Influence and Power in Group Interactions

Tomek STRZALKOWSKI[1], *Samira SHAIKH*[1], *Ting LIU*[1], *George Aaron BROADWELL*[1], *Jenny STROMER-GALLEY*[1], *Sarah TAYLOR*[2], *Veena RAVISHANKAR*[1], *Umit BOZ*[1] *and Xiaoai REN*[1]

(1) State University of New York – University at Albany, NY 12222 USA

(2) Sarah M. Taylor Consulting, LLC

tomek@albany.edu

Abstract. In this article, we present a novel approach towards the detection and modeling of complex social phenomena in multiparty interactions, including leadership, influence, pursuit of power and group cohesion. We have developed a two-tier approach that relies on observable and computable linguistic features of conversational text to make predictions about sociolinguistic behaviors such as Topic Control and Disagreement, that speakers deploy in order to achieve and maintain certain positions and roles in a group. These sociolinguistic behaviors are then used to infer higher-level social phenomena such as Influence and Pursuit of Power, which is the focus of this paper. We show robust performance results by comparing our automatically computed results to participants' own perceptions and rankings. We use weights learned from correlations with training examples to optimize our models and to show performance significantly above baseline.

Keywords. computational sociolinguistics, online dialogues, social phenomena, linguistic behavior, influence, pursuit of power, multi-disciplinary artificial intelligence, social computing

1 Introduction and Related Work

Our objective is to model high-level sociolinguistic phenomena such as Influence, Pursuit of Power, Leadership and Group Cohesion in discourse. This research aims to develop a computational approach that uses linguistic features of conversational text to detect and model sociolinguistic behaviors of conversation participants in small group discussions. Given a representative dialogue of multi-party conversation, our prototype system automatically classifies the participants by the degree to which they engage in such sociolinguistic behaviors as Topic Control, Disagreement, and several others discussed in this paper. These mid-level sociolinguistic behaviors are deployed by discourse participants in order to assert and maintain certain higher-level social roles such as leader, influencer, or pursuit of power, among others. Our approach to this problem combines robust computational linguistics methods and established empirical social science techniques. In this paper, we discuss robust detection of influence and pursuit of power in discourse.

Following social science theory (e.g., [11], [4], [7]), we define an *influencer* as a group participant who has credibility in the group and who introduces ideas that others pick up on or support. Thus, an influencer exercises a degree of control over the topic and content of a conversation. This may be contrasted with the more explicit *pursuit of power* in the group, which is an attempt to seize control of the group's agenda or actions. Both behaviors often correlate with, and may be considered as dimensions of, group leadership ([4, 16]); however, they are manifested quite differently in interaction. Neither behavior necessarily implies leadership, and often may be seen as a challenge to the group leader. In order to fully understand the social dynamics of a group it is important to model these behaviors independently. This is the focus of our paper, and we are particularly interested in groups involved in online interactions.

Internet-enabled interaction is particularly interesting to study because in this reduced-cue environment, the only means of engaging in and conveying social behaviors is through written language. As such, online discussion relies on the more explicit linguistic devices to convey social and cultural nuances than is typical in face-to-face or telephonic conversations. Relevant recent research in this area includes Freedman et al. ([6]) who developed an approach to detect behaviors such as persuasion in online discussion threads, and Bracewell et al. ([2]) who categorize several types of social acts (e.g. agreement and disagreement) to detect pursuit of power in online groups. These approaches, however, depend on discovering specific linguistic markers that may indicate a type of behavior rather than looking for a more sustained demonstration of sociolinguistic behavior by each speaker over the course of entire discourse. Our research takes that latter approach, and the work presented here builds on Strzalkowski et al. ([15]) and Broadwell et al. ([3]), who also proposed the two-tiered approach to sociolinguistic modeling and have demonstrated that a subset of mid-level sociolinguistic behaviors may be accurately inferred by a combination of low-level language features. Our work successfully extends their approach to modeling of influence and pursuit of power in group interactions. The models discussed in this paper were developed and implemented based on online chat and threaded discussions in English and Mandarin.

2 Sociolinguistic Behaviors in Discourse

In the two-tier modeling approach, we use linguistic elements of discourse to first unravel sociolinguistic behaviors, and then, use the behavior models, in turn, to determine complex social roles, as shown in Figure 1.

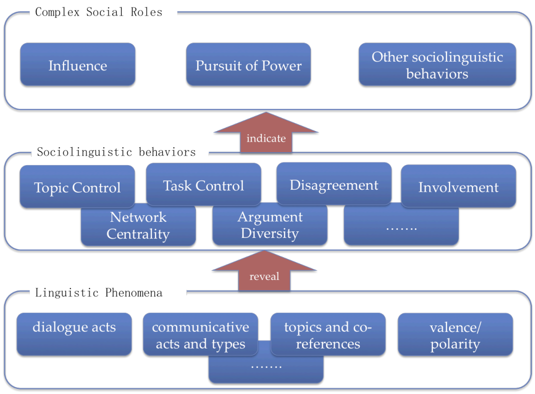

Fig. 1. Two-tier approach applied to model social roles in discourse.

It is important to note that, at both these levels, our analyses are solidly grounded in sociolinguistic theory. In this section, we briefly describe the mappings (which we call *measures*), between the sociolinguistic behaviors (mid-level in Fig 1) and the complex social roles. These measures operationalize the second tier in our system. We discuss the first tier only briefly in Section 4.

- **Topic Control Measure (TCM)** is defined as attempts of participants to impose a topic of conversation. In any conversation, whether it is focused on a particular issue or task or is just a social conversation, the participants continuously introduce multiple topics and subtopics. These are called *local topics*. Local topics, following the notion put forth by Givon ([8]), may be equated with any substantive noun phrases introduced into discourse that are subsequently mentioned again via repetition, synonym, or pronoun. Who introduces local topics into conversation and who continues to talk about them, and for how long are some of the indicators of topic control in dialogue.
- **Cumulative Disagreement Measure (CDM)** has a role to play with regard to influence and pursuit of power in that it is possible that a person in a small group engages in disagreements with others in order to control the topic by way of identifying or correcting what they see as a problem ([5], [13]). While each utterance where a participant disagrees with another is a vivid of expression of disagreement, we are interested in a sustained phenomenon where participants repeatedly disagree, thus revealing a social relationship between them.
- **Involvement Measure (INVX)** is defined as a degree of engagement or participation in the discussion of a group. A degree of involvement may be estimated by how much a speaker contributes to the discourse in terms of substantive content. Contributing substantive content to discourse includes introduction of new local topics, taking up the topics introduced by others, as well as taking sides on the topics being discussed. As previously defined, a local topic is a concept or a thing or an event referred to in a conversation, and is typically introduced with a noun phrase (a name, noun or pronoun).

- **Network Centrality Measure (NCM)** is the degree to which a participant is a "center hub" of the communication within the group. In other words, someone whom most others direct their comments to as well as whose topics are most widely cited by others.
- **(Measure of) Argument Diversity (MAD)** is displayed by the speakers who deploy a broader range of arguments in conversation. This behavior is signaled by the use of more varied vocabulary, including specialized terms and citations of authoritative sources, among others. A person who uses more varied vocabulary and introduces more unique words into a conversation is considered to have a higher degree of Argument Diversity.
- **Tension Focus Measure (TFM)** is defined as the degree to which a speaker is someone at whom others direct their disagreement, or with whose topics they disagree the most. Similar to Network Centrality Measure, Tension Focus reflects the actions of other members of the group towards the speaker.
- **Social Positioning (SPM)** is a sociolinguistic behavior defined as a degree to which the speaker attempts to position oneself as central in the group by committing to some future activity and by getting others to confirm or re-affirm what the speaker stated, as well as what the speaker already believes. This behavior is reflected in the speaker's conversational moves that aim at increasing their centrality in the group.
- **Involved Topic Control Measure (ITCM)** is defined with combined measure of Topic Control <u>and</u> Involvement. This measure identifies speakers who are both highly involved as well as control topics under discussion to a greater extent than other speakers in the group.

By computing the score of each *measure*, we obtain a full ranking of participants on each sociolinguistic behavior. The measures used to compute Pursuit of Power are ITCM, CDM, TFM and SPM. Measures used to compute Influence are TCM, CDM, NCM and MAD. We shall elaborate this further in Section 5.

3 Correlations between the measures and human assessment

We computed the correlations among our proposed measures of Influence and Pursuit of Power. As described before, Topic Control Measure (TCM), Cumulative Disagreement Measure (CDM), Network Centrality Measure (NCM) and Measure of Argument Diversity (MAD) are calculated for Influence. Involved Topic Control Measure (ITCM), Cumulative Disagreement Measure (CDM), Tension Focus Measure (TFM), and Social Positioning Measure (SPM) are combined to predict Pursuit of Power.

Table 1 shows the correlations between all proposed Influence measures and proposed Pursuit of Power measures. We note that Cumulative Disagreement Measure correlations are lower than the other measures for influence, pointing to evidence of it being the discriminant variable. We have observed similar correlation patterns across the sessions we have looked at. Computing the correlation against human rankings elicited using survey questionnaire provides us with evidence that indeed the proposed behaviors are measuring the correct phenomena. The correlation between rankings produced by annotated data and ranking induced by participant ratings holds quite strongly across a significant proportion of data sets in our corpus with an average of over 0.80 Cronbach's alpha.

Influence	NCM	MAD	TCM	CDM
NCM	1.0			
MAD	0.86	1.0		
TCM	0.98	0.86	1.0	
CDM	0.58	0.59	0.48	1.0

Pursuit of Power	ITCM	CDM	TFM	SPM
ITCM	1.0			
CDM	0.78	1.0		
TFM	0.67	0.91	1.0	
SPM	0.7	0.88	0.67	1.0

Table 1. – Correlation among measures of Influence and Pursuit of Power for a sample chat dialogue

Using this evidence of high correlations among behaviors and their measures, as well as measures against human survey ratings, we can be confident about our approach in measuring and detecting Influence and Pursuit of Power. We demonstrate our evaluation and results in section 5.

4 Linguistic Components on the First Tier

The first tier of our system comprises a series of basic linguistic components that support models for computing sociolinguistic behaviors in the mid-level of Fig. 1. These models are derived from a corpus of online dialogues that includes chat and threaded discussions. The chat portion of the corpus, the MPC chat corpus is described in ([14], [10]), consists of over 90 hours of online chat dialogues in English and Mandarin. The threaded discussions portion of the data consists of 70 asynchronous thread discussions collected from public Wikipedia discussions in English and Mandarin. A substantial subset of the corpus was annotated using trained annotators who are native speakers of the respective language. Annotators were trained extensively so that inter-annotator agreement level was sufficiently high (0.8 or higher Krippendorf's alpha). The annotated data was used to train the linguistic components and to calibrate mappings from linguistic components to sociolinguistic behaviors that constitute the first tier of our system. Space limitations do not allow us to discuss the first tier mappings here; instead, we briefly describe the key linguistic components involved: (1) communication links between participants, (2) dialogue acts, and (3) local topics in conversation.

Communication Links

It is important and very challenging to determine automatically who speaks to whom in multiparty discourse. In our annotation process, we ask annotators to classify each utterance in the chat by marking it as either a) addressed to someone or everyone; b) a response to someone else's specific prior utterance; or c) a continuation of one's own prior utterance. Using annotated data from this layer of annotation, we can train a communication link classification module, which uses context, inter-utterance similarity and proximity of utterances as some of the features in a Naïve Bayes classifier to automatically classify utterances in one of the above-mentioned three categories. The current performance of this module is 61% accuracy as measured against annotated ground truth data.

Dialogue Acts

We have developed a hierarchy of 15 dialogue acts in order to annotate the functional aspect of an utterance in discourse. The tag set adopted is based on DAMSL ([1]) and SWBD-DAMSL ([9]), but compressed to 15 tags tuned towards dialogue pragmatics and away from more surface characteristics of utterances. A detailed description of dialogue act tags and annotation procedure has been described in a separate publication. Some dialogue acts that are note-worthy are: Disagree-Reject, Offer-Commit, Confirmation-Request, Assertion-Opinion. Annotated data from this process is used to train a cue-phrase based dialogue act classifier adapted from Webb and Ferguson's ([18]) approach, which currently performs at 64% accuracy. Our Cumulative Disagreement Measure (CDM) is calculated using the proportion of disagreement dialogue act utterances detected for each participant by this automatic module.

Local Topics

Local topics are defined as nouns or noun phrases introduced into discourse that are subsequently mentioned again via repetition, synonym, or pronoun. Annotators were asked to mark all nouns and noun phrases of import from the discussion. We use Stanford part-of-speech tagger ([17]) to automatically detect nouns from text. Princeton's Wordnet ([12]) is consulted to identify synonyms commonly used in co-references. Since POS taggers are typically trained on well-formed text, performance of POS tagging on chat text – where grammar may be disorganized, use of abbreviations and symbols etc. may be quite frequent – would affect the accuracy of POS tagging. Our automatic local topic detection module performance is at 70% in the current system prototype.

We note here that it is not the goal of this research to develop the best computational modules such as POS taggers or the most accurately performing dialogue act or communication link classifier. In spite of the shortcomings in modules that support our index calculations, we are able to achieve very robust performance in our intended task of modelling complex social roles. This is because we base our claims of sociolinguistic behavior on repeated counts of each linguistic phenomenon over the length of entire discourse. When computational modules such as local topic detection fail, such errors are systematic, and would be replicated consistently for each participant. If the count for each participant were not fully accurate, nevertheless, the *distribution* of counts for all participants would still hold, thus giving us the desired ranking or the degree of sociolinguistic behavior for each participant.

Having multiple indices for each behavior helps us account for the error introduced from our automatic modules. If the predictions on individual indices are not always consistent, we can still combine them into a single output by using different weighting schemes, albeit with lesser confidence. In order to validate our proposed indices and measures, we analyzed their correlation with each other, both from human annotated data as well as our automatic process, as we shall discuss next.

5 Evaluation and Results

We compute the scores for each participant for all proposed measures. Although we have a full ranking of participants, both from survey ratings, human assessment as well as system output, we are only interested in participants who have the highest Influence and Pursuit of Power. This means, the top-ranking participant on both rankings should match in order evaluate system performance. In cases where the top two individuals are quite close in the survey scores, we may consider top two participants.

We calculate the Influence and Pursuit of power score for all participants by taking the mean of our measures and deriving an Influencer and Pursuit of Power score for each participant. We devised a weighting scheme that reflects the evidence found from our analysis of correlations against survey ratings. So, the weighting scheme for English dialogues is:

Influencer score = (α_{TCM}* TCM) + (α_{CDM}* CDM) + (α_{NCM}* NCM) + (α_{MAD}* MAD)

Where $\alpha_{TCM} > \alpha_{NCM} > \alpha_{MAD} > \alpha_{CDM}$

PoPScore = α_{ITCM}*ITCM + α_{CDM}*CDM + α_{TFM}*TFM + α_{SPM}*SPM

Where: $\alpha_{SPM} < \alpha_{ITCM} = \alpha_{TFM} < \alpha_{CDM}$

Similar combinations are derived for Chinese chat dialogues as well. For different data types and different languages, we have learnt different weighting schemes where the sociolinguistic behaviors may be combined differently to compute scores. In essence, the higher the correlation, the greater the weight given to the measure. As expected and shown in Table 2, different correlations hold across languages – hence possibly cultures, since the participants are native speakers. Where the scores are 0, it signifies that the behavior is found to not correlate well with the other measures that comprise the phenomena being modeled; hence we do not include these behaviors while taking linear combinations.

Weights	TCM	ITCM	CDM	MAD	TFM	SPM	NCM
English Influence	0.75	0	0.15	0.4	0	0	0.5
Chinese Influence	0.75	0	0.1	0.75	0	0	0.4
English PoP	0	0.1	0.6	0	0.1	0.09	0
Chinese PoP	0	0.08	0.08	0	0.84	0	0

Table 2. Weighting schemes for combining social behaviors learnt from correlation analyses

Table 3 shows performance accuracy of our automated system in detecting the top influencers and those who pursue power in group dialogues, computed across languages and compared to a random selection baseline. We could choose another baseline, such as selecting the participant with the most turns as an Influencer, etc.; however, we see similar performance for such baselines as the random one. The unweighted model represents the first approximation, untrained option of the system.

Performance	BaseLine	Without Weight	With Weight
English Influence	17.85%	71.4%	78.5%
Chinese Influence	12.5%	69%	90%
English PoP	10%	52.6%	84.2%
Chinese PoP	5%	60%	73.3%

Table 3. Performance accuracy against random baseline, with and without weighting scheme

6 Conclusion and perspectives

We have shown a novel, robust method for modeling social phenomena in multi-party discourse. We have combined established social science theories with computational modeling to create a two-tier approach that can detect high-level sociolinguistic phenomena such as Influence and Pursuit of Power in language with a high degree of accuracy. In future work, we have planned for a larger scale evaluation, testing index stability, and resilience to errors in automated language processing, including topic detection, coreference resolution, and dialogue act classification. Current performance of the system is based on versions of these linguistic modules, which perform at about 70% accuracy, so these need to be improved as well.

The advantage of applying a two-tier approach is that we can add or remove mid-level sociolinguistic behaviors efficiently when applying our models to different data types and languages. This would be impractical in a straightforward machine-learning approach where one can add all features to a learning algorithm to decide how features may best be combined. A machine-learning approach modeled directly on linguistic features would not be easily transferable to other data types and could prove brittle. Some measures turn out to be more predictive in a given data genre, and when applied appropriately, perform well at predicting phenomena as rated and understood by human assessors. We note that there may be some variance as to how humans perceive the concept of Influence and Pursuit of Power and rate a participant based on their intuitive notion of the concept. The fact that we have multiple indicators in the form of measures helps us overcome the potential variance in this perception.

Acknowledgement: This research was funded by the Office of the Director of National Intelligence ODNI), Intelligence Advanced Research Projects Activity (IARPA), and through the U.S. Army Research Lab. All statements of fact, opinion or conclusions contained herein are those of the authors and should not be construed as representing the official views or policies of IARPA, the ODNI or the U.S. Government.

References

1. Allen, J. and M. Core. (1997). Draft of DAMSL: Dialog Act Markup in Several Layers. www.cs. rochester.edu/research/cisd/resources/damsl/
2. Bracewell, D. B., Tomlinson, M., and Wang, H. “A Motif Approach for Identifying Pursuit of Power in Social Discourse,” in Proceedings of *the 6th IEEE International Conference on Semantic Computing* (ICSC 2012), 2012.

3. Broadwell, G., Stromer-Galley, J., Strzalkowski, T., Shaikh, S., Taylor, S., Liu T., Boz, U., Elia, A., Jiao, L., and Webb, N. (2012). Modeling SocioCultural Phenomena in Discourse. *Journal of Natural Language Engineering, Cambridge Press.*
4. Bradford, L. P. (1978). Group development. La Jolla, Calif., University Associates.
5. Ellis, D. G., & Fisher, B. A. (1994). *Small group decision making: Communication and the group process*. New York: McGraw-Hill
6. Marjorie Freedman, Alex Baron, Vasin Punyakanok and Ralph Weischedel. (2011). "Language Use: What it can tell us?" In *Proceedings of the Association of Computational Linguistics.* Portland, Oregon.
7. French, J. R. P., and Raven, B. "The Bases of Social Power," in *Studies in Social Power*, ser. Studies in social power, D. Cartwright, Ed. Institute for social research, 1959, vol. 35, ch. 9, pp. 150–167.
8. Givon, T. (1983). *Topic continuity in discourse: A quantitative cross-language study.* Amsterdam: John Benjamins.
9. Jurafsky, D., E. Shriberg and D. Biasca. (1997). Switchboard SWBD-DAMSL Shallow-Discourse-Function Annotation Coders Manual. http://stripe.colorado.edu/~jurafsky/manual.august1.html
10. Liu, T., Samira Shaikh, Strzalkowski, T., Broadwell, A., Stromer-Galley, J., Taylor, S., Boz, Umit., Ren, Xiaoai. and Jingsi Wu. (2012) Extending the MPC corpus to Chinese and Urdu-Multi-party Multi-lingual Chat Corpus for Modeling Social Phenomena in Language. In *Proceedings of the Eight International Conference on Language Resources and Evaluation (LREC'12).*
11. McClelland, D. C. 1987. *Human motivation*. Cambridge University Press, 1987, vol. 23.
12. Miller, G. A., Beckwith, R., Fellbaum, C. D., Gross, D., and Miller, K. (1990). WordNet: An online lexical database. *Int. J. Lexicograph*. 3(4): 235--244.
13. Sanders, R. E., Pomerantz, A., Stromer-Galley, J. (2010). Some ways of taking issue with another participant during a group deliberation. Paper presented at the annual meeting of the National Communication Association, San Francisco, CA
14. Shaikh, S., T. Strzalkowski, S. Taylor and N. Webb. (2010). MPC: A Multi-Party Chat Corpus for Modeling Social Phenomena in Discourse, in proceedings of the 7th International Conference on Language Resources and Evaluation (LREC2010), Valletta, Malta. 2010.
15. Strzalkowski, T., Broadwell, G. A., Stromer-Galley, J., Shaikh, S., Taylor, S., & Webb, N. (2010) Modeling socio-cultural phenomena in discourse. In *Proceedings of The 23rd International Conference on Computational Linguistics. Beijing, China.*
16. Strzalkowski, T., Shaikh, S., Liu, T., Broadwell, G. A., Stromer-Galley, J., Taylor, S., Ravishankar, V., Boz, U., & Ren, X. (2012) Modeling Leadership and Influence in Multi--party Online Discourse. In Proceedings of *The 24rd International Conference on Computational Linguistics*. Mumbia, India.
17. Kristina Toutanova, Dan Klein, Christopher Manning, and Yoram Singer. (2003). Feature-Rich Part-of-Speech Tagging with a Cyclic Dependency Network. In *Proceedings of HLT-NAACL 2003*, pp. 252-259.
18. Webb, N. and M. Ferguson. (2010). Automatic Extraction of Cue Phrases for Cross-Corpus Dialogue Act Classification, in the proceedings of the 23rd International Conference on Computational Linguistics (COLING-2010), Beijing, China. 2010