Temporal Analysis of Stages in Negotiation

by
Reshma P. Khilnani

Submitted to the Department of Electrical Engineering and Computer Science
in Partial Fulfillment of the Requirements for the Degrees of
Bachelor of Science in Electrical [Computer] Science and Engineering
and Master of Engineering in Electrical Engineering and Computer Science
at the Massachusetts Institute of Technology
August 16, 2004

Author ____
Department of Electrical Engineering and Computer Science
August 17, 2004

Certified by ________
Alex (Sandy) Pentland
Thesis Supervisor

Accepted by ____________________
Arthur C. Smith
Chairman, Department Committee on Graduate Theses

Temporal Analysis of Stages in Negotiation

by
Reshma Khilnani
reshma_k@mit.edu
Submitted to the
Department of Electrical Engineering and Computer Science
August 16, 2004

Abstract

This thesis investigates the problem of computer voice analysis as applied to the negotiation setting. In our study we collected audio recordings of negotiations in a controlled experiment, outlined distinct behaviors within the negotiations, and built a program that can automatically annotate a recorded negotiation. In addition, we analyzed patterns of behaviors within the samples we collected. We found seven distinct behaviors present in the voice signal, and found that the role of the participant in the negotiation affects behavior to some degree.
Thesis Supervisor: Alex (Sandy) Pentland
Title: Toshiba Professor of Media Arts & Science

Contents

1 Introduction

The negotiation is the primary method by which progress is made and consensus is reached on a multi-person level. It is defined as a discussion intended to produce an agreement, and manifests itself every day as part of organizational behavior, as part of personal life, and between hostile parties. However, negotiation is just as mysterious as it is common. Humans use common sense, and heuristic techniques to navigate negotiations, but quantitative analysis of negotiations has yet to be explored.

1.1 Purpose

The purpose of this investigation at the highest level is to understand negotiations using recordings of the participants' voices throughout the negotiation. What stages does the negotiation go through? How does behavior of participants depend on their role? What behaviors allow negotiators to make progress on their position?

At a more specific level the purpose is to use the tools available in voice analysis, such as stress indicators, turn-taking parameters, and influence [6] to understand the meaning of the subjective stages that the human brain goes through when observing such a discourse.

Understanding negotiations is the first step toward influencing them, leveraging the techniques that humans have learned over time and the information gained from voice analysis to steer and manage the dynamics of negotiation.

1.2 Problem

When analyzing human behavior using mathematical models there is always a marriage of practical knowledge, or common sense knowledge, and a mathematical structure. Once the mathematical structure and practical knowledge have been reconciled in a way that seems appropriate, the model can be probed for information that provides insight into the real system it mimics.

The study of negotiation described in this thesis is no different. Here, the human behavior of negotiation is being modeled, a marriage of subjective ideas on negotiation with mathematical structure. This means outlining what negotiation experts see as the stages and behaviors within a negotiation, combining those with parameters from voice analysis, and finally with a mathematical model.

The model should answer the following questions:

- What behaviors do individuals exhibit during negotiation? When? In what order?
- How do individual's behaviors influence others' behavior within the negotiation setting?
- Which behaviors correspond to participants' roles in the negotiation?

The problem of building a model that answers the questions above is to merge common sense and mathematical models carefully, in a way that reflects what is known thus far in both domains.

1.3 Relevance and Importance of Problem

One way of looking at the world is to consider each individual as having a perspective, or simply a set of state about the world. Each individual also has a set of actions that he/she can use to add to the state of the world, and methods of communication between people cause different mixtures action and state change. For example, a teacher educating a student can cause a change in state, but typically does not prompt much action; the techniques that the student learns are stored and tested, but sometimes never applied. Another way to communicate is advertising, where advertisers try and change state in the mind to favor the product as well as invoke action to purchase that product. Smalltalk with an aquaintance typically does not amount to much of a change in state, and not much action either. Negotiation falls into the category where both parties have strong ideas and their change in state is small, but participants are attempting to change the action sets of eachother.

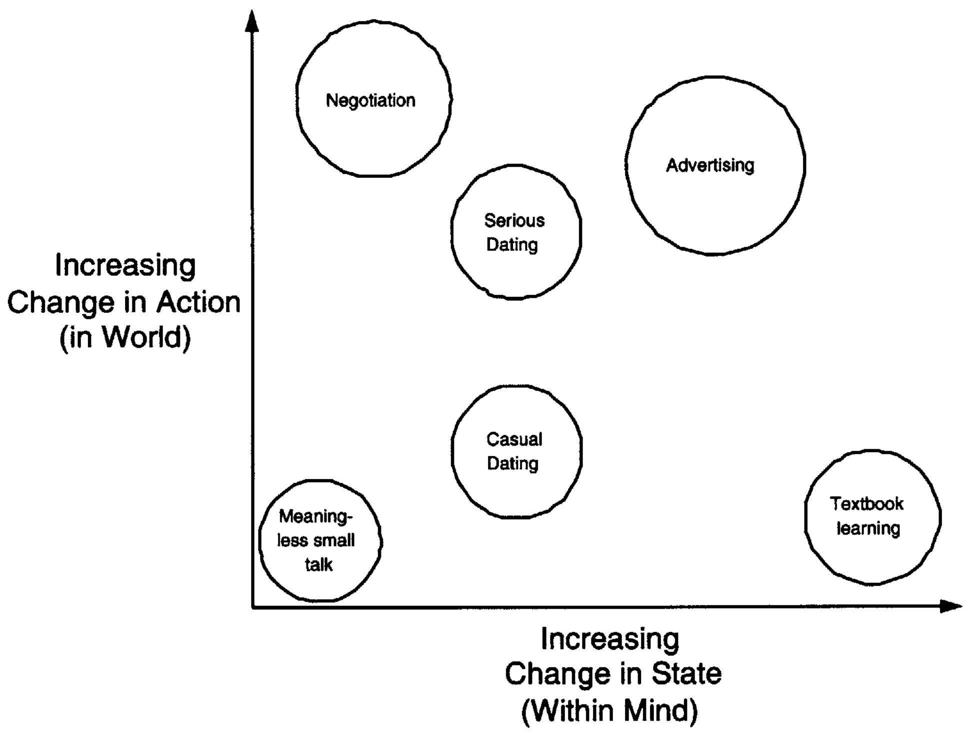

Figure 1: Some areas on the Δ State vs. Δ Action plane

The dichotomy of state change versus action change suggests that these axes are orthogonal, and common knowledge supports this hypothesis. For example, the quality of advertizing is not necessarily correlated with the quality of a product, whereas in teaching, content is very important because there is little action component. When applied to negotiation it follows that the quality of negotiation is not strongly linked to the content of the negotiation, and so one would suppose that there are aspects to the process of negotiation itself that can be studied fruitfully independent of the content of the negotiation. Intuition seems to coordinate with this hypothesis. Commonly in negotiations, parties come with a fairly rigid view of what they should and can get, and leave with a set of agreed on actions that are either in accordance with their beliefs or make them feel cheated.

In the interest of learning how humans might best argue their position it makes sense to study negotiation. As all parties in a negotiation increase their savvy in the relevant techniques, the content component becomes more significant with regard to the outcome of the negotiation. This point is not lost on professionals in the business and academic world. Harvard Law School has a Program on Negotiation, [1] police officers learn negotiation techniques as part of the their training, and business people publish and read volumes on negotiation from a number of perspectives. Even historical figures such as Machiavelli can be interepreted as applied to negotiation.

However, the analysis to date has been primarily by humans who have experience with negotiations, and explain methods they employ to educate others on how to negotiate. The approach employed by this thesis uses automated audio processing to study negotiations. This is important for two reasons. First, acess to experts is often difficult and an automated methods could provide access and ubiquity that is impossible with human experts. Second, machine analysis provides a different perspective on a negotiation than a human one. Humans minds have many specialized and varied functions that can process voice signals at a high level, speech processing does not have these functions, and so has the potential to do processing at a more fundamental level.

Because machine analysis has properties different from those of human interpretation, and negotiations are interesting phenomena, automated study of negotiation is an interesting and worthwhile endeavor.

1.4 General Approach of Study

One categorization technique professors use to classify theses is the "synthesis thesis" and the "novel thesis." In theses of the novel variety, a new process, object, or idea is explored. In theses of the synthesis variety many elements are brought together in a meaninful way. This thesis is of the synthesis variety. It attempts to bring together elements from multiple diciplines and use these elements to understand negotiations, as well as understand our capabilities to study negotiations from the point of view of voice analysis.

Our general approach is to collect recorded signals from negotiations in a controlled environment. Using these signals we look for ways to extract information relevant to a negotiation (for example, who's asking questions of whom?). Once we find a set of features that satisfies our criteria for extracting the relevant elements of a voice signal, we look at how those features vary over time and vary with the roles of participants.

[1] www.pon.harvard.edu

1.5 Criteria for Success

Our analysis here rests on several assumptions. The first, and most fundamental is that it is possible to extract things like who is influencing who, or whether the particpants are copacetic from a simple voice signal as collected by a microphone. A similar study, showing that a large amount of emotional content is present in the voice has been done, but nothing about business or negotiation dynamics [8].

Subsidiary assumptions are that humans exhibit micro-behaviors on the order of minutes during conversation. An example of such a behavior is a person explaining to another, and the other listening. Some guides on negotiation refer to these behaviors when giving advice on negotiation [5], and from our own experiences we suspect they exist.

A final assumption is that when humans listen to a recorded negotiation, they can recreate in their minds what was going on at the time. This is crucial to our analysis, and though we have attempted this recreation, and feel fairly confident in it, it is difficult to evaluate ourselves. We refer to the audio clips throught this thesis as "the ground truth" because we assume we can recreate the scenarios in our mind using those clips.

2 Theory

2.1 The Ground Truth

At this point in time, negotiation is more of an art than a science. Many academics and professionals have spent time investigating patterns and behaviors that lead to success in negotiations, and in parallel research on speech analysis computer scientists are investigating ways of extracting emotional/mood expressors from speech. This thesis mixes these investigations to explore negotiation. The most important part of this study is that there is a constant back-and-forth between the sociological studies and the speech analysis that can be achieved by listening to the recording of the speech signal while looking at the agglomerated data.

2.2 Extracting Relevant Information from Voice Signal

A major component of this analysis is extracting features from the recordings of the negotiation. There are a variety of features that can be chosen. The classic work in psycholinguistics requires consideration: in 1971 Scherer et al. showed that pitch, amplitude, and rate of articulation were sufficient for human listeners to judge the emotional content of speech [8]. Some more behavioral indicators, such as the rate of speaking, the pauses in a person's speech, and the rate of back and forth between the two participants also are significant indicators within a conversation. There are well-established ties between physical conditions such as stress and indicators in the voice.

The choice of features described in this thesis are based on voice analysis techniques developed by others, a preliminary study done on this data set, and advice from individuals knowledgeable about both negotiation and speech processing.

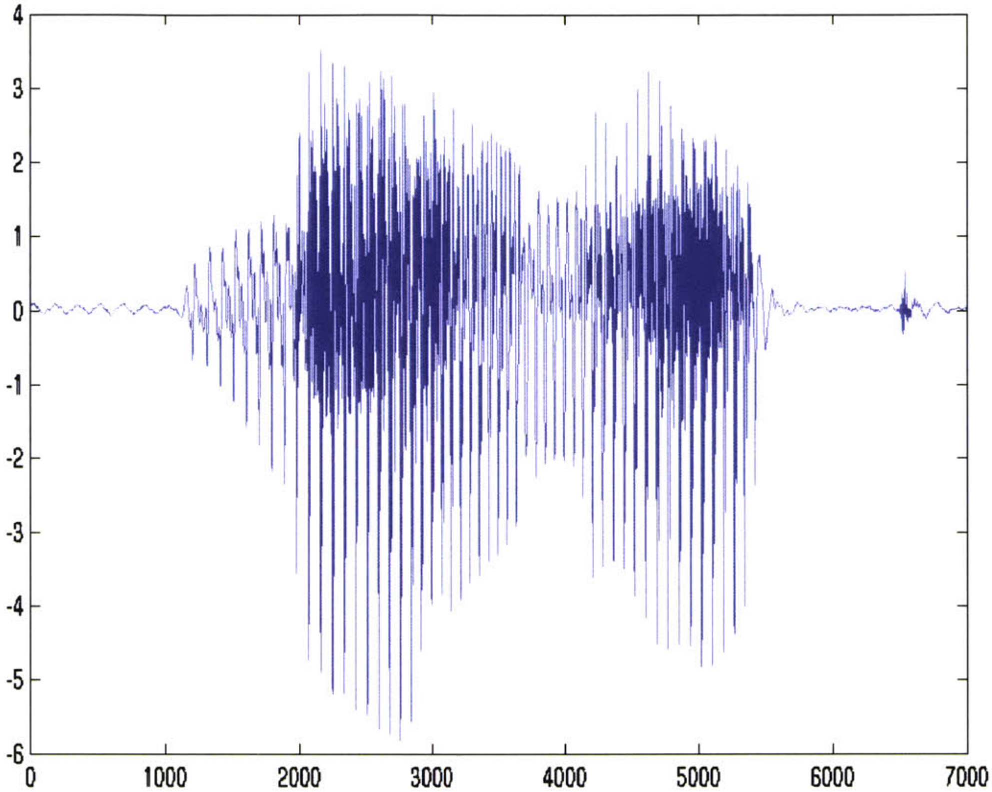

Figure 2: A plot of almost a second of human voice, sampled at 8000Hz

One axis that is important to think about is who's talking? And how much? Informally, this might be a measure of who is holding the floor. A low value of talking time might mean a participant could be listening to an explanation or relaxing. A medium value of talking time might indicate a balanced exchange, questions asked and answered. A high value might indicate more of a monologue, telling, and pushy behavior.

Another axis to think about is the interaction between the two participants. Are they having a friendly interaction? Is one person pushy? Is the interaction rushed? Are the participants having trouble conversing, or making conversation? This seems like a difficult thing to measure, and indeed it is. In this thesis we use an approximation. This approximation is obtained using the influence model [1].

The influence model is a tractable way to solve complex graphical models with many interdependencies. Consider modeling a negotiation as a Hidden Markov Process, the hidden variable being H_i a binary variable. If $H_i = 1$ then the person is the influensor, the participant who is affecting the other. If $H_i = 0$ then the person in not influencing the other participant. Whether a person is the influensor affects another random variable V_i, whether the sounds a person is making are voiced, unvoiced, or silence, see figure 3.

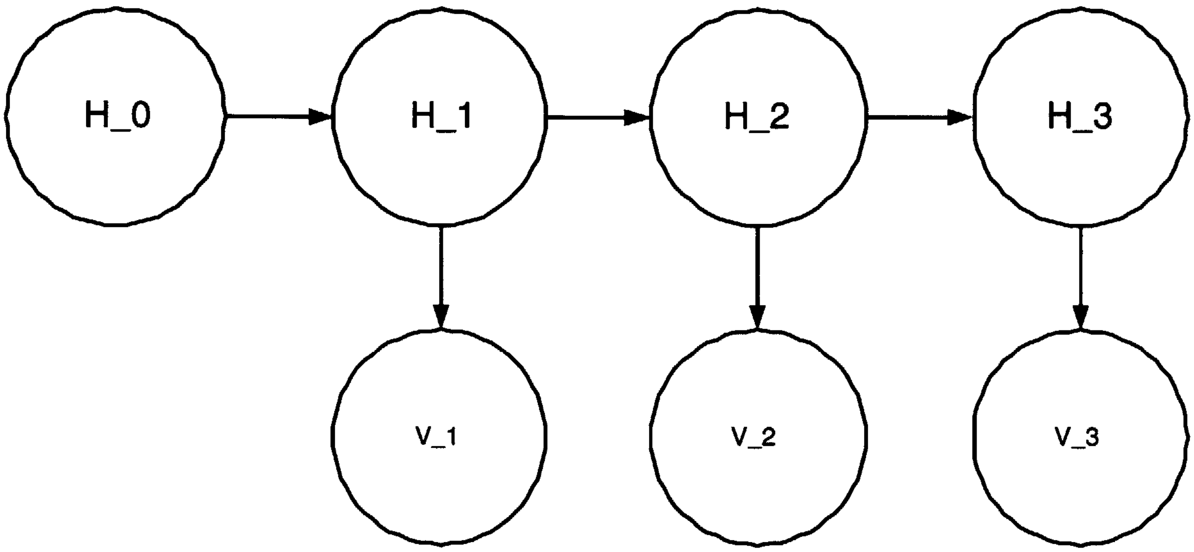

Figure 3: A diagram of the Markov Chain representing a negotiation from one participant's perspective, a part of the influence model as applied to the negotiation analysis. $H_i = 1$ means that the participant is the influencer. $V_i = 1$ means that the time period is voiced.

The H_i random variables, however, do not live in isolation. They are influenced by (among other things) the H_i of the other participant, lets call them G_i. In fact, two Markov chains of the type diagrammed in 3 are coupled in the negotiation scenario. A diagram showing the influences on H_i is shown in 4. The influence model provides a way of estimating all influences on the nodes H_i by a single parameter, shown in figure 4.

The influence factor thus tells a story about the negotiation it is calculated for. When a person is in the position of influence are they likely to hold it for a number of time steps? Does the influence get bounced back and forth? Are neither parties influential, deadlocking the negotiation? The use of voiced and unvoiced snippets as indicators for the influence dynamics in a negotiation was pioneered by Pentland, et al. in [6] [3].

The formant frequencies of a signal is the major frequency component of said signal, or pitch of voice. Though this has been debated, observation of behavior suggest that there is a possible correlation between variation in formant frequency and attention or boredom. For example, if a person says many syllables with the same formant frequency it may correlate to their being bored or babbling. If the formant frequency varies a lot then a person is engaged, making a point. Using this observation as a basis, we include a measure of the variation of formant frequency as part of our analysis.

The pace of a negotiations dialogue and friendly dynamics are both important parts of the

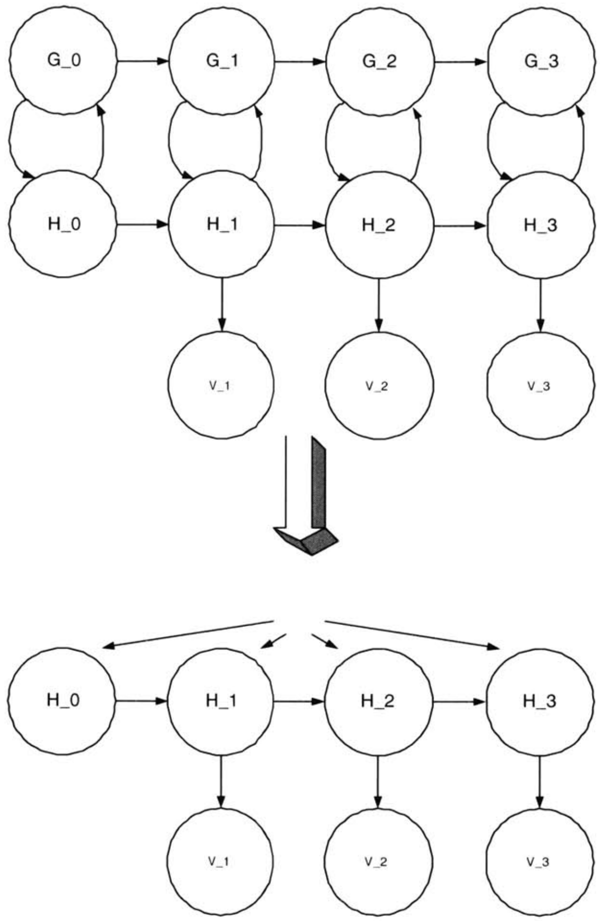

Figure 4: The role of the influence model in this setting is to estimate the influence of G_i on H_i and encapsulate it in a single parameter α

negotiation process [5]. In fact putting undue pressure on time is one of the hallmark pitfalls of amateurs. It makes sense then to have a measure of the speed at which the negotiation is going. Are there many pauses and silences? This is also an indicator of an unfriendly environment which is negatively correlated with level of satisfaction of both parties with the result [7]. It makes sense that our analysis would take pace into account. For this reason we use a measure of changes in signal energy level to get an idea of the position and duration of energetic exchange within a negotiation.

One final consideration that we make regards a behavior called mirroring. Mirroring is a conversational phenomena that humans take part in. Consider the following dialogue:

1 Alice: So you agree that we should follow plan A.
2 Bob: Sure.
3 Alice: Great.

Table 1: Some basic facts about the features we analyze

What We Want to Find Out	How We Measure It
Who's Holding the floor	Talking Time P
Dyad Dynamics	Influence of Self on Other I_{so}
Dyad Dynamics	Influence of Other on Self I_{os}
Changes in Pitch	Std. Dev. of Formant Freq. σ_f
Stress Measure	Std. dev. of Spectral Entropy σ_s
Speaking Energy	Std. dev. of Energy σ_e
Number of Short Exchanges	Shorts S

4 Bob: You can start on Monday?

The mirroring behavior is seen in lines 2 and 3, when Bob makes a terse response and Alice makes another short response in return. This mirroring behavior has been anecdotally linked to agreement and progress [6].

All of these features are interdependent at some level. This interdependence is taken into consideration when the features extracted from the signals are used in analysis.

2.3 Finding Patterns in the Data

Parties enter negotiation with a purpose, as described in section 1. In order to bring about some action, participants lay down the terms and agreements must be made.

Humans speak on average at about three phonemes per second. (A phoneme is defined as: the smallest phonetic unit in a language that is capable of conveying a distinction in meaning. For our purposes, we can think of a phoneme as a syllable.)So if we estimate that communication is happening at one word per second, the substance of a negotiation is happening on the level of minutes, taking into account pauses and switches between speakers. Using this inference, it makes sense to study the data at a minute by minute level.

As a part of this approach we split the original signal (example Figure 2) into one minute chunks, and extract the information discussed in the previous section from each of the one minute chunks. We end up with a vector of the measurements for each minute of a negotiation, from the perspective of one of the participants, the elements of the vector being the different feature whose extraction was noted in the previous section. Here is an example of a vector for the third minute of a negotiation between two subjects, A and B as computed from subject A's perspective:

ID	P	I_{so}	I_{os}	σ_f	σ_s	σ_e	S
A_3	4.893	.353	3.787	18.890	6.904	2.670	3

These minute vectors give us a picture of what is going on each minute, and provide a basis for looking at temporal patterns. There are several things that we would like to know about these minutes. How do these minute vectors relate to each other? Do minutes that belong to the same negotiation have the similar features? Do minutes from negotiations of male speakers differ from those of female speakers? Of course, what we are hoping is that the minutes break up according to the structure of a negotiation, i.e. the minutes where agreement is reached, or questions are asked, or explanations are made will have similarity in the minute vector characterization.

Our study is an investigation of the natural categories that these minutes fall into, and the meaning of this categorization. For example, suppose two minute vectors are almost identical in their values, how are they related on the psychological and behavioral front? The section 3 goes into our categorization and evaluation process.

Once we have these categories, we study how they are linked in time and to each other, that is, what stages a negotiation goes through. How this is achieved is described in section 3.

3 Materials, apparatus and procedures

3.1 Methods and Procedure: Mock Negotiations

The mock negotiation experiment was designed to create an environment in which we, the researchers, could study the process of negotiation in a controlled and observable setting.[2]

3.1.1 Participant Selection

Participants for this study were chosen from a class of first-year students at MIT Sloan School of Management. These students had to complete a mock negotiation as part of course work. Almost all of these students had work experience, and were interested in participating in a study about negotiation. Because of their interest in management, we assume that the majority of them have participated in negotiations sometime before this study. Possible settings they negotiated in may have been salary negotiations for a job, or negotiating work on a project in a group setting.

3.1.2 Setup

The participants were paired by gender (18 female pairs and 28 male pairs) and one member of each pair was designated a Vice President (VP), and the other a Middle Manager (MM). In the fictional scenario the VP and the MM work for the same company, and the MM is applying for a transfer to the VP's division. Because this is a switch, not an advancement or demotion, compensation and placement are flexible. The subjects negotiated a number of points, some having fixed choices and ranges. The participants were compensated monetarily to give them incentive to negotiate vigorously for a better position.

[2]The mock negotiation experimental design was created by Jared Curhan an MIT Assistant Professor of Organization Studies

Table 2: Points Negotiated

Subjects Covered as Part of Negotiation
Salary
Vacation
Company Car
Division
Health Care
Benefits

It is important to note that each participant had their own microphone into which they were speaking. Analysis in this thesis is done with a perspective, either the *owner* of the microphone or the *other* participant. These descriptors *owner* and *other* refer to the microphone input being processed.

3.2 Extracting Features

The goal of extracting features from the signals captured by the microphones during mock negotiation, as described in section 2, is to create a basis for temporal analysis. When humans talk to each other they communicate in bursts of information and action surrounded by lulls and descriptors, during which, for many people a "sinking in" of information happens [7]. There is a tradeoff that needs to be made within this framework. It is the hope that we can capture the essence of this subjective oscillation of negotiations whatever their periodicity, but must also remain mindful of the limitations of the sensors and algorithms that we are using to study the dynamics of interest. To allow our algorithms and sensors to have high accuracy, as well as being able to study the negotiation dynamics with an adequate amount of granularity, we decided to split the negotiation recordings into one-minute segments and extract the following data from the one-minute chunks:

1 Talking time of owner
2 Influence of owner on other
3 Influence of other on owner
4 Spectral Entropy
5 Energy
6 Formant Frequency
7 Shorts (5 minute intervals)

There are many subtleties involved in calculating some of the simple features, such as "Talking Time" or "Shorts," and the process of extracting the features is described in detail in this section.

3.2.1 Basic Processing

During the negotiation, each participant is wearing a microphone, but variation in volume, person proximity and other factors can cause a person's microphone to pick up a number of other signals including noise and the voices of both participants. Some of the features require computations based on the voice of a single individual, some require both participants' voices, and others require both silence and sound. Basic processing techniques take care of these cases so that more complicated analysis can be done with the refined signals.

A fundamental process in speech analysis is determining the voiced and unvoiced regions of a signal. Voiced regions (roughly) correspond with regions of the signal that are vowel sounds, whereas unvoiced regions correspond with consonant sounds. Silence is a third category in which human speech is not present in the signal. In this study, voiced and unvoiced regions are computed as described in [2]. This process is robust to noise and variability in microphone distance and performance. The annotation of the signal into voiced, unvoiced, and silent regions is a base for higher level features. It is important to note that voiced and unvoiced sections correspond to the time intervals during which speech is happening.

After a signal has been separated into voiced/unvoiced/silent chunks, another important operation that needs to be done on a signal is to determine at which intervals each participant is talking. The idea of this algorithm is that if a microphone is closer to one person than another, than the closer person should sound louder in the signal captured by their microphone. The algorithm employed in this case takes the energy of the signal E, and then makes a histogram of the computed E signal.

$$E = \int |x(t)|^2 \, dt. \tag{1}$$

In general the histogram of the energy signal looks like a mixture of three Gaussians, one corresponding to the microphone's owner's speech, another to when both participants are talking, and the third when both parties are talking. The parts of the signal corresponding to the two higher energy Gaussian are marked as belonging to the speaker.

Consider a signal $x(t)$, the raw voice signal. Here are some symbolic variables for representative values produced by the analyses described in this section, lets call this function $b(t)$. At the end of these two steps we have a signal that has the following annotations:

1 Voiced ($b(t) = t_V$),
2 Unvoiced ($b(t) = t_{UV}$),
3 Silent regions ($b(t) = t_S$),
4 Microphone owner speaking ($b(t) = t_{\text{me}}$),
5 Both speaking ($b(t) = t_{\text{me\&you}}$),
6 Other negotiation participant speaking ($b(t) = t_{\text{you}}$).

These serve as building blocks for several of the other features that are used as a basis for the negotiation analysis described in this thesis.

3.2.2 Talking Time

Using the basic processing techniques described in the previous section calculating the talking time is relatively simple. When the signal is chunked into sections by speaker (silent zones are not counted), the ratio can be taken of each speaker's talking time to the total talking time [4]. The equation used to calculate the talking time P of an individual is calculated using the following equation:

$$P = \frac{\#\text{ samples in } b(t) \text{ s.t. } t = t_{\text{me}} || t = t_{\text{me\&you}}}{\#\text{ samples in } b(t) \text{ s.t. } t = t_V || t = t_{UV}} \tag{2}$$

This will provide a ratio of the time that the microphone owner is speaking to the total time spent speaking. In this study recordings from the microphones were chunked into one minute intervals and the talking time was calculated for each of these intervals, for each person's microphone in all negotiations.

3.2.3 Influence Factors

Influence factors, of both participants on each other are an important part of the analysis of negotiations described in this thesis. Here the influence factors are computed in accordance with the influence model created by C. Asavathiratham, [1].

Consider the graphical model that is shown in figure 3. The influence model uses a tractable algorithm to estimate the effect of all of the edges going into the node H_i considering the values of the nodes that connect to it.

3.2.4 Spectral Entropy

The standard deviation of the spectral entropy of the signal, the energy of the signal, and the formant frequency of the signal are calculated from the raw signal itself. In all the explanations for feature extraction, in the examples $x(t)$ of is the original signal, as in figure 5.

Calculating spectral entropy involves looking at the signal in the frequency domain. Here is the frequency content of the signal $x(t)$, figure 6. The spectrum is calculated using a Fast Fourier Transform algorithm.

As the name suggests, spectral entropy is a measure of the entropy in the frequency domain of $x(t)$, or simply, the entropy of $X(s)$. The entropy of a signal is calculated as follows:

$$H(s) = -\sum_{s \epsilon S} X(s) \lg X(s) \tag{3}$$

The raw audio is a digital signal sampled at 8000 samples per second and the spectral entropy is calculated for each 128 samples; so for one minute 3,750 spectral entropy values are calculated each minute. The standard deviation of all 3,750 samples is the σ_s extracted from each minute of the signal.

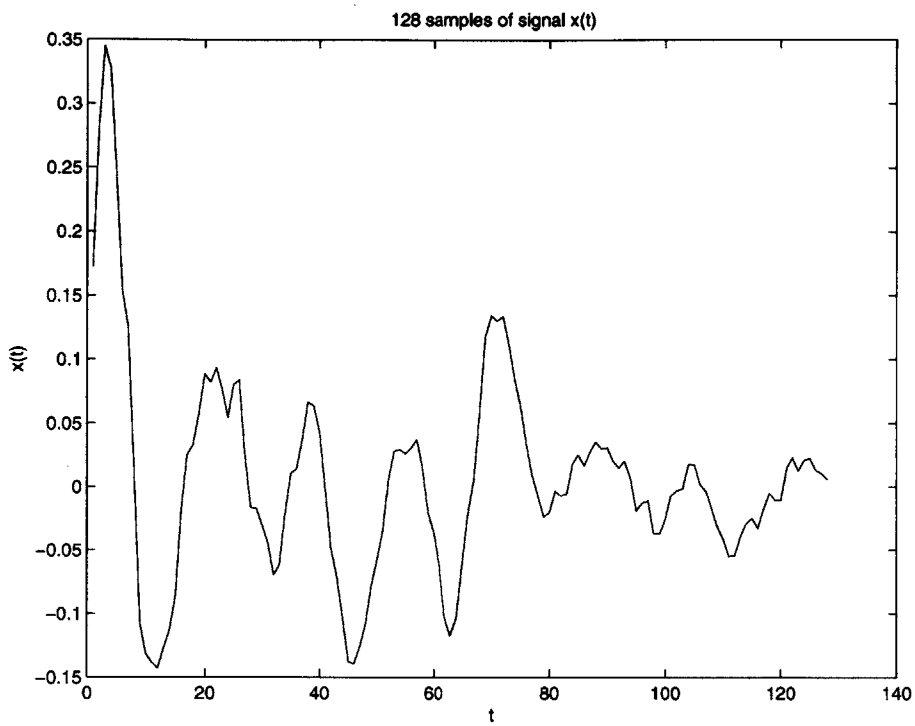

Figure 5: Plot of one minute of negotiation audio, $x(t)$

3.2.5 Standard Deviation of Energy

The energy of the signal is computed in a strait forward way. Consider the original signal $x(t)$ as seen in figure 5. The energy is calculated according to the following formula:

$$E = \sum_{k=0}^{n} \|x(t)\|^2 \tag{4}$$

Again, as in the spectral entropy the digital signal sampled at 8000 samples per second and the energy is calculated for every 128 samples. The standard deviation of the energy of the 3,750 frames is the σ_e calculated from each minute of the signal.

3.2.6 Standard Deviation of Formant Frequency

The formant frequency is the major frequency component within the voice signal. Consider the spectrum of the 128 sample windows in figure 6. The formant frequency is the frequency of the highest peak within any 2π interval. As in σ_e and σ_s, 3,750 formant frequencies are calculated per minute and the standard deviation σ_f is extracted from each minute.

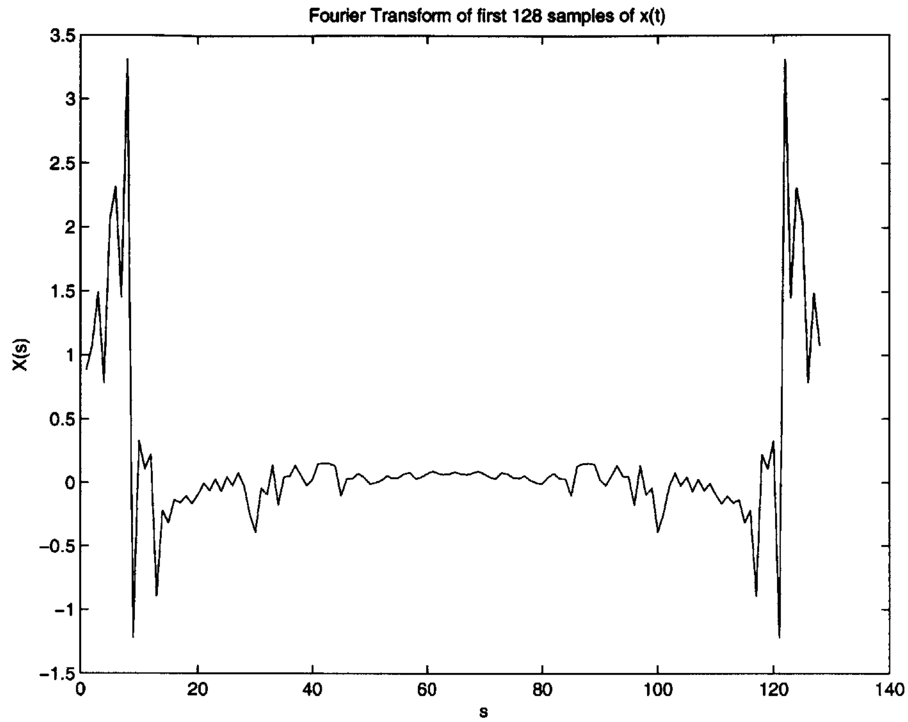

Figure 6: Spectral components of signal $x(t)$, in the frequency domain becomes $X(s)$

3.2.7 Shorts

Shorts are the number of short exchanges there are within a recorded signal $x(t)$. To determine this, we survey the signal and count the number of "patches" where t_{me} is uninterrupted. The calculation of this value differs from the calculation of other values used in this analysis because it is calculated over a period of five minutes instead of one minute, and each one minute feature vector gets its value as one-fifth of the calculated value for the five minute interval to which it belongs. This is done to increase the accuracy of the calculation. Over the interval of only one minute the algorithm that counts the number of shorts within that minute can be highly noisy.

The algorithm simply counts the number of times where t_{me} for less than a second followed by t_{you} for less than a second. This is calculated for a 5 minute interval and than divided by 5.

The accuracy of the number of shorts can be audited by listening to the original audio, slowly.

During the experiments we recorded a total of 46 2-person negotiations. Each person had a personal microphone, so we collected 92 raw signals. Each of these signals were split into one minute chunks, a total of 2050 minutes. The features described in this section were extracted from each of the minutes creating a total of 2050 feature vectors, that provide the basis for the classification and temporal model.

3.3 Classification

At this point in the process we have 2050 vectors all the person minutes that were recorded in the negotiation experiment. The next problem to solve is how to classify the vectors. This is a non-trivial process for a number of reasons. First, we don't know what categories to put minutes in. Classical negotiation analysis looks more at behavioral traits as belonging to a person, but being static over time [7], but our study includes indicators that vary minute by minute. The non-existence of objective categories leaves us with a trial and error approach, making categories that *seem* right, or in some cases, do not seem wrong. We acknowledge the difficulty and subtlety of the classification process and built into the cluster-making process a very tight development and audit process to prevent wild deviation from the ground truth that is the signal captured.

3.3.1 Exploring High Dimensional Vector Space

First of all, we were interested in how the data organizes itself in vector form. With the feature space that we accumulated from experiences of our own and others, are there inherent patterns in the data? The first step is to look at the distributions of the data, which we accomplished using a rough histogram.

The Talking Time feature in each of the vectors appears to have a Gaussian distribution, figure 7.

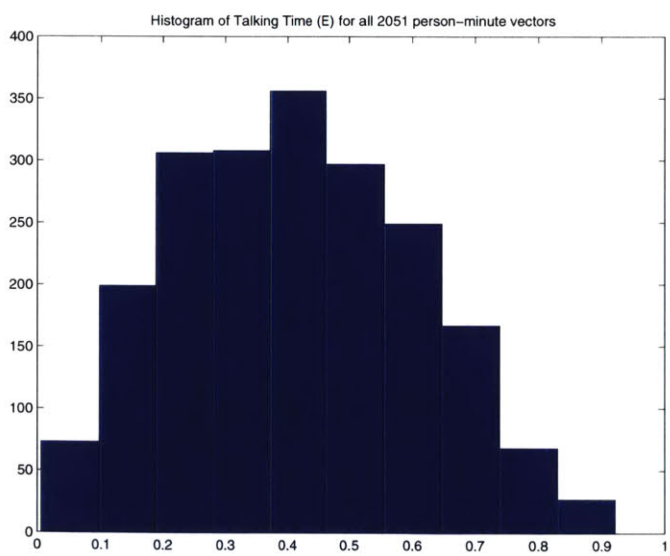

Figure 7: Distribution of Talking Time (P) Values in the data set

The influence factors are interesting to look at because appear to have an exponential distribution, figures 8 and 9. This suggests that during most minutes of the negotiation is little influence of one party on another, but during some minutes, presumably interesting ones, the influence level is high.

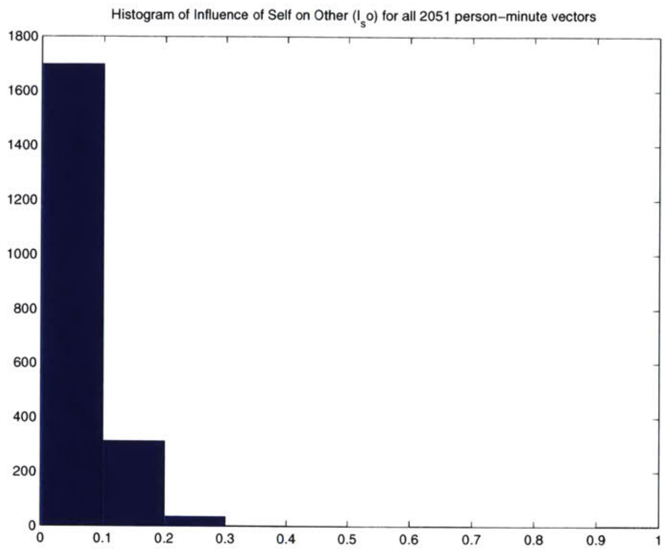

Figure 8: Distribution of Influence of Self on Other (I_{so}) Values in the data set

The next three features, the standard deviation of formant frequency (σ_f), the standard deviation of spectral entropy (σ_s), and the standard deviation of energy (σ_e) had an interesting relationship to the data set. These features had a roughly exponential distribution, similar to the ones in figures 8 and 9. On minutes where any one or more of σ_f, σ_s, and $sigma_e$ were in the 75th percentile or above the minutes sounded very similar to listeners. All minutes with these clusters had tension and stress, pushing and resistance. If taken seperately, elevated values of σ_f, σ_s, and σ_e seemed to have special significance, such as female-tension, or foreign-accent tension. Because we were only interested in the behavior, and not the demographics, we combined these values in the feature vector by adding them, converting the feature vector from:

$$\begin{array}{cccccccc} ID & P & I_{so} & I_{os} & \sigma_f & \sigma_s & \sigma_e & S \\ A_3 & 4.893 & .353 & .787 & 18.890 & 6.904 & 2.670 & 3 \end{array}$$

,

to a new vector with fewer features:

$$\begin{array}{cccccc} ID & P & I_{so} & I_{os} & \sigma_f + \sigma_s + \sigma_e & S \\ A_3 & 4.893 & .353 & .787 & 18.890 + 6.904 + 2.670 & 3 \end{array}$$

This modifications of features in the vector eliminated some uninteresting distinctions while preserving interesting ones. A histogram of the new feature $\sigma_f + \sigma_s + \sigma_e$ is shown in

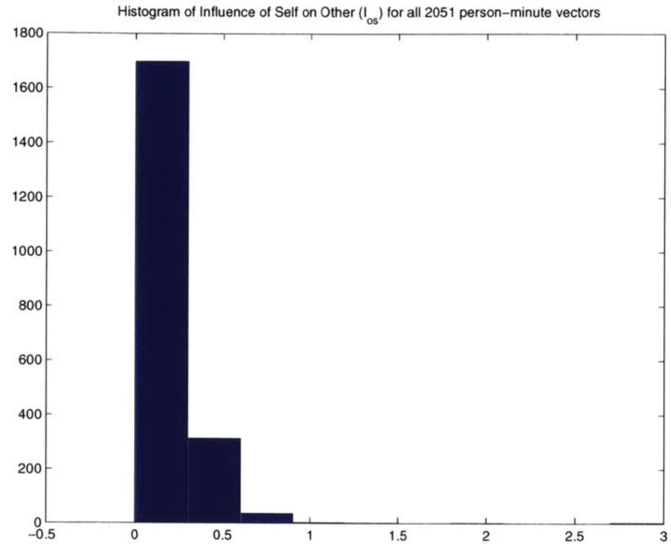

Figure 9: Distribution of Influence of Other on Self (I_{os}) Values in the data set

Figure 10. This new feature, because of its link to tension, became colloquially known as the stressor feature.

The distribution of shorts (S) in the data set also seems to indicate that most minutes have very few shorts, but the few that do have the potential to be very interesting. A plot of their distribution is given in Figure 11.

Studying these distributions, and checking that their attributes were not due to other obvious trends in the data was one check that was performed to confirm that we indeed had something of interest extracted from the raw signals. With the feature vectors pared down to five features we then sought to make categories out of the 2050 vectors with 5 features each.

Because some features are an order of magnitude larger than others, we normalized the data so that the average of each of the features was 0, with a standard deviation of 1 (zscore). This was done over the whole dataset of 2050 vectors. With the scaled vectors we looked to see how the features were related to eachother. To do this we performed a heirarchical nearest neighbor clustering, using a centroid linkage function, and measuring the distance between two vectors as the Euclidean distance between them. Based on the topology of the obtained tree we extimated that the data fell naturally into 5, 6, 7 or 8 clusters, depending where you transect the heirarchical tree.

3.3.2 Making and Auditing the Clusters

We now had to decide how many clusters to break the data into, 5, 6, 7, or 8. Because we did not have a set of clear criteria of how to break up clips based on psychological, or verbal content, so we decided to use numerical analysis to choose the best categorization.

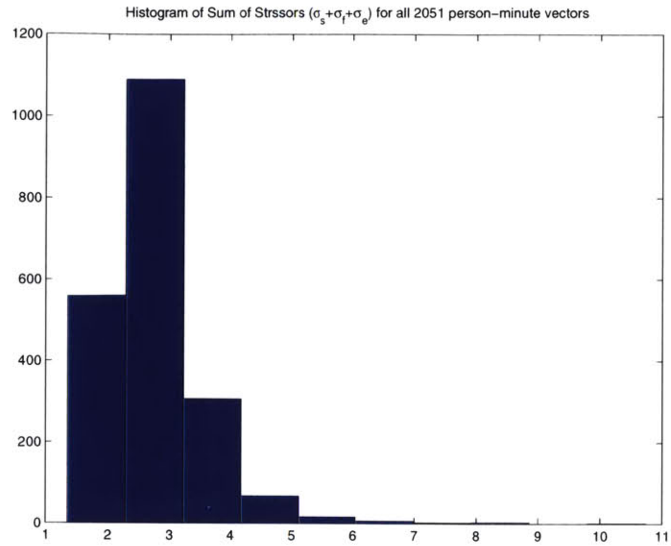

Figure 10: Distribution of Stressor $(\sigma_f + \sigma_s + \sigma_e)$ Values in the data set

Our goal was to find the natural clusters that existed in the data. We used an algorithm called K-means that takes as inputs a dataset of same length vectors and an integer number n, and iteratively tries to find the n mean vectors or centroids that represent the centers of n clusters within the data. We ran K-means on our dataset of 2050 vectors with values of $n = 5$, $n = 6$, $n = 7$, and $n = 8$. In each of these clusterings we measured the average distance of each data point to its respective centroid, and found that this value was a minimum at $n = 7$. This indicated to us thet he clusters were tightest at $n = 7$ and so it was the logical choice for the number of centroids.

There are many implementations of the K-means algorithm, and so we will describe the one we used. The K-means algorithm we used selects $n = 7$ points randomly from the data set and uses these as the initial centroids. The other points in the data set are put in a line in random order. One at a time, a point is pulled off of the line and its distance to each of the 7 centroids is computed. It is added to the cluster corresponding to the closest centroid. The centroid (average of all elements) of the cluster to which the new point is added is updated. This is done until all the points in the line are used. Then all of the points in the data set are put in line in random order again, and using the centroids from the previous iteration the assignment and update process is done again. This is done until the centroids change only a small percentage from one iteration to the next.

We subjected our data set to the K-means algorithm of the order of 20 times for $n = 7$ to check that the centroids did not vary a large amount during each application of the algorithm. We discovered that in most of the instances the centroids came out quite similar, though there were a handful of cases in which the centroids varied substancially. We chose arbitrarily chose a categorization from the majority similar centroids. Thus we broke the

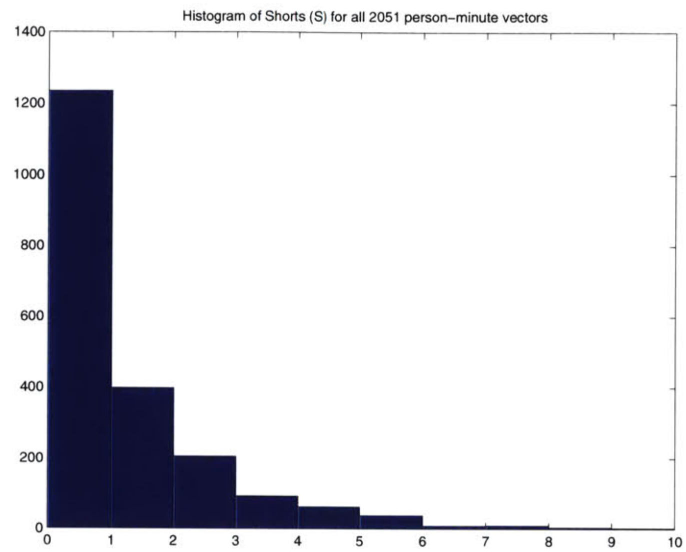

Figure 11: Distribution of Shorts (S) Values in the data set

data set into categories.

The next step was to look at what was going on in the categories that we had discovered in the data set. This process was undertaken by playing clips of the negotiations corresponding to points in the data set. In groups of 3 or 4 we listened to the clips and tried to discern what we heard going on during the clips. We called this the auditing process. (A version of this was done to examine the feature space, and is how we learned the role of the stressors.) We first did a free-form session in which we made notes about the clips based on instinct, how we interpreted the vocal cues that were present in the negotiation. Later, for fine tuning we looked at the feature vectors while listening to the clips and evaluated our ideas of what the features signified (see section 2) and what we actually heard in the clips. We also did a differencial analysis. We listened to clips from the same negotiation, but different clusters and tried to tease out what we were hearing that was different. This helped us find outliers, provide labels to the categories, and rule out some of the categorizations.

3.4 Making the Model

Once we had all the data points and a classification the next step was to make a model so that if a new negotiation were recorded we could apply our model and decode the action that was going on. This involved making an engine that would perform the same feature extraction on the audio, and find the point it was closest to in the data set and record to which cluster it belonged.

As part of building the model, we also present some facts about it. For example, what stage is the VP in when the MM is in stage x? What patterns do these stages come in?

These and other interesting patterns are described in section 4.

The model is the synthesis of the study and provides a picture about what is happening in a negotiation with easy reference back to the ground truth, to ensure that analysis does not stretch too far from reality.

4 Results

4.1 The Clusters

After listening to the clips in different categories, and looking at the feature vector values in each category we were able to get a sense of what was going on in each of the 7 clusters.

The first cluster contains minutes that had a lot of tension between the two parties, it seemed to be a review of points to be discussed or details of a pending agreement. There tend to be lots of pauses in points in this cluster, and the discussion sounds strained and tense. Below is a table showing the mean and variance of this cluster, both in the original and normalized coordinates. Here normalized means zscored, or made so that all elements of all vectors in the data set have a mean zero and standard deviation of 1.

Description	E	I_{so}	I_{os}	$\sigma_f + \sigma_s + \sigma_e$	S
Centroid 1(zscore)	0.20	-0.47	-0.46	2.48	0.20
Variance (zscore)	0.83	0.25	0.29	2.01	0.85
Centroid 1	0.45	0.03	0.08	4.68	1.81
Variance	0.03	0.00	0.01	1.20	2.20

The second cluster is the contains minutes in which the owner is talking a lot. This is usually explanatory in nature. One feature about these clips is that it appears that the other party, the one who is not speaking very much is not very engaged. He or she interjects with a "right..." or "mmhmmm", but retains low engagement overall. Below is the summary of cluster 2 both in original and normalized (zscored) coordinates:

Description	E	I_{so}	I_{os}	$\sigma_f + \sigma_s + \sigma_e$	S
Centroid 2(zscore)	1.00	-0.39	-0.33	-0.13	-0.29
Variance (zscore)	0.36	0.24	0.24	0.48	0.35
Centroid 2	0.61	0.03	0.11	2.66	1.03
Variance	0.01	0.00	0.01	0.29	0.90

The third cluster seems to represent to owner questioning the other party. The other party here is explaining how they see it in return, and doing most of the talking. Though it appears that the owner is doing little talking in clips in this section, he/she seems to be dominating the conversation and calling the shots. It is also interesting to note that here the influence of self on other feature, I_{so} is, on average, approximately three standard deviations above the mean for the data set as a whole. This cluster evidently contains many entries with high influence scores, as seen in its distribution in see section 3.

Description	E	I_{so}	I_{os}	$\sigma_f + \sigma_s + \sigma_e$	S
Centroid 3(zscore)	-0.85	2.94	-0.06	-0.28	-0.02
Variance (zscore)	0.87	5.84	0.54	0.45	0.68
Centroid 3	0.25	0.24	0.16	2.55	1.46
Variance	0.03	0.02	0.02	0.27	1.74

The fourth cluster is particularly interesting. From a human listeners standpoint, the both participants are speaking and engaged in the conversation. They are exploring points in the negotiation. It seems that minutes that fall in this category are interesting ones in the negotiation. A summary of cluster 4 is given below:

Description	E	I_{so}	I_{os}	$\sigma_f + \sigma_s + \sigma_e$	S
Centroid 4(zscore)	0.18	0.75	0.76	-0.37	-0.37
Variance (zscore)	0.46	0.37	0.34	0.42	0.35
Centroid 4	0.45	0.10	0.31	2.48	0.88
Variance	0.02	0.00	0.01	0.25	0.92

The fifth cluster distinguishes itself with a large number of shorts. Shorts are usually affirmations of points, consequently one would expect that this cluster would be participants agreeing to large extent on different points. Listening to these clusters gave many of us think of polite chatter.

Description	E	I_{so}	I_{os}	$\sigma_f + \sigma_s + \sigma_e$	S
Centroid 5(zscore)	-0.07	-0.36	-0.34	0.17	1.91
Variance (zscore)	0.58	0.29	0.33	0.65	0.85
Centroid 5	0.40	0.03	0.11	2.90	4.56
Variance	0.02	0.00	0.01	0.39	2.19

In cluster six someone listening to these clips it seems that the microphone owner is being probed on his or her point of view, and he or she is rationalizing her perspective to the other party. Here, the owner talks a lot, but seems to be highly influenced by the other participant. This is interesting and is in accordance with the dynamics of the interview setting, where the subject is heavily influenced by the interviewer. One more thing to note is that the influence of other on self, I_{os} is very high, which tells us about the relevance of the influence factor. A summary of this cluster is below:

Description	E	I_{so}	I_{os}	$\sigma_f + \sigma_s + \sigma_e$	S
Centroid 6(zscore)	1.06	-0.04	2.96	-0.29	-0.22
Variance (zscore)	0.82	0.56	5.81	0.43	0.45
Centroid 6	0.62	0.05	0.71	2.54	1.13
Variance	0.03	0.00	0.19	0.26	1.15

The final cluster, cluster 7, is comprised of minutes in which the owner is relatively unengaged. Here, the other participant is explaining, or talking about background while

Table 3: Summary of Cluster descriptions

Cluster Number	Description
1	Tension build
2	Owner explains
3	Owner questions Other
4	Explore Possibilities
5	Tension release
6	Other questions owner
7	Other explains

the microphone owner is asking a few questions or voicing agreement. From a numerical perspective the owner has very little talking time over these minutes.

Description	E	I_{so}	I_{os}	$\sigma_f + \sigma_s + \sigma_e$	S
Centroid 7(zscore)	-0.92	-0.33	-0.39	-0.16	-0.34
Variance (zscore)	0.27	0.24	0.20	0.41	0.33
Centroid 7	0.24	0.04	0.10	2.64	0.95
Variance	0.01	0.00	0.01	0.25	0.84

Overall, these clusters have tangible distinction, and groups listening to them were able to relate to the situations that they heard in the audio clips. Listeners also noted that there are several opposing pairs among the clusters. Cluster 2 seems to be the opposite of cluster 7. In cluster 2, owner is talking to other, explaining some detail. In cluster 7 other is explaining to owner. Clusters 3 and 6 also appear to be opposing. In cluster 3 owner in questioning other, while the opposite happens in cluster 6. Clusters 1 and 5 have a different kind of opposite nature. Cluster 1 is a time of tension, while cluster 5 is a time of relief. Anecdotally, the laughter is cluster 5 seems more genuine than that in cluster 1. Cluster 4 stands on its own, which is another reason why it is a cluster of heightened interest.

4.2 Relationship between Clusters

The with the clusters established, the next step of the analysis is to look for patterns in the negotiations. Specifically, we are interested in what order the minutes in a negotiation come in, and what state the owner is in when the other is in a given state. Also, (as described in section 3) we provided roles of Vice President (VP) and Middle Manager (MM) to the participants, and we are interested in seeing how those roles effect the temporal patterns.

Let us consider $h(t)$ to be the cluster that a negotiator is in at time t. Therefore $h(t) = \{1, 2, 3, 4, 5, 6, 7\}$. Also, if the negotiator is a MM this will be indicated as a subscript $(h_{MM}(t))$, and similarly for VP $(h_{VP}(t))$. Looking at all of the negotiations in the dataset we created a Markov chain by calculating the inter-state transfers. That is,

$$P(h(t+1) = x|h(t) = y)\forall x, y \qquad (5)$$

This provided us with a seven-by-seven matrix shown below. The entries here are for row r and column c, $(r, c) = P(h(t+1) = r|h(t) = c)$.

From↓ To→	1	2	3	4	5	6	7
1	0.3226	0.1935	0.0081	0.0403	0.1935	0.0161	0.2258
2	0.0490	0.3229	0.0267	0.1893	0.1247	0.0290	0.2584
3	0.0139	0.1806	0.0417	0.4722	0.0139	0.1389	0.1389
4	0.0141	0.1972	0.0610	0.3357	0.0728	0.0540	0.2653
5	0.1073	0.2275	0.0258	0.1416	0.2060	0.0172	0.2747
6	0.0278	0.0972	0.1250	0.4722	0.0278	0.0556	0.1944
7	0.0411	0.1986	0.0188	0.1575	0.1353	0.0188	0.4298

This matrix shows several interesting results, which are best summarized in pictorial form. Several transitions have very small values, indicating that they almost never happen, but others appear to have high probability of interchange. A diagram showing the transitions between states is shown in Figure 12.

There are many scenarios that become evident from looking at the Markov chain. It appears that the questioning or interviewing process (clusters 3 and 6) lead to the exploration of points (cluster 4) during which the meat of the negotiation seems to happen. Also, it appears that once tension appears in the negotiation, it is likely to be tense for a matter of several minutes.

The transition probabilities for VPs and MMs do show some interesting differences. For example, looking at the VP transition matrix 8 and Markov Chain, 13 you notice that the edge between nodes 3 and 4 is lighter in figure 13 than it is in figure 12. This signifies that when a VP is questioning it is less likely to lead to exploring possibilities. Other differences are also apparent in the chain.

Similar analysis can be done on the Markov chain for the MMs, figure 14(the transition matrix is in section 9). An interesting thing to notice here is that the MMs seem to allow the VPs more entry into the conversation, look at the edges between nodes 2 and 7.

It is also interesting to explore the state the owner is in when the other is in some given state. Below is a table that shows the probability that a MM is in state c given that the VP is in state r. The matrix is organized by rows (r) and columns(c) so that

$$(r, c) = P(|h_{MM}(t) = c|h_{VP}(t) = r) \qquad (6)$$

Here is the actual table showing the probabilities, and a picture of the probability matrix is shown in Figure 15. This analysis shows some interesting results. When the VP is in a state of tension (cluster 1), the MM is either explaining (cluster 2) or being explained to (cluster 7). If the VP is explaining (cluster 2) the MM is most likely to be listening (cluster 7). The rest of the dynamics show some interesting results that reinforce some of the theories

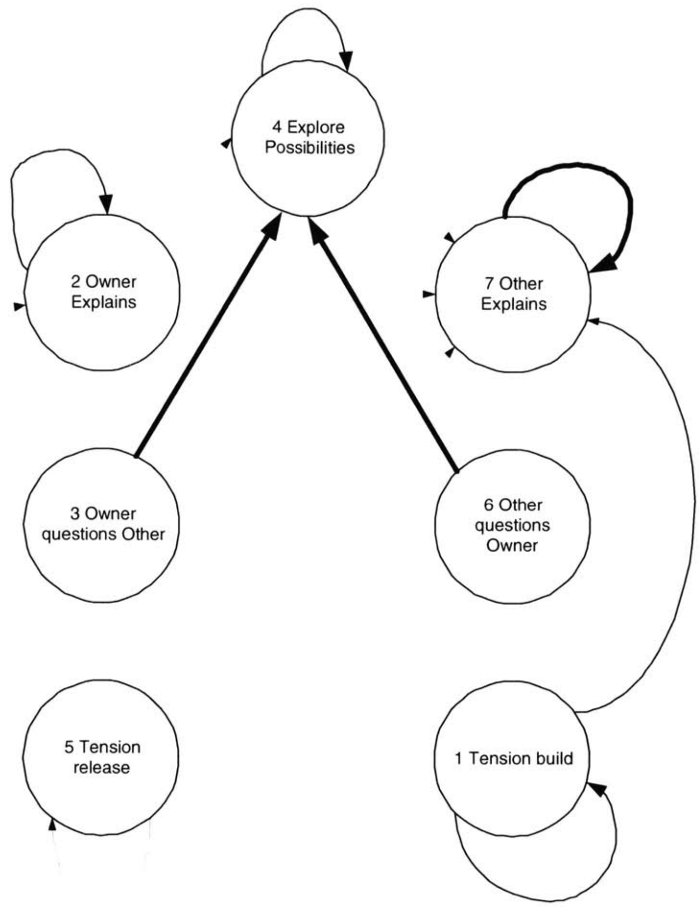

Figure 12: A diagram of transitions between states, transitions with higher probabilities are shown with heavier edges. Only transitions with .2 probability or better are shown on the graph.

about the feature space that we expressed in section 2. When the VP is questioning (cluster 3) the MM, the MM is usually being questioned (cluster 7). When the VP is exploring points (cluster 4) the MM is usually doing the same. When the VP is laughing or chatting (cluster 5) the MM is usually doing the same. When the VP is being questioned (cluster 6), the MM is questioning. Finally, when the VP (cluster 7) is explaining some point, the MM is usually explaining back (cluster 2), or listening (cluster 7). As discussed in section 3, the features were calculated on different microphone streams during the negotiation, each person had their own microphone, so it is an affirmation of our choice of features and clusters that these intuitive dynamics occur.

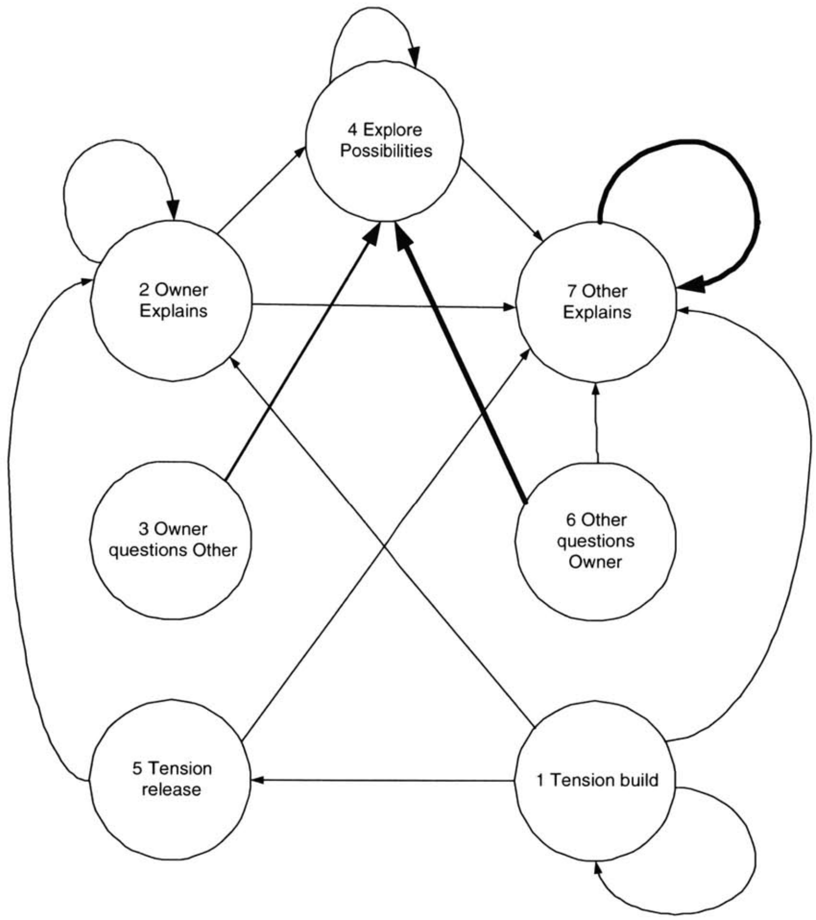

Figure 13: A diagram of transitions between states for VPs only, transitions with higher probabilities are shown with heavier edges. Only transitions with .2 probability or better are shown on the graph.

MM ↓ VP →	1	2	3	4	5	6	7
1	0.2424	0.2727	0	0.0606	0.1515	0	0.2727
2	0.0333	0.2000	0	0.0200	0.0333	0	0.7133
3	0	0	0	0.0500	0	0.9500	0
4	0	0.0205	0.0137	0.8699	0.0137	0.0068	0.0753
5	0.0345	0.0920	0	0	0.7816	0	0.0920
6	0	0	0.9091	0.0909	0	0	0
7	0.0455	0.5390	0	0.0519	0.0390	0	0.3247

Looking from the perspective of the MM (figure 16), the dynamics are similar to the perspective VP, figure 15. Below is a table that shows the probability that a VP is in state c given that the MM is in state r. The matrix is organized by rows (r) and columns (c) so

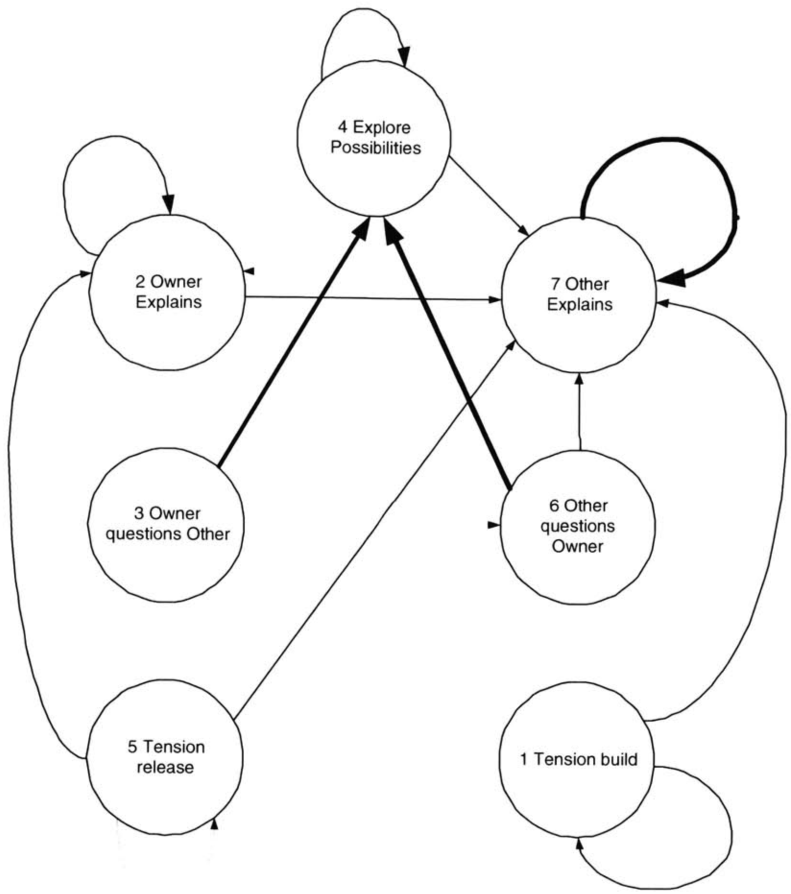

Figure 14: A diagram of transitions between states for MMs only, transitions with higher probabilities are shown with heavier edges. Only transitions with .2 probability or better are shown on the graph.

that

$$(r, c) = P(|h_{VP}(t) = c|h_{MM}(t) = r) \tag{7}$$

Here is the actual table showing the probabilities, and a picture of the probability matrix is shown in Figure 16. The similarities between figure 16 and figure 15 are striking, but they do have some marked differences. It seems that from the VP's perspective, when the VP is experiencing a tension build (cluster 1), the MM is more likely to be experiencing tension than the other way around (when the MM is experiencing tension, the VP is probably not). In addition, when the MM is explaining (cluster 2), the VP is equally likely to either explain back (cluster 2) or listen; while when the VP is explaining the MM is more likely to explain back than to listen. This indicates that the MM is under more scrutiny than the VP in most cases.

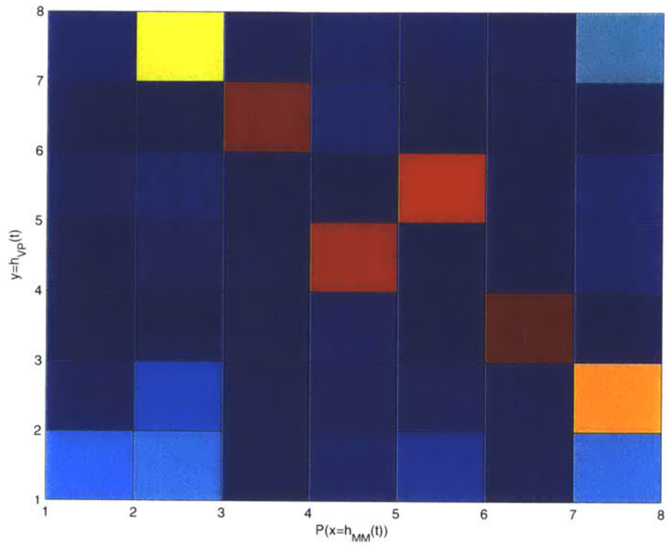

Figure 15: The color of square (x, y) is indicative of the $P(h_{MM}(t) = x | h_{VP}(t) = y)$. All the rows in this diagram sum to 1.

VP ↓ MM →	1	2	3	4	5	6	7
1	0.1186	0.2881	0	0.0508	0.1695	0	0.3729
2	0.0333	0.1556	0	0.0111	0.0222	0	0.7778
3	0	0	0	0.0769	0.0769	0.8462	0
4	0	0.0139	0	0.9306	0	0.0139	0.0417
5	0.0732	0.1220	0	0	0.6585	0	0.1463
6	0	0	0.9444	0.0556	0	0	0
7	0.0083	0.4583	0	0.0417	0.0417	0	0.4500

The differences in dynamics can be further explored by building a Markov chain like the one in 12, but specific to VPs and MMs. This might help expose attributes of the power structure inherent in the negotiation.

Originally, we wanted to correlate the patterns in negotiations with success of the negotiation, but we very quickly realized that in the experimental design 3 we let people choose what was more important to them, a nice car, or a high salary for example. This changes the negotiations so that they are no longer a zero-sum game, but can have mutual profit for both parties [7], making correlation of this type less meaningful.

One trend that we did notice, by listening to the negotiations was that the VPs seemed to be in a more secure and comfortable position than the MMs. This might be a reflection of the subjects we used, and the current business environment. It might also reflect on the differences in dynamics between VPs and MMs, that is why MMs might experience more tension than VPs and why more VP states result in MMs experiencing tension.

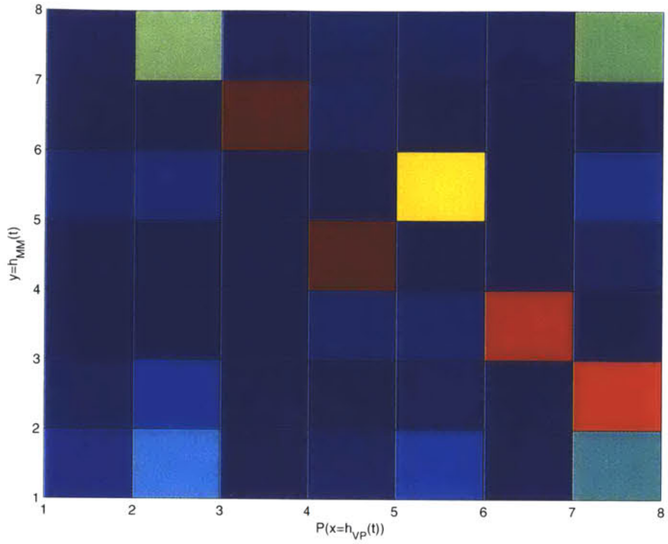

Figure 16: The color of square (x, y) is indicative of the $P(h_{VP}(t) = x | h_{MM}(t) = y)$. All the rows sum to 1.

5 Discussion

As in most quantitative studies of social situations, it is very difficult to get an "objective" measure of success or failure. We are attempting to merge complex common-sense information with signal analysis to create a model of a phenomenon, negotiation. Our minds, subconsciously know how to parameterize a negotiation. We know when we are feeling tense, or when we are explaining, etc. We also know the effects on the other party, to some extent.

This thesis sought to test the effectiveness of the features that we hypothesized would be useful, and to find a numerical way to analyze the patterns and dynamics that negotiation experts observe in the field. The methods described in this thesis show that we are able to do this to some extent.

The best way to gauge the actual contributions of this study is to look at the byproduct coincidences that emerged. First, we note that in this feature space the data has a propensity to organize itself into seven clusters (six might also be possible, and we would have to do another study to definitively choose). Because of the distributions of the features, the majority of the clusters are comprised of fairly average feature scores, with the exception of one of the features which is "off the charts." The data tends to organize itself that way, and as a result we can easily sift out the role of our features in the audio.

Our second way to gauge our contribution is to look at the symmetry of the model. Six of our clusters form pairs (1,6) (2,7) (1,5) and the remaining cluster seems to be the one where both participants share a similar role within the negotiation. It is very telling that even through we had no knowledge of the calculations at the time we were auditing the clusters we note that the clusters have an intuitive relationship to each other, much in accordance with common sense.

6 Conclusions

The product of this investigation is a model of a negotiation that annotates the discourse with labels that humans can understand without human intervention. The interesting parts to note about the model is that it appears to have a universality to it. When referring back to section 3, you might note that the experimental protocol was relatively loose. Research subjects were given a topics to discuss, but otherwise they were given free reign on what order, and how to conduct the negotiation. This leads us to believe that this model may actually be more general than we have described it.

We have also reached some subsidiary conclusions while developing said model. First, we noted the role of stressors ($\sigma_f, \sigma_e, \sigma_s$), and we able to manipulate them to serve our purposes. Second we found out some interesting facts about the meaning of the influence factor. As described in section 2, the influence factor is an approximation of the influence of other nodes on a given node in a complex graphical network. Originally, this analysis was used to approximate the fluxuating voltages within a network of power stations, and so it was unclear exactly how it translated to the world of human behavior. This study gives us a peek into its meaning through clusters 3 and 6, which have very high influence values, and seem to indicate the scenario when one participant questions another with the aim to really understand his or her situation, not just small talk. A final subsidiary conclusion that there are segments of the negotiation where the details are worked out, and action is agreed upon (cluster 4).

7 Reccomendations

There are several avenues for further exploration within this realm. First, a subsequent experiment that might be interesting would be to create a new negotiation experiment and apply the model to the results. In this new experiment, the points of the negotiation would be more tightly controlled so that there would be a clear "winner" and "loser" as opposed to the possibilities for mutual benefit. At this point we could start to success with patterns of states.

The universality of the model also should be explored. Perhaps analysis on other data sets, recordings of hostage negotiations, or sentencing negotiation episodes of Law and Order spring to mind. These new data sets should be subjected to the same feature extraction, clustering and auditing for comparison. Later, migration to more general data sets like tapped telephone conversations may be in order if the model holds up.

The feature space should be explored as well, particularly talking time, influence and stressors. This can be done with a differential analysis, careful choice of speech segments and then correlation with calculated values is in order.

Once more about the domain and features are known, it is important to think about possible applications. An obvious area that emerges is the arena of multimedia indexing for things like legal dispositions and arbitration. These negotiations tend to be lengthy and an easy way to filter through to the important parts can be of great benefit. There are many applications along these lines that could serve us well in a domain where it is difficult to manage the amount of data.

If it was found that there were correlations between successful negotiation and certain patterns of the categories enumerated in the model, we could start on another application stream where individuals are able to modify their behavior to maximize their chances of successful outcome.

References

[1] Chalee Asavathiratham. *The Influence Model: A Tractable Representation for the Dynamics of Networked Markov Chains.* PhD thesis, Massachusetts Institute of Technology, 2000.

[2] Sumit Basu. *Conversational Scene Analysis.* PhD thesis, Massachusetts Institute of Technology, 2002.

[3] Martin C. Martin. Power point presentation: Matlab code.

[4] Martin C. Martin. Power point presentation: Speech features.

[5] James G. Patterson. *How to Become a Better Negotiator.* Amacom, 1996.

[6] Alex (Sandy) Pentland, Jared Curhan, Nathan Eagle, and Martin C. Martin. Money and influence. Technical Report 557, MIT Media Lab, 2004.

[7] Howard Raiffa, John Richardson, and David Metcalfe. *Negotiation Analysis.* The Belknap Press of Harvard University Press, 2002.

[8] Klaus R. Scherer, Judy Kovumaki, and Robert Rosenthal. Minimal cues in the vocal communication of affect. *Journal of Psycholinguistic Research*, 1972.

8 Appendix 1: Cluster Summary

8.1 Vector Format

ID	P	I_{so}	I_{os}	$\sigma_e + \sigma_f + \sigma_s$	S

For clustering, all vectors in data set zscored (this means normalized so that they have a mean 0 and std dev 1)

8.2 Cluster Summary Legend

```
Centroid centroid#:   a vector representing the average vector in
this cluster (zscored)
Var (z):  the variance of all vectors in this cluster (zscored)
Mean:  the vector representing the average vector
in this cluster (non-zscored)
Var:  the variance of all vectors in this cluster (non-zscored)
Size: number of vectors in this cluster
(of 2050=total number of vectors in the data set)
Male: ratio of males to total in cluster vs
ratio of males to total in data set.
VPs : ratio of VPs to total in cluster vs
 ratio of VPs to total in data set
When: average time the vector is in the conversation
+/- std dev of the time the vector is in the conversation

VPs: (list of 5 closest VPs to the centroid)
Minute#_Name_Negotiation# :  feature vector representing this minute
(rank, how close to centroid, 1 is the closest)

MMs: (list of 5 closest MMs to the centroid)
47_lal_36          :  0.29 -0.82 -0.86  2.40  0.32 (1)
```

8.3 Data

```
Centroid 1  :  0.20 -0.47 -0.46  2.48  0.20
Var (z)  :  0.83  0.25  0.29  2.01  0.85
Mean  :  0.45  0.03  0.08  4.68  1.81
Var  :  0.03  0.00  0.01  1.20  2.20
Size: 129 of 2050
Male: 0.434 vs 0.597
VPs : 0.450 vs 0.500
```

```
When: 0.502 +/- 0.324

VPs:
35_roreilly_30    :  0.74 -0.24 -0.51  2.30  0.32 (3)
30_ethan_15       : -0.25 -0.69 -0.92  1.97  0.32 (7)
27_roreilly_30    : -0.45 -0.60 -0.50  2.95 -0.30 (8)
12_sbiggar_34     :  0.21 -0.85 -0.92  3.12 -0.30 (9)
36_roreilly_30    : -0.44 -0.76 -0.58  1.79 -0.30 (14)
MMs:
47_lal_36         :  0.29 -0.82 -0.86  2.40  0.32 (1)
39_kmatsuda_8     :  0.09 -0.83 -0.83  2.28  0.32 (2)
19_lal_36         :  0.72 -0.75 -0.79  2.40  0.32 (4)
40_lal_36         :  0.47 -0.48 -0.67  2.02 -0.30 (5)
53_lal_36         :  0.64 -0.72 -0.87  2.09  0.32 (6)


Centroid 2  :  1.00 -0.39 -0.33 -0.13 -0.29
Var (z)  :  0.36  0.24  0.24  0.48  0.35
Mean  :  0.61  0.03  0.11  2.66  1.03
Var  :  0.01  0.00  0.01  0.29  0.90
Size: 465 of 2050
Male: 0.609 vs 0.597
VPs : 0.557 vs 0.500
When: 0.504 +/- 0.291

VPs:
21_boo_4          :  0.99 -0.62 -0.12 -0.30 -0.30 (1)
13_bsm_19         :  1.21  0.01 -0.31 -0.23 -0.30 (3)
27_akhosla_25     :  0.59 -0.34 -0.62 -0.26 -0.30 (6)
17_pravoor_22     :  1.07  0.11 -0.12 -0.21 -0.30 (9)
23_akhosla_25     :  1.44 -0.45 -0.56 -0.42 -0.30 (10)

MMs:
17_gburke_33      :  0.94 -0.76 -0.29 -0.25 -0.30 (2)
1_mfitzgerald_28  :  0.69 -0.60 -0.33 -0.48 -0.30 (4)
10_andera_25      :  0.92 -0.20  0.12 -0.26 -0.30 (5)
16_tgreenlaw_40   :  1.13 -0.49 -0.69 -0.47 -0.30 (7)
3_wilsonb_20      :  1.12 -0.22  0.15 -0.05 -0.30 (8)


Centroid 3  : -0.85  2.94 -0.06 -0.28 -0.02
Var (z)  :  0.87  5.84  0.54  0.45  0.68
```

```
Mean  :  0.25  0.24  0.16  2.55  1.46
Var  :  0.03  0.02  0.02  0.27  1.74
Size: 72 of 2050
Male: 0.583 vs 0.597
VPs : 0.431 vs 0.500
When: 0.423 +/- 0.288

VPs:
13_justinafan_35 : -0.92  2.65 -0.54 -0.39 -0.30 (2)
20_ocedar_18     : -1.17  2.17 -0.31 -0.17 -0.30 (3)
9_ocedar_18      : -1.09  3.26 -0.39  0.02 -0.92 (5)
7_shourong_2     : -0.22  2.09  0.06 -0.58  0.32 (8)
33_ocedar_18     : -0.89  1.82  0.21 -0.72  0.32 (9)

MMs:
24_agisiger_11   : -0.98  2.63 -0.14  0.05 -0.30 (1)
11_rjiang_2      : -1.45  2.57  0.34 -0.71 -0.30 (4)
8_mfitzgerald_28 : -0.49  2.36  0.52 -0.93 -0.30 (6)
27_vivian.wong_35: -0.85  2.32  0.25 -0.37 -0.92 (7)
3_rjiang_2       : -1.46  3.62 -0.15 -0.40  0.94 (13)


Centroid 4  :  0.18  0.75  0.76 -0.37 -0.37
Var (z)  :  0.46  0.37  0.34  0.42  0.35
Mean  :  0.45  0.10  0.31  2.48  0.88
Var  :  0.02  0.00  0.01  0.25  0.92
Size: 439 of 2050
Male: 0.658 vs 0.597
VPs : 0.497 vs 0.500
When: 0.465 +/- 0.264

VPs:
13_bpellegr_16    :  0.39  0.93  0.78 -0.54 -0.30 (2)
22_shourong_2     :  0.34  0.67  0.41 -0.57 -0.30 (4)
11_mcnerd_46      :  0.27  0.72  0.57 -0.78 -0.30 (5)
51_rgreenblatt_28:  0.05  0.50  0.64 -0.90 -0.30 (9)
15_arnoldos_13    :  0.14  0.83  0.92 -0.15 -0.92 (10)

MMs:
11_heaps_46       :  0.08  0.57  0.72 -0.15 -0.30 (1)
2_gburke_33       : -0.10  0.89  0.82 -0.24 -0.30 (3)
1_cguo_15         :  0.22  0.29  0.82 -0.22 -0.30 (6)
```

```
7_jmsmith_41      :  0.16  1.13  1.01 -0.14 -0.30 (7)
29_mfitzgerald_28:  0.09  1.04  0.50 -0.80 -0.30 (8)


Centroid 5  : -0.07 -0.36 -0.34  0.17  1.91
Var (z)  :  0.58  0.29  0.33  0.65  0.85
Mean  :  0.40  0.03  0.11  2.90  4.56
Var  :  0.02  0.00  0.01  0.39  2.19
Size: 259 of 2050
Male: 0.668 vs 0.597
VPs : 0.514 vs 0.500
When: 0.581 +/- 0.304

VPs:
12_echow_1        : -0.01 -0.27 -0.43  0.63  2.19 (3)
14_jauerbock_33   : -0.54 -0.39 -0.57  0.06  1.56 (5)
6_neville_45      : -0.33 -0.12 -0.73 -0.03  1.56 (6)
10_sfox_7         : -0.41  0.08 -0.08  0.49  2.19 (8)
21_ocedar_18      : -0.48 -0.92 -0.28  0.22  2.19 (10)

MMs:
12_beckerley_1    : -0.12 -0.43 -0.27  0.50  2.19 (1)
10_kmcheng_9      : -0.26 -0.56 -0.36 -0.09  2.19 (2)
8_kmcheng_9       : -0.04 -0.92 -0.42  0.26  2.19 (4)
23_jmc_45         : -0.03 -0.43 -0.55  0.78  2.19 (7)
21_myfeng_18      : -0.33 -0.28 -0.92  0.42  2.19 (9)


Centroid 6  :  1.06 -0.04  2.96 -0.29 -0.22
Var (z)  :  0.82  0.56  5.81  0.43  0.45
Mean  :  0.62  0.05  0.71  2.54  1.13
Var  :  0.03  0.00  0.19  0.26  1.15
Size: 72 of 2050
Male: 0.583 vs 0.597
VPs : 0.583 vs 0.500
When: 0.425 +/- 0.294

VPs:
11_shourong_2     :  1.48  0.34  2.57 -0.69 -0.30 (2)
24_jfernandez_11 :  1.44 -0.14  2.63  0.37 -0.30 (3)
8_rgreenblatt_28 :  0.90  0.52  2.36 -0.22 -0.30 (4)
3_shourong_2      :  1.51 -0.15  3.62 -0.36  0.32 (5)
```

```
3_ocedar_18        :  0.76 -0.18  2.56 -0.81 -0.92 (6)

MMs:
13_vivian.wong_35:  0.92 -0.54  2.65 -0.55 -0.30 (1)
10_mfitzgerald_28:  0.34  0.37  2.52 -1.03 -0.30 (7)
27_mfitzgerald_28:  1.09 -0.02  2.25 -0.99 -0.92 (9)
2_nancyhu_26      :  0.55 -0.78  2.10 -0.19 -0.30 (10)
7_rjiang_2        :  0.67  0.06  2.09 -1.16 -0.30 (11)


Centroid 7  : -0.92 -0.33 -0.39 -0.16 -0.34
Var (z)  :  0.27  0.24  0.20  0.41  0.33
Mean  :  0.24  0.04  0.10  2.64  0.95
Var  :  0.01  0.00  0.01  0.25  0.84
Size: 614 of 2050
Male: 0.552 vs 0.597
VPs : 0.463 vs 0.500
When: 0.578 +/- 0.273

VPs:
28_sfox_7          : -0.88 -0.33 -0.34  0.06 -0.30 (1)
10_djackson_41     : -0.65 -0.36 -0.46 -0.37 -0.30 (3)
5_hlindell_40      : -0.87 -0.40 -0.27 -0.49 -0.30 (4)
3_olivas_10        : -1.12 -0.57 -0.43  0.08 -0.30 (5)
31_tgojo_8         : -0.76  0.00 -0.58 -0.26 -0.30 (6)

MMs:
35_kkh_30          : -1.09 -0.51 -0.24  0.00 -0.30 (2)
10_jmsmith_41      : -0.96 -0.46 -0.36  0.28 -0.30 (7)
9_gfung_4          : -0.87  0.06 -0.30 -0.40 -0.30 (8)
8_ehoff_23         : -0.64 -0.13 -0.32 -0.48 -0.30 (9)
9_pmingardi_43     : -0.61 -0.46 -0.37  0.19 -0.30 (10)
```

9 Appendix 2: VP and MM Transition Matrices

9.1 Transition Matrix for VPs

$$(i,j) = P(h_{VP}(t+1) = j | h_{VP}(t) = i) \tag{8}$$

From↓ To→	1	2	3	4	5	6	7
1	0.2444	0.2444	0.0222	0.0222	0.2222	0	0.2444
2	0.0424	0.3347	0.0169	0.2161	0.1186	0.0339	0.2373
3	0.0270	0.1892	0.0541	0.4865	0.0270	0.0811	0.1351
4	0.0138	0.1972	0.0734	0.3440	0.0780	0.0459	0.2477
5	0.1102	0.2373	0.0254	0.1356	0.1949	0.0254	0.2712
6	0	0.0882	0.1765	0.5000	0	0.0294	0.2059
7	0.0240	0.1884	0.0137	0.1370	0.1438	0.0171	0.4760

9.2 Transition Matrix for MMs

$$(i,j) = P(h_{MM}(t+1) = j | h_{MM}(t) = i) \tag{9}$$

From↓ To→	1	2	3	4	5	6	7
1	0.3671	0.1646	0	0.0506	0.1772	0.0253	0.2152
2	0.0563	0.3099	0.0376	0.1596	0.1315	0.0235	0.2817
3	0	0.1714	0.0286	0.4571	0	0.2000	0.1429
4	0.0144	0.1971	0.0481	0.3269	0.0673	0.0625	0.2837
5	0.1043	0.2174	0.0261	0.1478	0.2174	0.0087	0.2783
6	0.0526	0.1053	0.0789	0.4474	0.0526	0.0789	0.1842
7	0.0582	0.2089	0.0240	0.1781	0.1267	0.0205	0.3836

10 Appendix 3: Cluster Graphs

Here are graphs of the clusters in the feature space. On the x-axis is talking time P, on the y-axis is the net influence $I_{so} - I_{os}$ and on the z-axis it the stressors minus the shorts $\sigma_e + \sigma_f + \sigma_s - S$.

Cluster	Color
1	red
2	cyan
3	yellow
4	black
5	blue
6	magenta
7	green

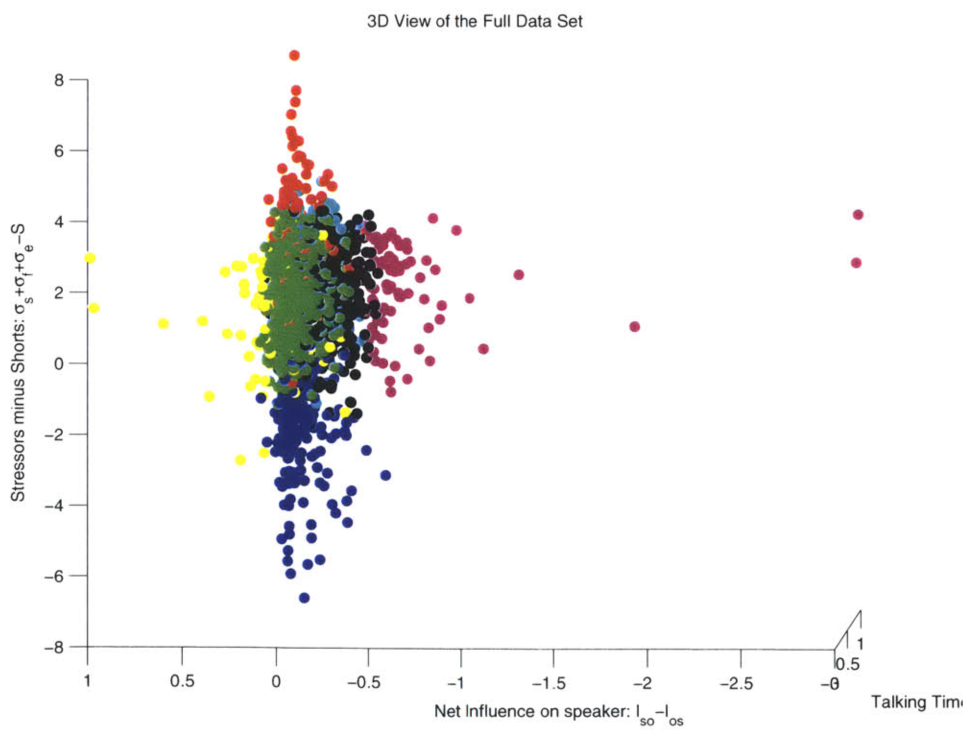

Figure 1: A plot of all of the clusters, looking down the talking time (P) axis.

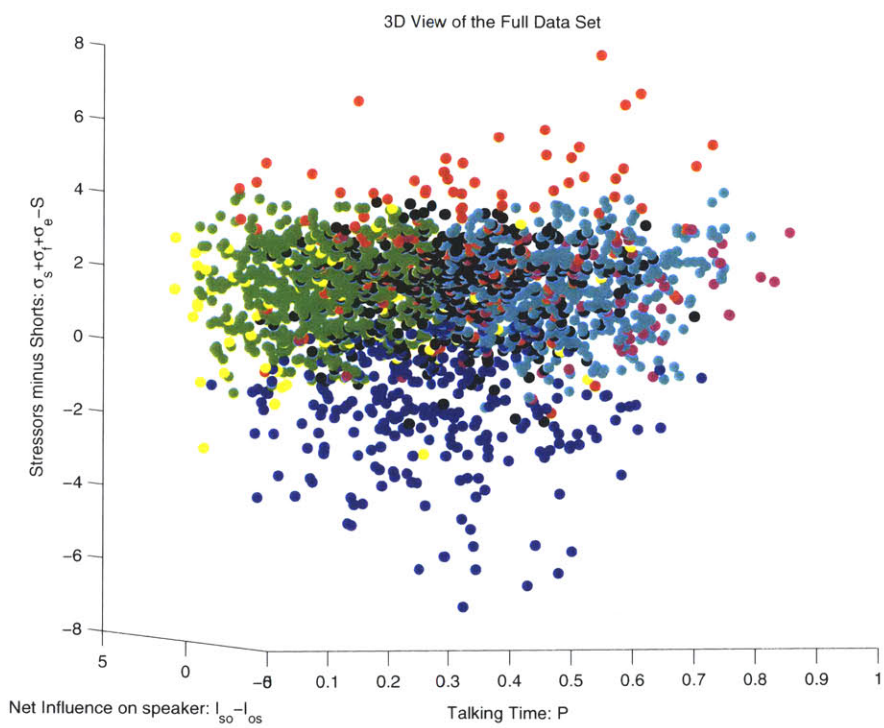

Figure 2: A plot of all the clusters looking down the net influence axis $(I_{so} - I_{os})$.

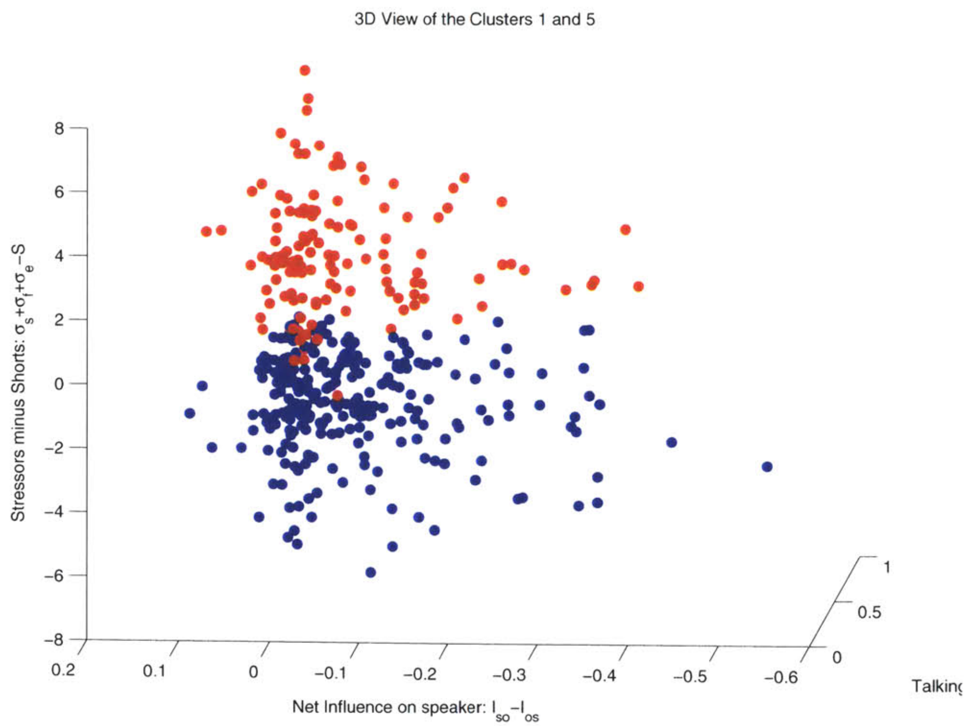

Figure 3: A plot of clusters 1 (red) and 5 (blue) looking down the talking time (P) axis. It seems that cluster 1 and cluster 5 are split along the stressor minus shorts axis.

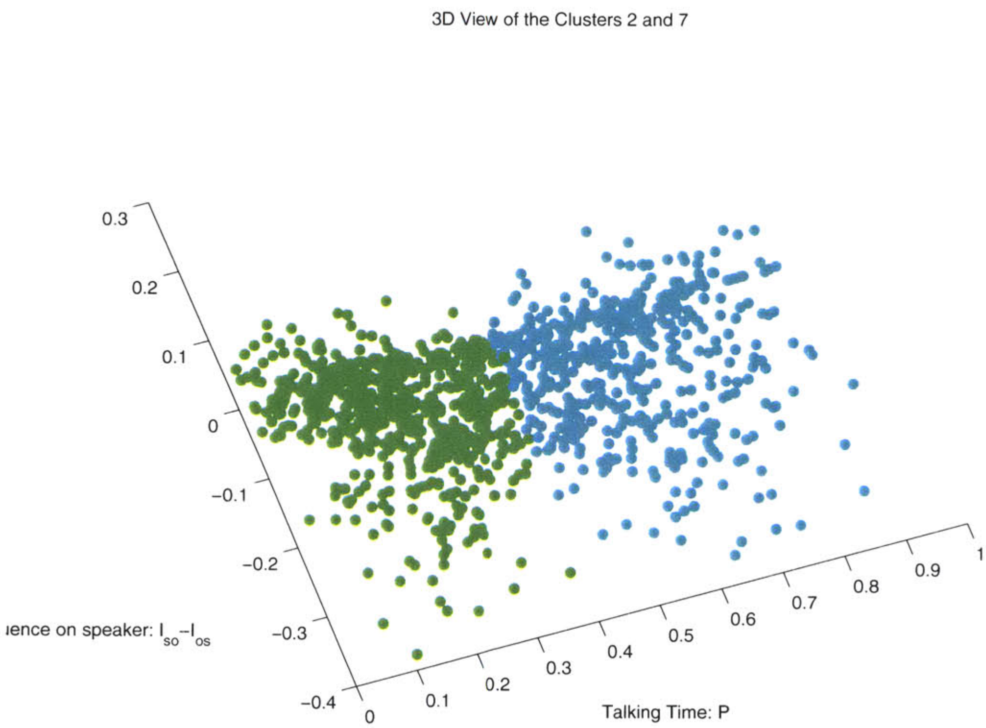

Figure 4: A plot of cluster 2 (cyan) and 7 (green) looking down the stressor minus shorts axis. Clusters 2 and 7 split along the talking time axis.

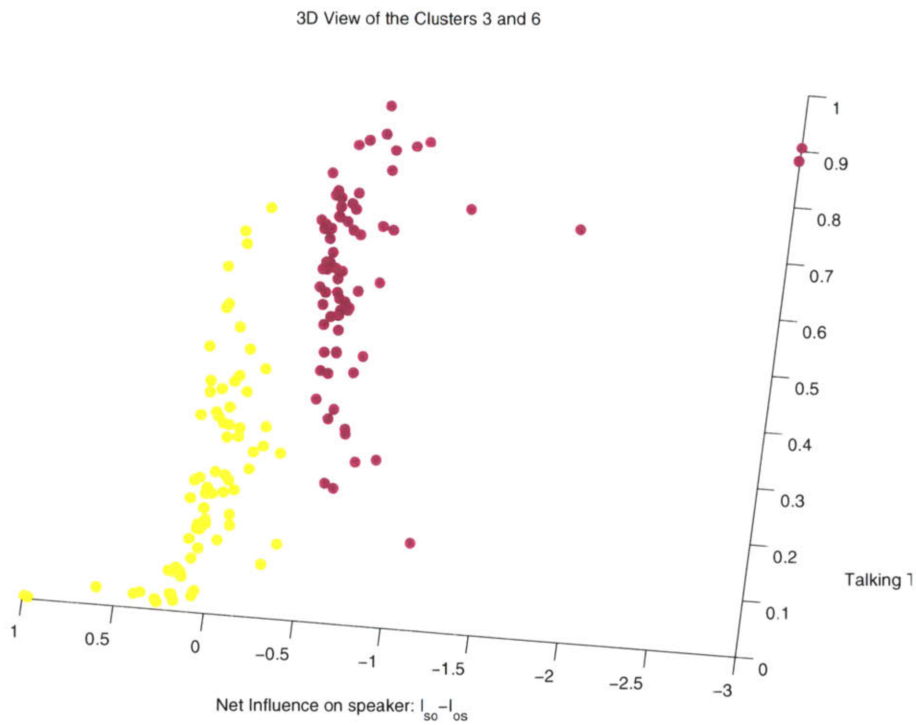

Figure 5: A plot of clusters 3 (yellow) and 6 (magenta) looking down the stress minus shorts axis. Clusters 3 and 6 split along the net influence axis.

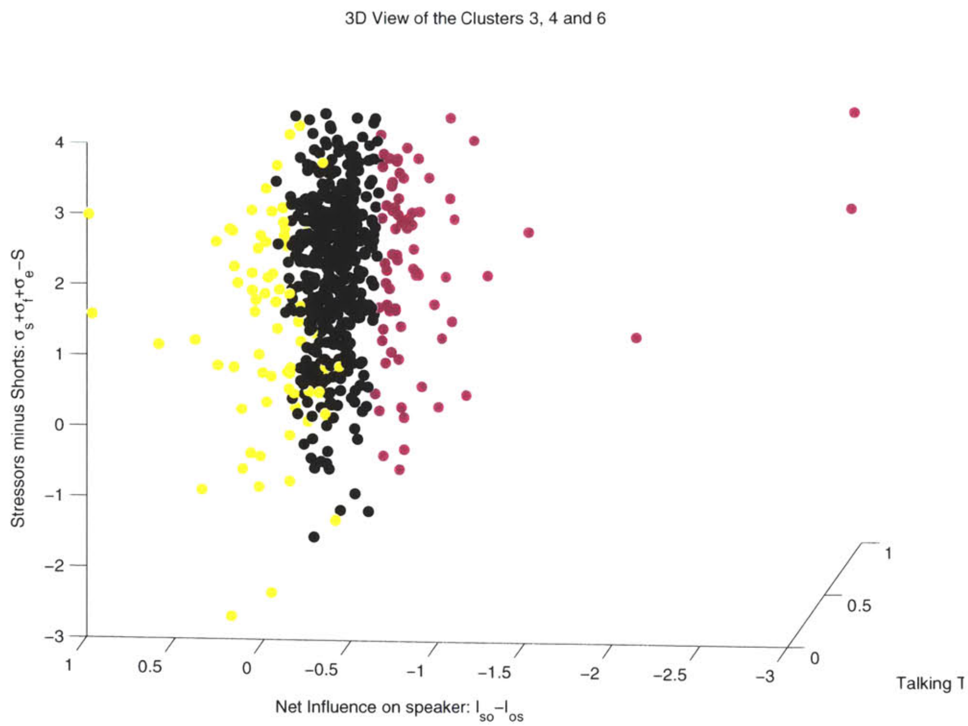

Figure 6: A plot of clusters 3 (yellow), 4 (black), and 6 (magenta) looking down the stress minus shorts axis.