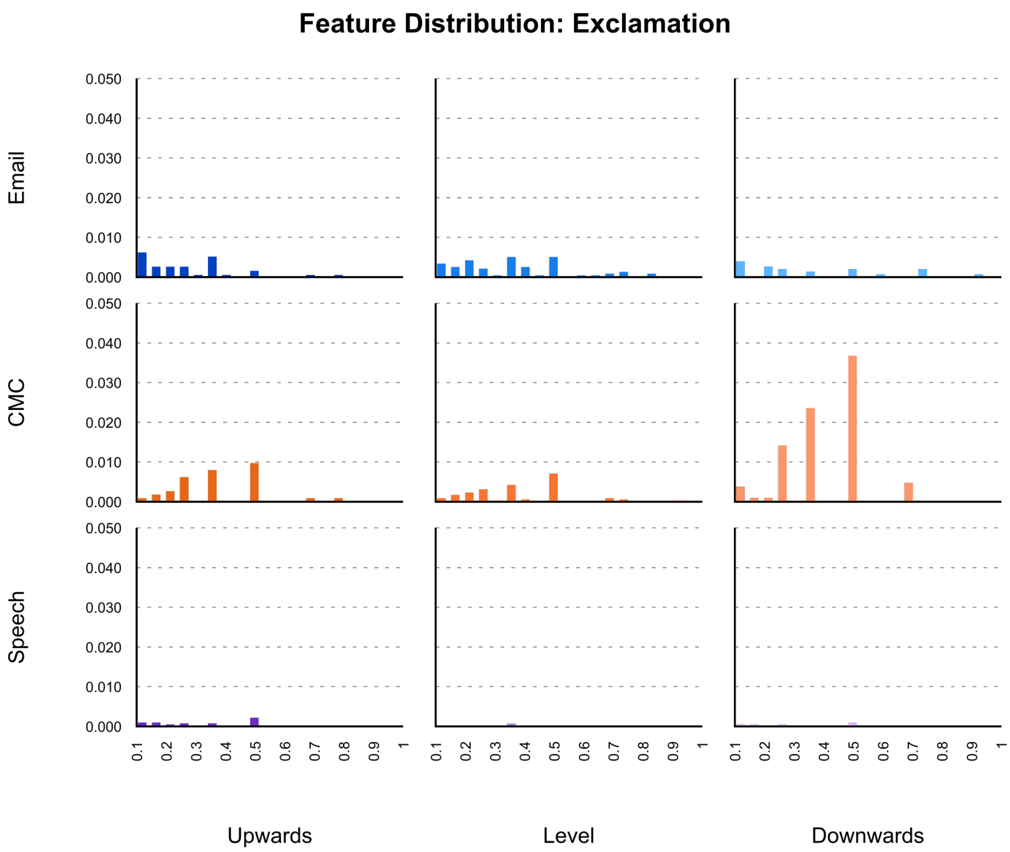

Figure 51: Frequency distributions for exclamation marks

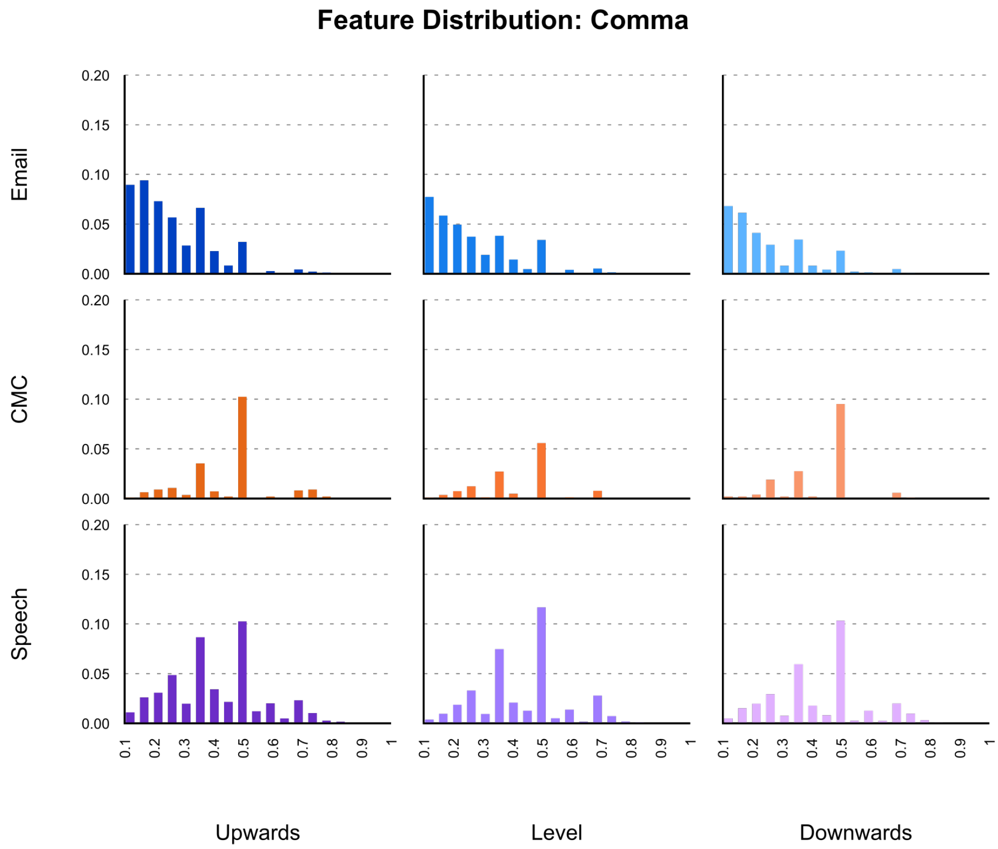

Figure 52: Frequency distributions for commas

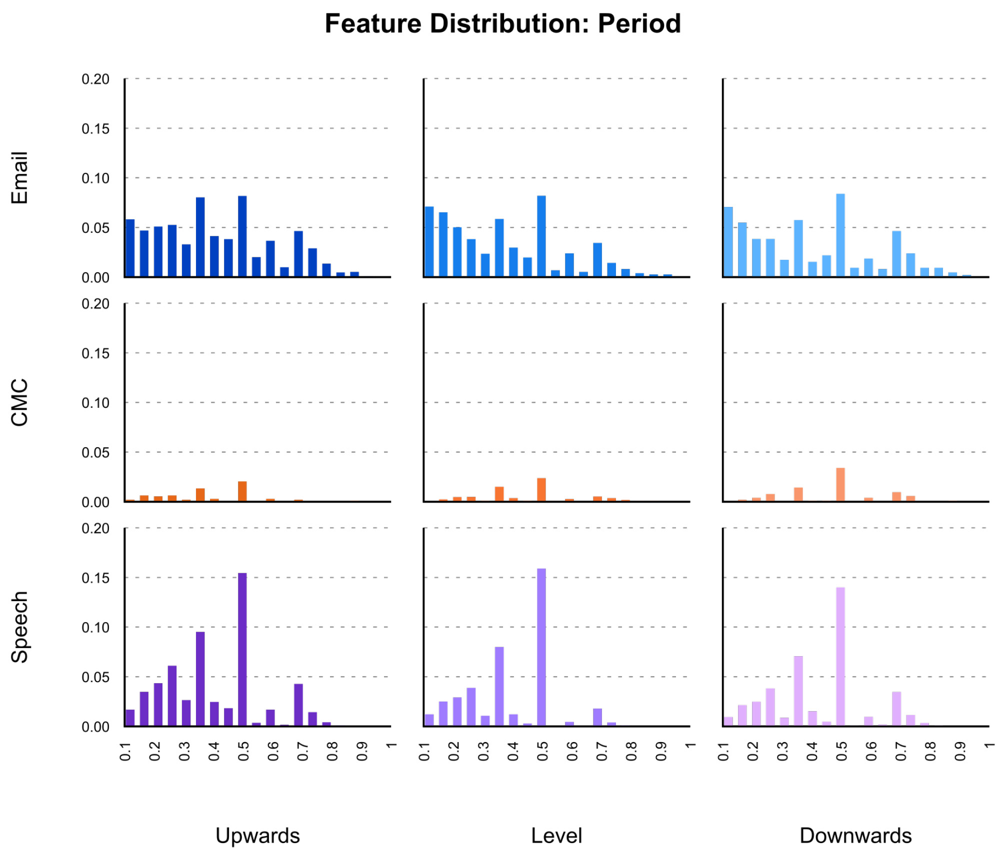

Figure 53: Frequency distributions for periods

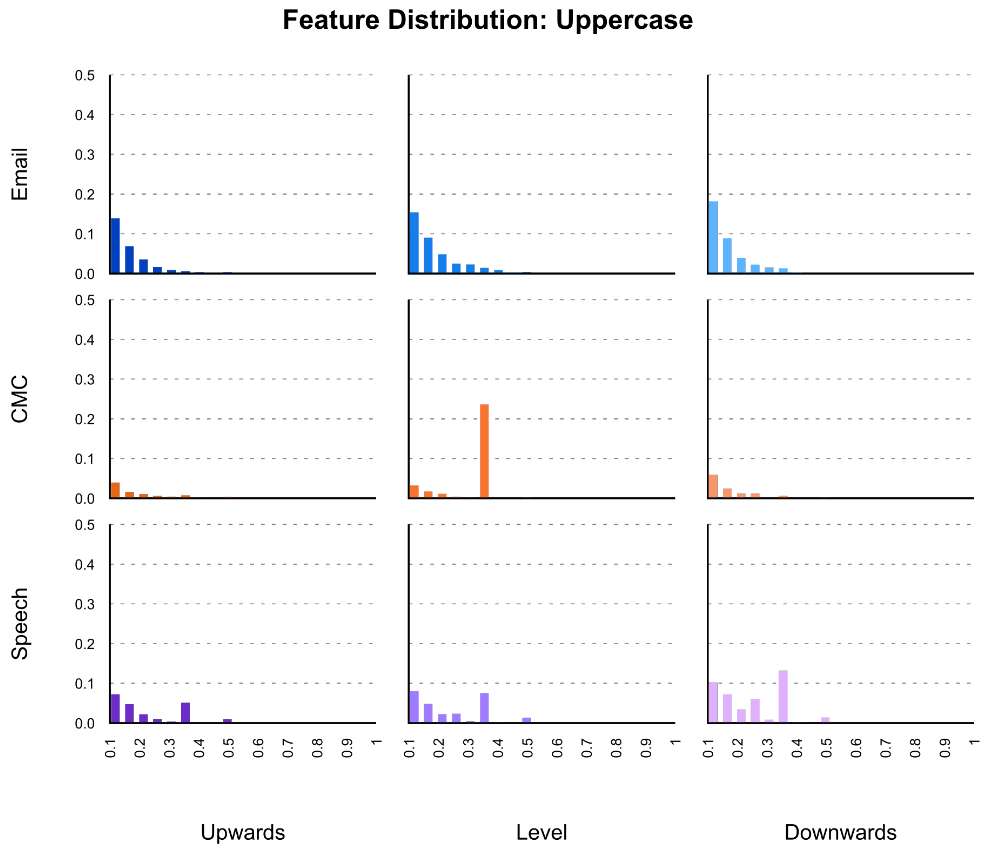

Figure 54: Frequency distributions for uppercase letters

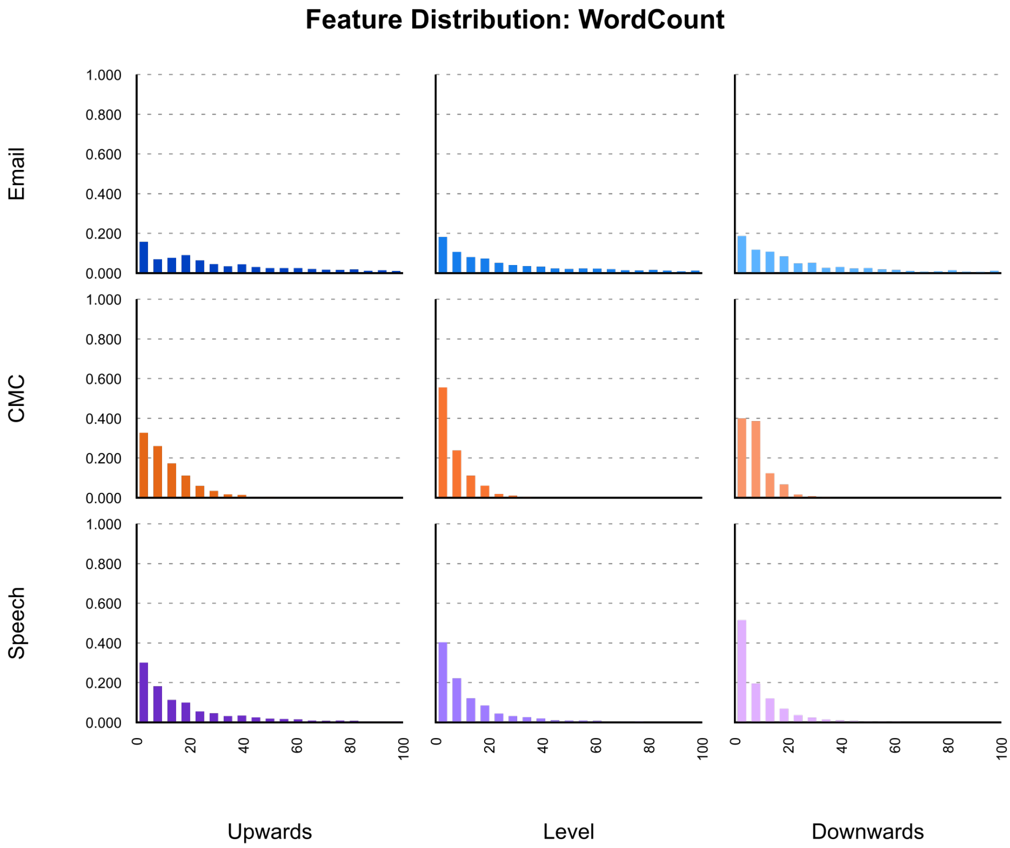

Figure 55: Frequency distributions for word counts

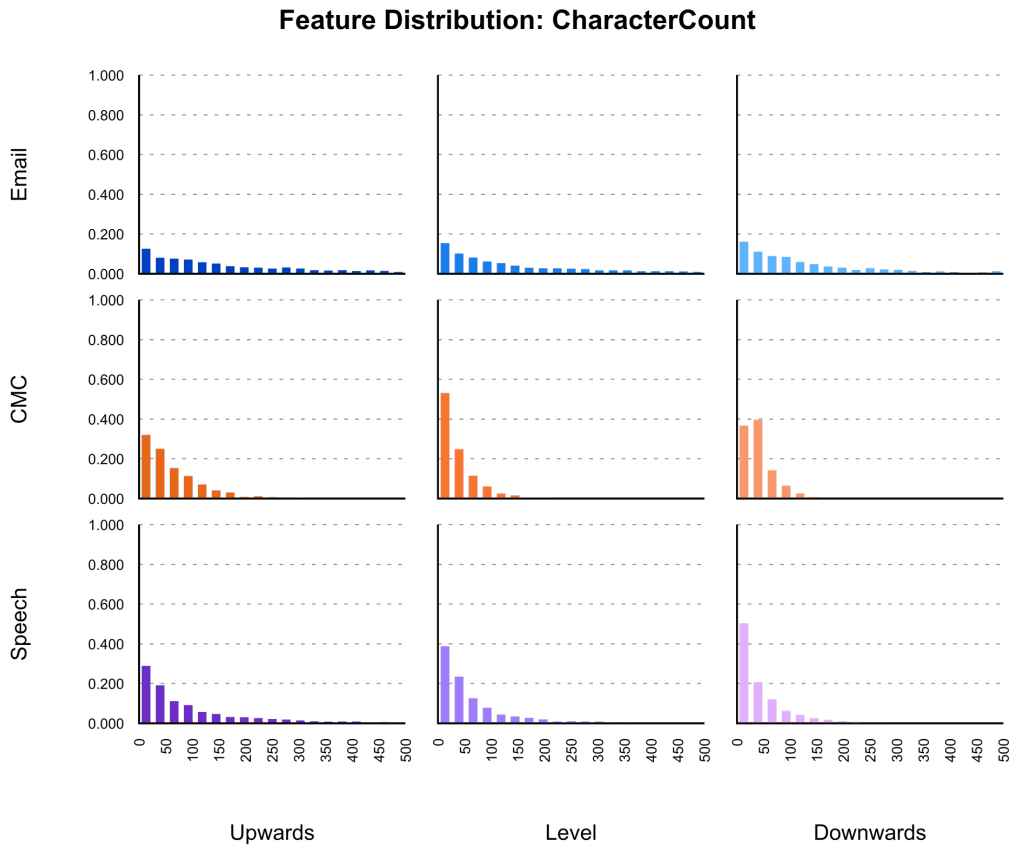

Figure 56: Frequency distributions for character counts

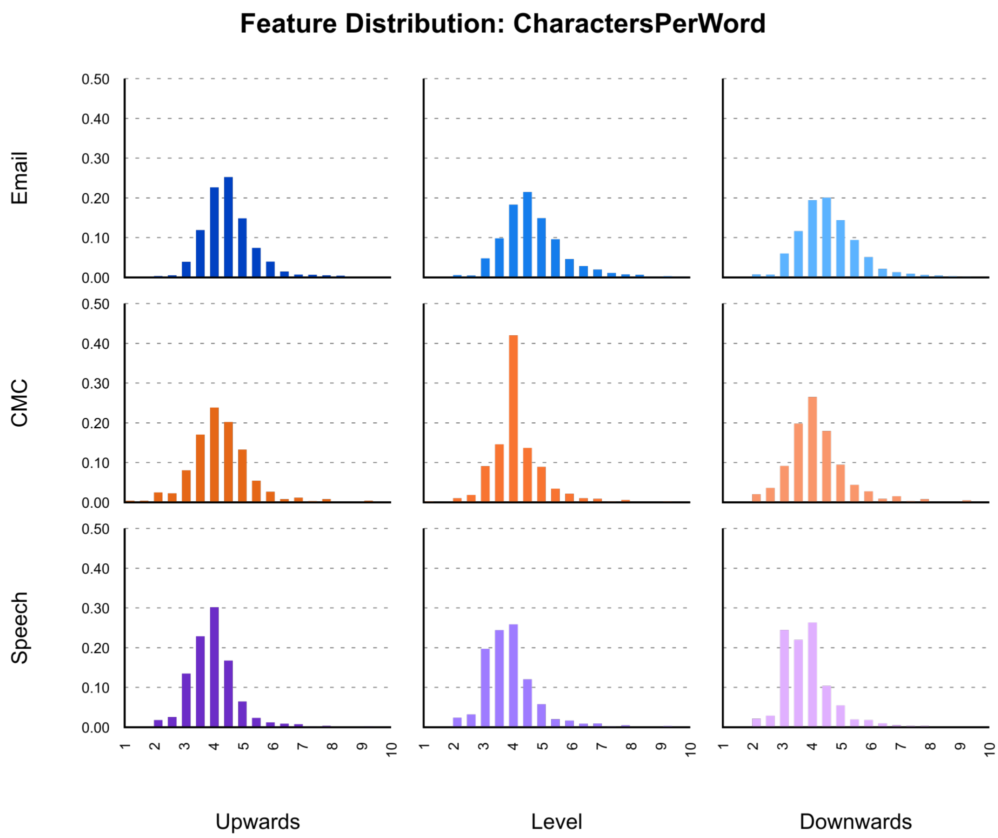

Figure 57: Frequency distributions for characters per word

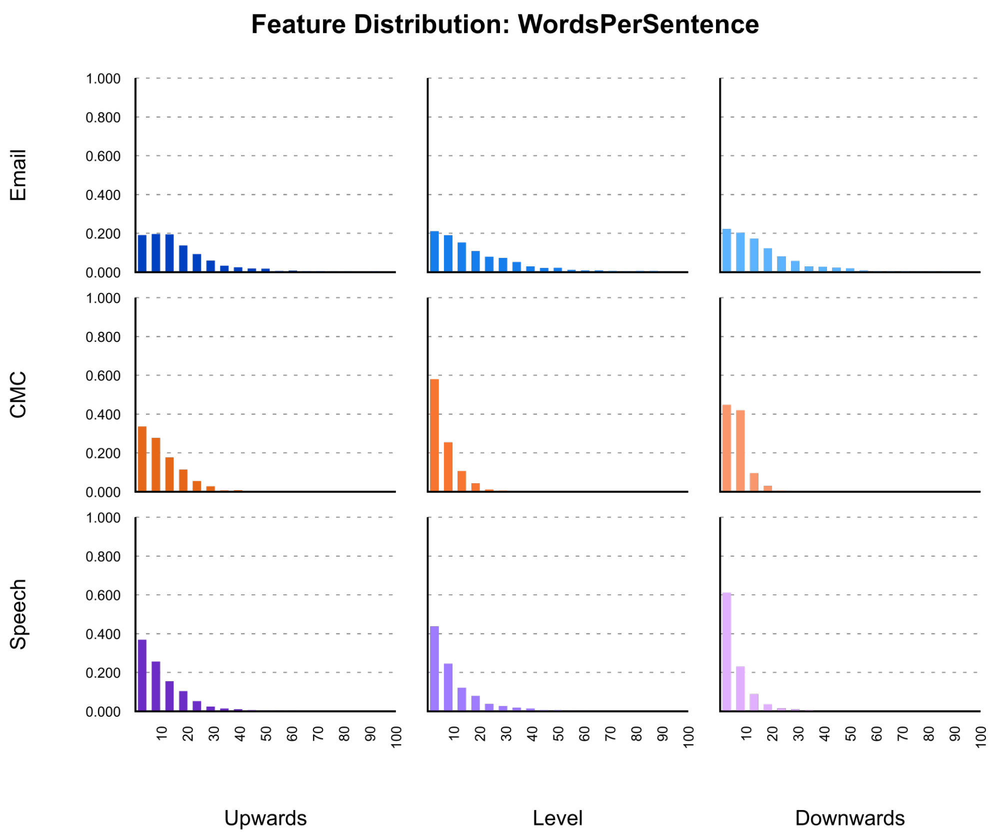

Figure 58: Frequency distributions for words per sentence

10. References

Abu Sheikha, F., & Inkpen, D. (2010). Automatic Classification of Documents by Formality. *IEEE*.

Agarwal, A., Omuya, A., Harnly, A., & Rambow, O. (2012). A Comprehensive Gold Standard for the Enron Organizational Hierarchy. In *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Short Papers - Volume 2*, 161–165.

Agarwal, A., & Rambow, O. (2010). Automatic Detection and Classification of Social Events. *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*.

Akiva, N., & Koppel, M. (2013). A generic unsupervised method for decomposing multi-author documents. *Journal of the American Society for Information Science and Technology*, *64*(11), 2256-2264.

Arefin, A. S., Vimieiro, R., Riveros, C., Craig, H., & Moscato, P. (2014). An Information Theoretic Clustering Approach for Unveiling Authorship Affinities in Shakespearean Era Plays and Poems. *PLoS One*, *9*(10).

Argamon, S., Dhawle, S., Koppel, M., & Pennebaker, J. W. (2005). Lexical Predictors of Personality Type. In *Proceedings of the 2005 Joint Annual Meeting of the Interface and the Classification Society of North America*.

Argamon, S., Koppel, M., Pennebaker, J. W., & Schler, J. (2007). Mining the Blogosphere: Age, gender and the varieties of self-expression. *First Monday*, *12*(9).

Bailenson, J. N., & Yee, N. (2005). Digital chameleons automatic assimilation of nonverbal gestures in immersive virtual environments. *Psychological science*, *16*(10), 814-819.

Bamman, D., Eisenstein, J., & Schnoebelen, T. (2012). Gender in Twitter: Styles, Stances, and Social Networks. *Arxiv CoRR*.

Baron, N. S. (1998). Letters by phone or speech by other means: the linguistics of email. *Language & Communication*, *18*.

Beers Fägersten, K. (2007). A Sociolinguistic Analysis of Swearword Offensiveness. *Saarland Working Papers in Linguistics*, *1*, 14–37.

Bell, A. (1984). Language style as audience design. *Language in society*, *13*(02), 145-204.

Bennett, P. N., & Carbonell, J. (2005). Detecting Action-Items in Email. *SIGIR*.

Bernardi, V. (2012). Swearing as an alignment technique. Presented at the Sociolinguistics Symposium 19: Language And The City, Berlin.

Biber, D. (1993). Using Register-Diversified Corpora for General Language Studies. *Association for Computational Linguistics*.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, *3*(1), 993–1022.

Blumer, H. (1986). *Symbolic Interactionism: Perspective and Method*. University of California Press.

Bourhis, R. Y., & Giles, H. (1977). The language of intergroup distinctiveness. *Language, ethnicity, and intergroup relations*, 119-36.

Bousfield, D. & Locher, M. A. (2008). *Impoliteness in Language*. De Gruyter.

Bracewell, D. B., Tomlinson, M. T., & Wang, H. (2012). Identification of Social Acts in Dialogue. *Proceedings of COLING 2012*.

Bradac, J. J., Mulac, A., & House, A. (1988). Lexical diversity and magnitude of convergent versus divergent style shifting: Perceptual and evaluative consequences. *Language & Communication*, *8*(3-4), 213-228.

Bramsen, P., Escobar-Molano, M., Patel, A., & Alonso, R. (2011). Extracting Social Power Relationships from Natural Language. *Proceedings of the Annual Meeting on Association for Computational Linguistics*, 773–782.

Brody, S., & Diakopoulos, N. (2011). Cooooooooooooooollllllllllllll!!!!!!!!!!!!!!: Using Word Lengthening to Detect Sentiment in Microblogs. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, 562–570. Stroudsburg, PA.

Brooke, J., & Hirst, G. (2012). Robust, lexicalized native language identification. In *Proceedings of COLING 2012*.

Brown, P., & Levinson, S. C. (1987). *Politeness: Some universals in language usage*. Cambridge: Cambridge University Press.

Buitkienë, J. (2006). Power Relations and Politeness / Valios raiðkos priemonës ir mandagumas. *Žmogus Ir Žodis*, *8*, 15–20.

Bunz, U., & Campbell, J. (2004). Politeness Accommodation in Electronic Mail. *Communication Research Reports*, *21*(1), 11–25.

Bykh, S., & Meurers, D. (2012). Native Language Identification Using Recurring N-grams – Investigating Abstraction and Domain Dependence. In *Proceedings of COLING 2012*.

Campbell, D. T. (1965). Ethnocentric and other altruistic motives. *Nebraska symposium on motivation*, 13, 283–311.

Carrick, T., Rashid, A., & Taylor, P. J. (2016). Mimicry in online conversations: an exploratory study of linguistic analysis techniques. In *IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining.*

Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception–behavior link and social interaction. *Journal of personality and social psychology*, *76*(6), 893.

Chen, S. F. & Goodman, J. (1999). An empirical study of smoothing techniques for language modeling. In *Computer Speech and Language*, *13*(4), 359–393.

Cheng, C. M., & Chartrand, T. L. (2003). Self-monitoring without awareness: using mimicry as a nonconscious affiliation strategy. *Journal of Personality and Social Psychology, 85*(6), 1170-1179.

Church, K. W. (2000). Empirical estimates of adaptation: The chance of two noriegas is closer to p/2 than p2. In *Proceedings of the 18th Conference on Computational Linguistics (COLING 2000)*, volume 1, pages 180–186.

Chung, C. & Pennebaker, J. W. (2007). The Psychological Functions of Function Words. In *Social Communication*. Psychology Press.

Cooley, C. (1902). *Human Nature and the Social Order*.

Coppersmith, G., Dredze, M., & Harman, C. (2014). Quantifying Mental Health Signals in Twitter.

Core, M., & Allen, J. (1997). Coding dialogs with the DAMSL annotation scheme. *AAAI Fall Symposium on Communicative Action in Humans and Machines*, 28–35.

Cotterill, R. (2011) Question classification for email. *Proceedings of the Ninth International Conference on Computational Semantics*, 330-334. Association for Computational Linguistics.

Cotterill, R. (2013) Using Stylistic Features for Social Power Modeling. *Computación y Sistemas*, 17(2), 219-227.

Cotterill, R., Muir, K., Joinson, A. and Dewdney, N. (2015) Identifying linguistic correlates of power. *International Journal of Computational Linguistics and Applications*, 6(1).

Daly, N., Holmes, J., Newton, J., & Stubbe, M. (2004). Expletives as solidarity signals in FTAs on the factory floor. *Journal of Pragmatics*, *36*(5), 945–964.

Danescu-Niculescu-Mizil, C., Lee, L., Pang, B., & Kleinberg, J. (2012). Echoes of Power: Language Effects and Power Differences in Social Interaction. In *Proceedings of the 21st International Conference on World Wide Web*, 699–708.

Danescu-Niculescu-Mizil, C., Sudhof, M., Leskovec, J., & Potts, C. (2013a). A computational approach to politeness with application to social factors. In *Proceedings of ACL*.

Danescu-Niculescu-Mizil, C., West, R., Jurafsky, Daniel, Leskovec, J., & Potts, C. (2013b). No Country for Old Members: User Lifecycle and Linguistic Change in Online Communities. In *Proceedings of WWW*.

DeForest, M., & Johnson, E. (2001). The density of Latinate words in the speeches of Jane Austen's characters. *Literary and linguistic computing*, *16*(4), 389–401.

Derks, D., Bos, A. E. R., & von Grumbkow, J. (2007). Emoticons and social interaction on the Internet: the importance of social context. *Computers in Human Behaviour*, *23*.

Dewaele, J., & Furnham, A. (1999). Extraversion: The Unloved Variable in Applied Linguistic Research. *Language Learning*, *49*(3), 509–544.

Dewaele, J.-M. (2010). “Christ fucking shit merde!” Language Preferences for Swearing Among Maximally Proficient Multilinguals. *Sociolinguistic Studies*, *4*(3).

Doherty-Sneddon, G., Anderson, A., O'Malley, C., Langton, S., Garrod, S., & Bruce, V. (1997). Face-to-face and video-mediated communication: A comparison of dialogue structure and task performance. *Journal of Experimental Psychology: Applied*, *3*(2), 105.

Dragojevic, M., Gasiorek, J., & Giles, H. (2015). Communication Accommodation Theory. In C. R. Berger & M. L. Roloff (eds.), *Encyclopedia of Interpersonal Communication*. New York: Blackwell/Wiley.

Duthler, K. W. (2006). Politeness of Requests Made via Email and Voicemail: support for the hyperpersonal model. *Journal of Computer-Mediated Communication*, *11*, 500–521.

Eggins, S., & Slade, D. (1997). Chapter 2: Relevant Approaches to Analysing Casual Conversation. In *Analysing Casual Conversation*. London: Cassell.

Eisenstein, J. (2013). What to do about bad language on the internet. *Proceedings of NAACL-HLT 2013*, 359–369.

Fang, A. C., & Cao, J. (2010). A Corpus-Based Quantitative Survey of the Etymological Composition of Contemporary British English: The Case of Latin, Greek and French. *Glottotheory*, *3*(1), 53-68.

Ferrucci, D., & Lally, A. (2004). UIMA: an architectural approach to unstructured information processing in the corporate research environment.*Natural Language Engineering*, *10*(3-4), 327-348.

Flekova, L., Carpenter, J., Giorgi, S., Ungar, L., & Preotiuc-Pietro, D. (2016) Analyzing Biases in Human Perception of User Age and Gender from Text. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics*.

Freeman, L. (1977). A Set of Measures of Centrality Based on Betweenness. *Sociometry*, 40, 35–41.

Freeman, L. C. (1979). Centrality in Social Networks: Conceptual Clarification. *Social Networks*, 1, 215-239.

French, J. R. P., & Raven, B. (1959). The Bases of Social Power. In D. Cartwright (Ed.), *Studies in Social Power*, 150–167. University of Michigan.

Gallagher, I. (2010). Bayesian Block Modelling for Weighted Networks. *Proceedings of the Eighth Workshop on Mining and Learning with Graphs*, 55–61.

Ganter, V., & Strube, M. (2009). Finding Hedges by Chasing Weasels: Hedge Detection Using Wikipedia Tags and Shallow Linguistic Features. *Proceedings of the ACL-IJCNLP 2009 Conference Short Papers*, 173–176.

Garlaschelli, D., & Loffredo, M. I. (2004). Patterns of link reciprocity in directed networks. *Physical Review Letters*, *93*(26).

Gilbert, E. (2012). Phrases That Signal Workplace Hierarchy. In *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work*, 1037–1046. New York, NY, USA: ACM.

Giles, H. (1973). Accent mobility: A model and some data. In *Anthropological linguistics*, *15*(2), 87–105.

Giles, H. (2008). Communication accommodation theory. In L. A. Baxter & D. O. Braithewaite (eds.), *Engaging theories in interpersonal communication: Multiple perspectives,* 161-173. Sage Publications.

Giles, H., Coupland, N., and Coupland, J. (1991). Accommodation theory: Communication, context, and consequence. In Giles, H., Coupland, J., and Coupland, N. (eds.), *Contexts of accommodation: Developments in applied sociolinguistics*, 1–68. New York: Cambridge University Press.

Giles, H., & Gasiorek, J. (2014). Parameters of nonaccommodation: Refining and elaborating communication accommodation theory. In J. P. Forgas, O. Vincze, & J. László (eds.), *Social cognition and communication,* 155-172. New York: Psychology Press.

Giles, H., Taylor, D. M., & Bourhis, R. (1973). Towards a theory of interpersonal accommodation through language: Some Canadian data. *Language in society*, *2*(02), 177-192.

Gill, A., Nowson, S. and Oberlander, J. (2009). What are they blogging about? Personality, topic and motivation in blogs. In *Proceedings of the 3rd International AAAI Conference on Weblogs and Social Media*, 18–25.

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

Goffman, E. (1959). *The Presentation of Self in Everyday Life*. Edinburgh: Anchor Books.

Goffman, E. (1974). *Frame Analysis: An essay on the organization of experience*. London: Harper & Row.

Golcher, F., & Reznicek, M. (2011). Stylometry and the interplay of topic and L1 in the different annotation layers in the FALKO corpus. In *Proceedings of Quantitative Investigations in Theoretical Linguistics 4*.

Gonzales, A., Hancock, J. T., & Pennebaker, J. W. (n.d.). Language style matching as a predictor of social dynamics in small groups.

Gouws, S., Metzler, D., Cai, C., & Hovy, E. (2011). Contextual Bearing on Linguistic Variation in Social Media. *Proceedings of the Workshop on Languages in Social Media*, 20–29.

Granger, S. (1998). The computer learner corpus: A versatile new source of data for SLA research. In In Granger, S. (Ed.) *Learner English on Computer.* Addison Wesley Longman.

Grant, T. (2007). Quantifying evidence in forensic authorship analysis. In *International Journal of Speech, Language & the Law*, *14*(1), 1–25.

Grice, H. P. (1975). *Logic and Conversation*. New York: Academic Press.

Grieve, J. (2007). Quantitative authorship attribution: An evaluation of techniques. *Literary and linguistic computing*, *22*(3), 251-270.

Guéguen, N. (2009). Mimicry and seduction: An evaluation in a courtship context. *Social Influence*, *4*(4), 249-255.

Gunning, R. (1969). The fog index after twenty years. *Journal of Business Communication*, *6*(2), 3–13.

Guthrie, D., Allison, B., Liu, W., Guthrie, L., & Wilks, Y. (2006). A Closer Look at Skip-gram Modelling. *Proceedings of the 5th international Conference on Language Resources and Evaluation.*

Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., & Witten, I. H. (2009). The WEKA data mining software: an update. *ACM SIGKDD explorations newsletter*, *11*(1), 10-18.

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. New York, NY: Guilford Press.

Heylighen, F., & Dewaele, J.-M. (2002). Variation in the contextuality of language: an empirical measure. *Foundations of Science*, *7*, 293–340.

Hill, J. (1995). Junk Spanish, covert racism, and the (leaky) boundary between public and private spheres. *Pragmatics*, *5*(2), 197-212.

Hogg, M. A., & Reid, S. A. (2006). Social identity, self-categorization, and the communciation of group norms. *Communication Theory*, *16*(1), 7–30.

Hogg, M. A., Terry, D. J., & White, K. M. (1995). A Tale of Two Theories: A Critical Comparison of Identity Theory with Social Identity Theory. *Social Psychology Quarterly*, *58*(4), 255–269.

Holmes, D. I. (1994). Authorship attribution. *Computers and the Humanities*, *28*(2), 87-106.

Huffaker, D. A., Swaab, R., & Diermeier, D. (2011). The language of coalition formation in online multiparty negotiations. *Journal of Language and Social Psychology*, *30*(1), 66-81.

Iacobelli, F., Gill, A. J., Nowson, S., & Oberlander, J. (2011). Large Scale Personality Classification of Bloggers. In S. D'Mello, A. Graesser, B. Schuller, & J.-C. Martin (Eds.), *Affective Computing and Intelligent Interaction*, 568–577. Springer Berlin Heidelberg.

Ireland, M. E., Slatcher, R. B., Eastwick, P. W., Scissors, L. E., Finkel, E. J., & Pennebaker, J. W. (2011). Language Style Matching Predicts Relationship Initiation and Stability. *Psychological Science*, *22*(1), 39–44.

Jabbari, S., Allison, B., Guthrie, D., & Guthrie, L. (2006). Towards the Orwellian Nightmare: Separation of business and personal emails. *Proceedings of the COLING/ACL*, 407–411.

Jacob, C., Guéguen, N., Martin, A., & Boulbry, G. (2011). Retail salespeople's mimicry of customers: Effects on consumer behavior. *Journal of Retailing and Consumer Services*, *18*(5), 381-388.

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

Jones, E., Gallois, C., Callan, V., & Barker, M. (1999). Strategies of accommodation: Development of a coding system for conversational interaction. *Journal of Language and Social Psychology*, *18*(2), 123-151.

Jurafsky, D., Shriberg, E., & Biasca, D. (1997). Switchboard SWBD-DAMSL Shallow-Discourse-Function Annotation Coders Manual, Draft 13.

University of Colorado, Boulder Institute of Cognitive Science Technical Report 97-02.

Kacewicz, E., Pennebaker, J. W., Davis, M., Jeon, M., & Graesser, A. C. (2013). Pronoun Use Reflects Standings in Social Hierarchies. *Journal of Language and Social Psychology*.

Kalman, Y. M., Ravid, G., Raban, D. R., & Rafaeli, S. (2006). Pauses and Response Latencies: A Chronemic Analysis of Asynchronous CMC. *Journal of Computer-Mediated Communication*, *12*(1).

Kankaanranta, A. (2005). *Hej Seppo, could you pls comment on this!" – Internal Email Communication in Lingua Franca English in a Multinational Company*. University of Jyväskylä.

Kano, Y., Dorado, R., McCrohon, L., Ananiadou, S., & Jun'ichi Tsujii. (2010). U-Compare: An Integrated Language Resource Evaluation Platform Including a Comprehensive UIMA Resource Library. In *Proceedings of LREC 2010*.

Karagiannis, T., & Vojnovic, M. (2009). Behavioral profiles for advanced email features. *Proceedings of the 18th International Conference on World Wide Web*, 711–720.

Kincaid, J. P., Fishburne Jr, R. P., Rogers, R. L., & Chissom, B. S. (1975). *Derivation of new readability formulas (automated readability index, fog count and flesch reading ease formula) for navy enlisted personnel* (No. RBR-8-75). Naval Technical Training Command Millington TN Research Branch.

Klimt, B., & Yang, Y. (2004). The enron corpus: A new dataset for email classification research. *ECML*, 217–226.

Karlgren, J. (2010). Textual Stylistic Variation: Choices, Genres and Individuals. In S. Argamon, K. Burns, & S. Dubnov (Eds.), *The Structure of Style*, 113–125. Springer Berlin Heidelberg.

Kasper, G., & Rose, K. R. (2002). *Pragmatic Development in a Second Language*. Blackwell.

Kochol, M. (2004). Tension-flow polynomials on graphs. *Discrete Mathematics*, *274*(1–3), 173–185.

Koppel, M., Argamon, S., & Shimoni, A. R. (2002). Automatically Categorizing Written Texts by Author Gender. *Literary and Linguistic Computing*.

Koppel, M., Schler, J., & Zigdon, K. (2005). Automatically Determining an Anonymous Author's Native Language. In *Intelligence and Security Informatics*. Berlin: Springer.

Koppel, M., Schler, J., & Argamon, S. (2009). Computational methods in authorship attribution. *Journal of the American Society for information Science and Technology*, *60*(1), 9-26.

Kurzius, E. (2015) The Extraverted Chameleon: Personality's effects on mimicry of verbal behaviour. *Journal of Individual Differences*, *36*(2), 80-86.

Ladd, D. R. (1981). A First Look at the Semantics and Pragmatics of Negative Questions and Tag Questions. *Papers from the Regional Meeting. Chicago Ling. Soc. Chicago, Ill.*, *17*, 164–171.

Lahiri, S., Mitra, P., & Lu, X. (2011). Informality Judgements at Sentence Level and Experiments with Formality Score. *CICLING 2011*, 446–457.

Lakin, J. L., & Chartrand, T. L. (2003). Using nonconscious behavioral mimicry to create affiliation and rapport. *Psychological Science, 14*(4), 334-339.

Lakoff, G. (1973). Hedges: A Study in Meaning Criteria and the Logic of Fuzzy Concepts. *Journal of Philosophical Logic*, *2*, 458–508.

Lam, W. S. E. (2008). Language Socialization in Online Communities. In N. H. Hornberger (Ed.), *Encyclopedia of Language and Education*, 2859–2869. Springer US.

Lampert, A., Dale, R., & Paris, C. (2006). Classifying Speech Acts using Verbal Response Modes. *Proceedings of the 2006 Australasian Language Technology Workshop*.

Lampert, A., Dale, R., & Paris, C. (2009). Segmenting Email Message Text into Zones. In *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing: Volume 2 - Volume 2*, 919–928. Stroudsburg, PA, USA: Association for Computational Linguistics.

Layton, R., Watters, P., & Dazeley, R. (2011). Automated unsupervised authorship analysis using evidence accumulation clustering. *Natural Language Engineering.*

Liu, F., Weng, F., & Jiang, X. (2012). A Broad-coverage Normalization System for Social Media Language. In *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Long Papers - Volume 1*, 1035–1044. Stroudsburg, PA, USA.

Locher, M. A. (2004). *Power and Politeness in Action.* De Gruyter.

Lutosławski, W. (1897). *The origin and growth of Plato's logic, with an account of Plato's style and of the chronology of his writings*. London: Longmans, Green and co.

McEnery, A. M., & Xiao, R. Z. (2004). Swearing in modern British English: the case of fuck in the BNC. *Language and Literature*, *13*(3), 235–268.

Manning, C. D., Surdeanu, M., Bauer, J., Finkel, J. R., Bethard, S., & McClosky, D. (2014). The Stanford CoreNLP Natural Language Processing Toolkit. In *ACL (System Demonstrations),* 55-60.

Mead, G. H. (1934). *Self and Society*. Chicago: University of Chicago.

Mehl, M. R., Gosling, S. D., & Pennebaker, J. W. (2006). Personality in its natural habitat: Manifestations and implicit folk theories of personality in daily life. *Journal of Personality and Social Psychology*, *90*(5), 862–877. http://doi.org/10.1037/0022-3514.90.5.862

Moore, E., & Podesva, R. (2009). Style, Indexicality and the Social Meaning of Tag Questions. *Language in Society*, *38*(4), 447–485.

Morand, D. A. (2000). Language and power: An empirical analysis of linguistic strategies used in superior–subordinate communication. *Journal of Organizational Behavior*, *21*(3), 235–248.

Morand, D. A., & Ocker, R. J. (2003). Politeness Theory and Computer-Mediated Communication: A Sociolinguistic Approach to Analyzing Relational Messages. *Proceedings of the 36th Hawaii International Conference on System Sciences*.

Mosteller, F., & Wallace, D. L. (1963). Inference in an Authorship Problem. *Journal of the American Statistical Association*, *58*(302), 275–309.

Muir, K., Cotterill, R., Joinson, A. and Dewdney, N. (2015) Power and personality in linguistic style accommodation. *British Psychological Society Developmental and Social Section Annual Conference*, Manchester, UK, 9 - 11 September 2015.

Muir, K., Joinson, A., Cotterill, R. and Dewdney, N. (2016a) Characterising the linguistic chameleon: Personal and social correlates of linguistic style accommodation. *Human Communication Research*, 42(3).

Muir, K., Joinson, A., Cotterill, R. and Dewdney, N. (2016b) When Communication Accommodation Backfires: Interpersonal Effects of Social Power and Linguistic Style Accommodation in Computer-Mediated-Communication. *Proceedings of the 66th Annual Conference of the International Communication Association.*

Muir, K., Joinson, A., Cotterill, R. and Dewdney, N. (2016c) English Speed Networking Conversational Transcripts LDC2016T16. *Linguistic Data Consortium.*

Munro, R., & Manning, C. D. (2012). Short Message Communications: Users, Topics, and In-language Processing. In *Proceedings of the 2Nd ACM Symposium on Computing for Development*. New York, NY, USA: ACM.

Murphy, M., & Levy, M. (2006). Politeness in Intercultural Email Communication: Australian and Korean Perspectives. *Journal of Intercultural Communication*, *12*(1).

Myslín, M., & Gries, S. T. (2010). k dixez? A corpus study of Spanish Internet orthography. *Literary and Linguistic Computing*, *25*(1), 85–104.

MySpell Dictionaries. (2008). https://code.google.com/archive/a/apache-extras.org/p/ooo-myspell

Natale, M. (1975). Social desirability as related to convergence of temporal speech patterns. *Perceptual and Motor Skills, 40*(3), 827-830.

Newman, M. E. (2001). The structure of scientific collaboration networks. *Proceedings of the National Academy of Sciences*, *98*(2), 404–409.

Newman, M. L., Groom, C. J., Handelman, L. D., & Pennebaker, J. W. (2008). Gender Differences in Language Use: An Analysis of 14,000 Text Samples. *Discourse Processes*, *45*(3), 211–236.

Nguyen, D., & Rosé, C. (2011). Language use as a reflection of socialization in online communities. *Proceedings of the Workshop on Languages in Social Media*, 76–85.

Nguyen, D., Trieschnigg, D. & Meder, T. (2014) TweetGenie: Development, Evaluation, and Lessons Learned. In *Proceedings of COLING 2014.*

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

Nowson, S. and Oberlander, J. (2007) Identifying more bloggers: Towards large scale personality classification of personal weblogs. In *Proceedings of the International Conference on Weblogs and Social Media (ICWSM 2007)*.

Oberlander, J., & Gill, A. J. (2006). Language With Character: A Stratified Corpus Comparison of Individual Differences in E-Mail Communication. *Discourse Processes*, *42*(3), 239–270. http://doi.org/10.1207/s15326950dp4203_1

Onorato, R. S., & Turner, J. C. (2004). Fluidity in the self-concept: the shift from personal to social identity. *European Journal of Social Psychology*, *34*, 257–278.

OpenNLP. (2011). *http://opennlp.apache.org*. Apache software foundation.

Panteli, N. (2002). Richness, power cues and email text. *Information and Management*, *40*, 75–86.

Pennebaker, J. W., Booth, R. J., & Francis, M. E. (2007). Linguistic Inquiry and Word Count (LIWC2007). *http://homepage.psy.utexas.edu/homepage/faculty/pennebaker/reprints/LIWC2007_OperatorManual.pdf*. Retrieved from https://researchspace.auckland.ac.nz/handle/2292/18109

Pennebaker, J. W., & King, L. A. (1999). Linguistic styles: Language use as an individual difference. *Journal of Personality and Social Psychology*, *77*(6), 1296–1312. http://doi.org/10.1037/0022-3514.77.6.1296

Pennebaker, J. W. & Stone, L. D. (2003). Words of wisdom: Language use over the life span. *Journal of Personality and Social Psychology*, *85*(2), 291–301. http://doi.org/10.1037/0022-3514.85.2.291

Pérez Sabater, C., Turney, E., & Montero Fleta, B. (2008). Orality and literacy, formality and informality in email communication. *Ibérica*, *15*.

Peterson, K., Hohensee, M., & Xia, F. (2011). Email Formality in the Workplace: A Case Study on the Enron Corpus. *Proceedings of the Workshop on Languages in Social Media*, 86–95.

Prabhakaran, V., Neralwala, H., Rambow, O., & Diab, M. (2013). Annotations for Power Relations on Email Threads. Retrieved from http://lrec.elra.info/proceedings/lrec2012/pdf/1006_Paper.pdf

Prabhakaran, V., & Rambow, O. (2013). Written dialog and social power: Manifestations of different types of power in dialog behavior. In *Proceedings of the IJCNLP*, 216–224.

Prabhakaran, V., & Rambow, O. (2014). Predicting Power Relations between Participants in Written Dialog from a Single Thread. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics*.

Rayson, P., Leech, G., & Hodges, M. (1997). Social differentiation in the use of English vocabulary: some analyses of the conversational component of the British National Corpus. In *International Journal of Corpus Linguistics*, *2*(1), 133–152.

Richardson, B. H., Taylor, P. J., Snook, B., Conchie, S. M., & Bennell, C. (2014). Language style matching and police interrogation outcomes. *Law and Human Behaviour*, *38*(4), 357–366.

Riordan, M. A., Markman, K. M., & Stewart, C. O. (2013). Communication accommodation in instant messaging an examination of temporal convergence. *Journal of Language and Social Psychology*, *32*(1), 84-95.

Rosenthal, S., & McKeown, K. (2011). Age Prediction in Blogs: A Study of Style, Content, and Online Behaviour in Pre- and Post-Social Media Generations. *Proceedings of the 49th Annual Meeting on Association for Computational Linguistics*, 763–772.

Rowe, R., Creamer, G., Hershkop, S., & Stolfo, S. J. (2007). Automated social hierarchy detection through email network analysis. *Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 Workshop on Web Mining and Social Network Analysis*, 109–117.

Sabidussi, G. (1966). "The centrality index of a graph". *Psychometrika*, 31, 581–603.

Santini, M. (2007). Characterizing Genres of Web Pages: Genre Hybridism and Individualization. In *Proceedings of the 40th Annual Hawaii International Conference on System Sciences (HICSS'07)*.

Sap, M., Park, G., Eichstaedt, J. C., Kern, M. L., Stillwell, D., Kosinski, M., Ungar, L. H. & Schwartz, H. A. (2014) Developing Age and Gender Predictive Lexica over Social Media. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*.

Schnoebelen, T. (2012a). Do You Smile with Your Nose? Stylistic Variation in Twitter Emoticons. *University of Pensylvannia Working Papers in Linguistics*, *18*(2).

Schnoebelen, T. (2012b). *Emotions are Relational: Positioning and the use of affective linguistic resources*. Stanford University.

Searle, J. R. (1969). *Speech acts: An essay in the philosophy of language*. Cambridge.

Setlock, L. D., Fussell, S. R., & Neuwirth, C. (2004). Taking it out of context: collaborating within and across cultures in face-to-face settings and via instant messaging. In *Proceedings of the 2004 ACM conference on Computer supported cooperative work*, 604-613.

Sexton, J., & Helmreich, R. (2000). Analyzing Cockpit Communications: The Links Between Language, Performance, Error, and Workload. *Journal of Human Performance in Extreme Environments*, *5*(1).

Shetty, J., & Adibi, J. (2004). The Enron Email Dataset Database: Schema and Brief Statistical Report. *Information Sciences Institute Technical Report, University of Southern California*.

Skovholt, K., & Svennevig, J. (2006). Email Copies in Workplace Interaction. *Journal of Computer-Mediated Communication*, *12*(1), 42–65.

Soliz, J., & Giles, H. (2014). Relational and Identity Processes in Communication: A Contextual and Meta-Analytical Review of Communication Accommodation

Theory. In E. Cohen (ed.), *Communication Yearbook 38*, 106-143. Sage Publications.

Sood, S. O., Antin, J., & Churchill, E. F. (2012). Profanity Use in Online Communities. In *Proceedings of CHI'12*.

Sotillo, S. M. (2000). Discourse functions and syntactic complexity in synchronous and asynchronous communication. *Language Learning & Technology*, *4*(1), 82-119.

Stamatatos, E., Fakotakis, N., & Kokkinakis, G. (2001). Computer-based authorship attribution without lexical measures. *Computers and the Humanities*, *35*(2), 193-214.

Stamatatos, E. (2009). A survey of modern authorship attribution methods. *Journal of the American Society for information Science and Technology*, *60*(3), 538–556.

Stenchikova, S. & Stent, A. (2007). Measuring adaptation between dialogs. In *Proceedings of the 8th SIGdial workshop on Discourse and Dialogue*, 166–173.

Tajfel, H., & Turner, J. (1979). An Integrative Theory of Intergroup Conflict. *The Social Psychology of Intergroup Relations*, *33*, 33–47.

Tang, J., Li, H., Cao, Y., & Tang, Z. (2005). Email Data Cleaning. *Proceedings of KDD'05*.

Tannen, D. (1984). *Conversational style: Analyzing talk among friends*. Oxford University Press.

Tausczik, Y. R., & Pennebaker, J. W. (2010). The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods. *Journal of Language and Social Psychology*, *29*(1), 24–54.

Taylor, P. J., & Thomas, S. (2008). Linguistic Style Matching and Negotiation Outcome. *Negotiation and Conflict Management Research, 1*(3), 263 - 281.

Taylor, D. M., Simard, L. M., & Papineau, D. (1978). Perceptions of cultural differences and language use: A field study in a bilingual environment. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement*, *10*(3), 181.

Teddiman, L. (2009). Contextuality and Beyond: Investigating an Online Diary Corpus. *Proceedings of ICWSM*.

Tsur, O., & Rappoport, A. (2007). Using Classifier Features for Studying the Effect of Native Language on the Choice of Written Second Language Words. In *ACL 2007 Workshop on Cognitive Aspects of Computational Language Acquisition*.

Turner, J. C., Hogg, M. A., Oakes, P. J., Reicher, S. D., & Wetherell, M. S. (1987). *Rediscovering the social group: A self-categorization theory*. Oxford: Blackwell.

Turner, J. C., Oakes, P. J., Haslam, S. A., & McGarty, C. (1994). Self and Collective: Cognition and Social Context. *Personality and Social Psychology Bulletin*, *20*, 454–463.

Turell, M. T. (2010). The use of textual, grammatical and sociolinguistic evidence in forensic text comparison. *International Journal of Speech, Language & the Law*, *17*(2).

Van Baaren, R. B., Holland, R. W., Kawakami, K., & van Knippenberg, A. (2004). Mimicry and prosocial behaviour. *Psychological Science*, *15*(1), 71–74.

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

Van der Wees, M., Bisazza, A., Weerkamp, W., & Monz, C. (2015). What's in a Domain? Analyzing Genre and Topic Differences in Statistical Machine Translation. In *Proceedings of the 53rd Annual Meeting of the Association of Computational Linguistics*.

Van den Berg, M. E. (1986). Language planning and language use in Taiwan: Social identity, language accommodation, and language choice behavior. *International journal of the sociology of language*, *59*, 97–116.

Varnhagen, C. K., McFall, G. P., Pugh, N., Routledge, L., Sumida-MacDonald, H., & Kwong, T. E. (2010). lol: new language and spelling in instant messaging. *Reading and Writing*, *23*(6), 719–733. http://doi.org/10.1007/s11145-009-9181-y

Waldvogel, J. (2007). Greetings and closings in workplace email. *Journal of Computer-Mediated Communication*, *12*(2), 456–477.

Walther, J. B., & D'Addario, K. P. (2001). The Impacts of Emoticons on Message Interpretation in Computer-Mediated Communication. *Social Science Computer Review*, *19*(3), 324–347.

Wasserman, S., & Faust, K. (1994). *Social Network Analysis: Methods and Applications*, Vol. 8. Cambridge University Press.

Waters, R. (2013). Flows and Tensions (private correspondence).

Watts, R. J. (2005). Linguistic Politeness Research: Quo Vadis? In R. J. Watts, S. Ide, & K. Ehlich (Eds.) *Politeness in Language: Studies in its History, Theory and Practice*.

Weaver, W. (1955). Translation. In *Machine Translation of Languages*, *14*, 15-23.

Welkowitz, J., & Feldstein, S. (1969). Dyadic interaction and induced differences in perceived similarity. *Proceedings of the Annual Convention of the American Psychological Association, 4*(1), 343-344.

Willemyns, M., Gallois, C., Callan, V. J., & Pittam, J. (1997). Accent accommodation in the job interview impact of interviewer accent and gender. *Journal of Language and Social Psychology*, *16*(1), 3-22.

Wilson, N. (2012). Swearing as a leadership tool. Presented at the Sociolinguistics Symposium 19: Language And The City, Berlin.

Wong, S. J., & Dras, M. (2011). Exploiting parse structures for native language identification. In *EMNLP '11 Proceedings of the Conference on Empirical Methods in Natural Language Processing*.

Wright, D., & Gomez, A. (2012). Forensic authorship identification in globalised digital communities of practice: the theoretical and methodological implications of a sociolinguistic approach to analysing authorship in Enron emails. In *Sociolinguistics Symposium 19*. Berlin.

Xiang, G., Bin, F., Ling, W., Hong, J. I., & Rosé, C. (2012). Detecting Offensive Tweets via Topical Feature Discovery over a Large Scale Twitter Corpus. In *Proceedings of CIKM'12*. Maui.

Zachary, W. W. (1977). An information flow model for conflict and fission in small groups. *Journal of Anthropological Research*, 452–473.

Zheng, R., Qin, Y., Huang, Z., & Chen, H. (2003). Authorship analysis in cybercrime investigation. In *International Conference on Intelligence and Security Informatics*, 59-73.