A Multimodal Corpus for Studying Dominance in Small Group Conversations

Oya Aran[1], Hayley Hung[1], and Daniel Gatica-Perez[1,2]

[1]Idiap Research Institute, Martigny, Switzerland
[2]Ecole Polytechnique Federale de Lausanne (EPFL), Switzerland
{oya.aran, hhung, gatica}@idiap.ch

Abstract

We present a new multimodal corpus with dominance annotations on small group conversations. We used five-minute non-overlapping slices from a subset of meetings selected from the popular Augmented Multi-party Interaction (AMI) corpus. The total length of the annotated corpus corresponds to 10 hours of meeting data. Each meeting is observed and assessed by three annotators according to their level of perceived dominance. We analyzed the annotations with respect to dominance, status, gender and behavior. The results of the analysis reflect the findings in the social psychology literature on dominance. The described dataset provides an appropriate testbed for automatic dominance analysis.

1. Introduction

Social interaction is a fundamental aspect of human life and is also a key research area in psychology and cognitive science. Although social psychologists have been researching the dimensions of social interaction for decades, the interest on the automatic analysis of social interaction, particularly small group conversations, is quite recent. It is an emerging field of research in several communities such as human computer interaction, machine learning, speech processing, and computer vision (Gatica-Perez, 2009; Pentland, 2005; Vinciarelli et al., 2009) and there is a crucial need for developing dedicated techniques and collecting necessary resources.

The social cues produced and exchanged during an interaction include verbal and nonverbal elements. In parallel to the verbal elements (the spoken words), the nonverbal information is conveyed as wordless messages through aural cues (voice quality, speaking style, intonation) and also through visual cues (gestures, body posture, facial expression, and gaze) (Knapp and Hall, 2009). These cues can be used to predict human behavior, personality, and social relations, in a very wide range of situations. It has been shown that, in many social situations, humans can correctly interpret nonverbal cues and can predict behavioral outcomes with high accuracy, when exposed to short segments or "thin slices" of expressive behavior (Ambady and Rosenthal, 1992). The length of these thin slices can change from a few seconds to several minutes depending on different situations.

Dominance is one of the fundamental dimensions of social interaction. It is signaled via both verbal and nonverbal cues. The nonverbal cues include vocalic ones such as speaking time (Schmid-Mast, 2002), loudness, pitch, vocal control (Dunbar and Burgoon, 2005b), turns, and interruptions (Smith-Lovin and Brody, 1989); and kinesic ones such as gesturing, posture, facial expressions, and gaze (Dovidio and Ellyson, 1982; Dunbar and Burgoon, 2005a). Dominant people are in general more active both vocally and kinesically, with an impression of relaxation and confidence (Hall et al., 2005; Burgoon and Dunbar, 2006). It has been shown that, they also have a higher visual dominance ratio (looking-while speaking to looking-while-listening ratio), i.e. they look at others more while speaking and less while listening (Dovidio and Ellyson, 1982).

There are a number of works in the literature that investigate techniques for the automatic estimation of dominance in small group conversations through nonverbal cues. (Rienks and Heylen, 2005) addressed the classification of participants dominance level (high, normal, low) and used a supervised approach based on Support Vector Machines with manually annotated audio nonverbal features. In (Jayagopi et al., 2009) a large number of automatically extracted nonverbal audio and visual activity cues were used to estimate the most dominant and least dominant participant. The difference in estimating the two dimensions of social verticality, dominance and status, is addressed in (Jayagopi et al., 2008). In (Hung et al., 2008), the authors investigated the use of visual attention cues for estimating dominance. A recent survey on the topic can be found in (Gatica-Perez, 2009). These initial works investigate the different features and models for the estimation of dominance. However for further advancement, there is a clear need for a large database, that can be used as a benchmark across different studies.

In this paper we present a new annotated multimodal dataset that can be used to assess dominance on small group conversations. The novelty of this dataset comes from the dominance annotations as the AMI meeting corpus is well known. In Section 2, we briefly describe the AMI meeting corpus. Section 3 details the dominance annotations and the experimental protocol. In Section 4, we present the resulting datasets and the estimation tasks. The detailed analysis of the annotations is given in Section 5.

2. Meeting Corpus

We use a subset of the publicly available Augmented Multi-party Interaction (AMI) corpus for this study (Carletta et al., 2005). The AMI meeting corpus includes two types of meetings: scenario meetings and non-scenario meetings. In scenario meetings, participants are given the task of designing a remote control over a series of sessions with roles assigned for each participant. One of the participants is the project manager who has the overall responsibility. These meetings are generally based on presentations followed by discussions. The participants are not always seated. It is common that one of the participants is presenting in front of the whiteboard or slide screen. In non-scenario meetings participants were free to choose their own topic beforehand. Participants are generally seated in these meetings. Each meeting has four participants.

Meetings in the AMI corpus were carried out in a multi-sensor meeting room as shown in Figure 1. The room contains a table for four participants, a slide screen, and a white board. The audio is recorded via several microphones: a circular microphone array placed on the table, another one with four microphones placed in the ceiling, headset and lapel microphones. The video is recorded via seven cameras: Three cameras mounted on the sides and back of the room capture mid range and global views, respectively; four cameras mounted on the table capture individual visual activity only. Example screen shots from the corpus, from each of the cameras are shown in Figure 2.

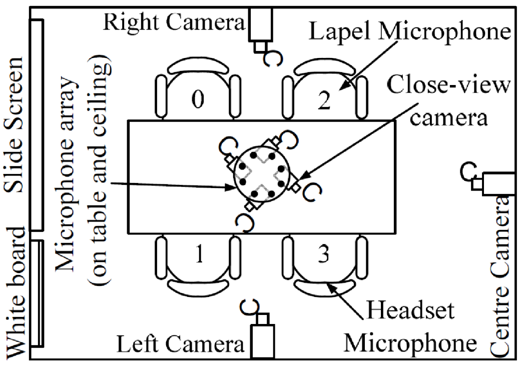

Figure 1: AMI meeting room setup.

Figure 2: AMI screen shots from seven available cameras. The top row shows the views from the right, center, and left cameras. The bottom row shows the views from the close up cameras.

3. Dominance Annotations

We collected a set of annotations on a subset of the meetings selected from the AMI corpus. We follow the "thin slice" approach and use five-minute meeting segments. Previous publications on dominance estimation on AMI corpus use a dataset that corresponds to 4.5 hours of recordings (Jayagopi et al., 2008; Jayagopi et al., 2009). We have enlarged the previously annotated data with a new set of annotations. With this new set, we double the size of our annotated dataset, which corresponds to more than 10 hours of recordings.

3.1. Annotation Questionnaire

The questionnaire asks the annotators about their perceived dominance of the meeting participants. There are two sections in the questionnaire: In the first section the annotators were asked questions on the participants' relative dominance; and in the second section, the questions are focused on evaluating each participant independently:

Dominance ranking: Each participant is ranked from 1 to 4, with 1 representing the most dominant person, and 4 representing the least dominant person in the meeting.

Dominance weight: 10 points are distributed among the participants reflecting annotator's impression of their relative dominance displayed during the meeting. More units signified higher dominance.

Confidence: To identify segments where the rankings were difficult to allocate, annotators were asked about their confidence in their rankings on a seven-point scale.

Participant characteristics: Annotators were requested to evaluate five specific characteristics of each participant independently: dominance (dominant/submissive), status (high/low), aggressiveness (aggressive/meek), dynamism (dynamic/passive), and talkativeness (talkative/silent), also on a seven-point scale. These questions were selected from social psychology work (Dunbar and Burgoon, 2005a).

3.2. Annotator Agreement

For each meeting segment, three annotators ranked the participants according to their level of perceived dominance. We then assessed the agreement between the three annotators for each meeting. If all annotators ranked the same participant as the highest (resp. lowest), we assume there is a full agreement on the most (resp. least) dominant person. If at least two annotators ranked the same participant as the highest (resp. lowest), we assume there is a majority agreement on the most (resp. least) dominant person. Following this procedure we obtained two annotated meeting datasets:

Meeting Set 1 (M1) The initial set of annotations is done on a total of 58 five-minute meeting segments with 21 independent annotators. The meetings were selected from the scenario meetings in AMI corpus. This set was previously used in several publications on automatic dominance estimation (Jayagopi et al., 2008; Hung et al., 2008; Jayagopi et al., 2009).

Meeting Set 2 (M2) We collected a new set of annotations with a completely new set of annotators. 21 annotators annotated a total of 67 five-minute meetings. The meetings were selected both from the scenario and non-scenario AMI meetings. Special care was taken to select segments where all participants were seated during the whole meeting.

Figure 3 shows the agreement statistics in M1 and M2 sets. The bars show the percentage of each type of agreement; three annotators agree (red/bottom), two annotators agree (green/middle), and no agreement (blue/top). The actual

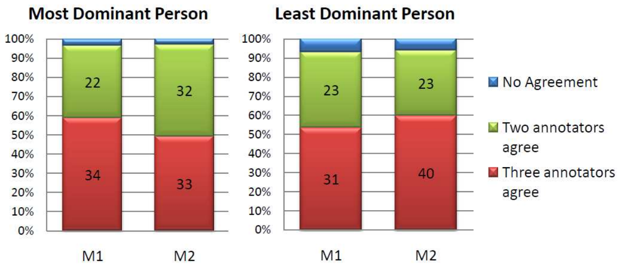

Figure 3: Distribution of agreement types in M1 and M2: Three annotators agree (red/bottom), two annotators agree (green/middle), and no agreement (blue/top)

number of meetings for each type of agreement is shown in the middle of the bars. On different meetings and with different sets of annotators, we observe similar agreement statistics: Full agreement is observed on around 50% of the meetings; whereas on almost all meetings, except a few, we observe majority agreement.

4. Experimental Protocol

4.1. Dominance Estimation Task

Following the recent work in (Jayagopi et al., 2009), we define two dominance estimation tasks:

1. Estimating the Most Dominant (MD) person: Among the participants of the meeting, we aim to estimate the most dominant person.

2. Estimating the Least Dominant (LD) Person: Among the participants of the meeting, we aim to estimate the least dominant person.

4.2. Datasets

The number of full and majority agreement meetings for MD and LD tasks for M1 and M2 sets, and also jointly, are summarized in Table 1. On the joint data, we define four datasets based on the dominance estimation tasks and annotator agreements. For each dataset, we also report the average annotator confidence (Conf - 1 being the highest, 7 being the lowest) and the average dominance weight of the agreed person (Weight - 10 being the highest, 1 being the lowest; all adding up to 10), as reported by the annotators:

FMD: **Full** agreement set, **most dominant** person estimation task (Conf: 1.85 - Weight: 4.57),

FLD: **Full** agreement set, **least dominant** person estimation task (Conf: 2.28 - Weight: 1.03)

MMD: **Majority** agreement set, **most dominant** person estimation task (Conf: 2.03 - Weight: 4.18),

MLD: **Majority** agreement set, **least dominant** person estimation task (Conf: 2.59 - Weight: 1.17)

The self-reported confidences show that the annotators gave higher confidence when annotating the most dominant person with respect to the least dominant one, indicating the latter can be a more difficult task. Furthermore, the full agreement datasets have higher self-reported confidence

	M1 (58)		M2 (67)		M1+M2 (125)	
	Full	Maj	Full	Maj	Full	Maj
MD	34	56	33	65	67	121
LD	31	54	40	63	71	117

Table 1: Number of meetings for tasks MD and LD with full and majority agreement in M1, M2, and jointly. The total number of meetings in each dataset is in brackets.

than the majority agreement datasets. The average relative weights assigned by the annotators also show the consistency of the dominance rankings.

5. Analysis of Annotations

5.1. Dominance and Status

Dominance and status are two aspects of the vertical dimension of human social interactions. Although related, these two concepts are different: dominance is a personality trait, which can be defined as the ability to control others; on the other hand, status is an achieved quality and does not directly relate to the ability to control (Hall et al., 2005).
In order to investigate this fact, we analyzed the relationship between the project manager, which is the highest status in the AMI meetings, and the dominance annotations. Figure 4 shows the project manager distribution among most/least dominant participants in full and majority agreement datasets (FMD, MMD, FLD, and MLD). It can be seen that only ~50% of the most dominant participants are also a project manager; whereas the number of least dominant participants who are also the project manager is extremely low. This shows the relation and also the difference between the concepts of dominance and status, as stated by social psychology: (1) high status is not a direct indicator of high dominance, (2) high status people are not totally submissive either.

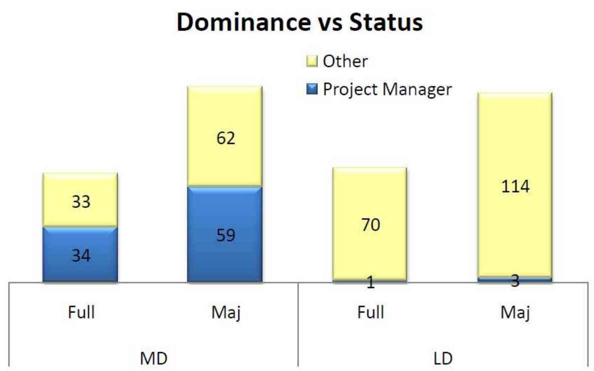

Figure 4: Distribution of project manager among most/least dominant participants in full and majority agreement datasets. Blue/bottom part shows the number of most/least dominant participants who are also the project manager. Yellow/top part shows the number of most/least dominant participants who have other roles than project manager.

5.2. Dominance and Gender

We also investigated our corpus to see the relationships between gender and dominance and gender and status.

Among the total meeting participants, the percentage of females is around 30% (156 females, 344 males). Among the project managers, it is around 50% (56 females, 69 males). We further investigated the distribution of gender for most dominant and least dominant participants. Figure 5 shows the number of males and females for most/least dominant participants in full and majority agreement datasets (FMD, MMD, FLD, and MLD), and also for the project manager (PM). It can be seen that for each case, the percentages of females and males are balanced (Percentages of females in FMD:52%, MMD:55%, FLD:56%, MLD:50%, and PM:45%).

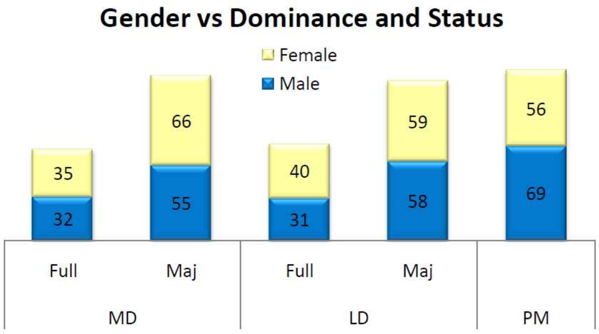

Figure 5: Gender distribution for most/least dominant participants infull and majority agreement datasets, and for project manager. Blue/bottom part shows the number of males and yellow/top part shows the number of females in each case.

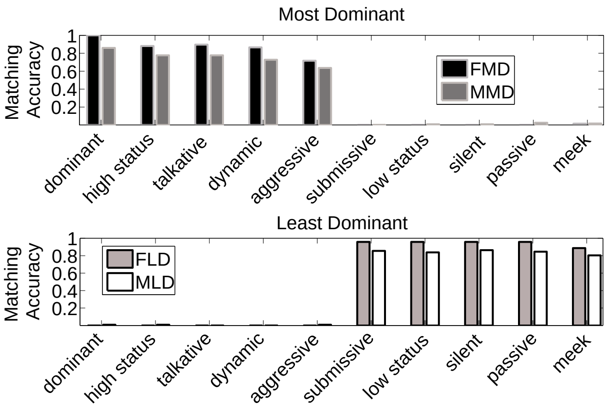

Figure 6: Person matching accuracy of the estimations based on participant evaluations with respect to most/least dominant participants for full an majority agreement datasets.

5.3. Participant Evaluations

We analyzed the participant evaluations in two aspects. The first analysis is based on comparing the participant selected with respect to the highest or lowest evaluation score in the questionnaire (e.g. talkative/silent, aggressive/meek) against the participant selected through the dominance rankings. This analysis aims to identify which of the participant characteristics are more related to perceptions of dominance. The second analysis is based on the correlations of the evaluation scores between the project manager and the most/least dominant person, which shows the relationship between these two concepts in more detail.

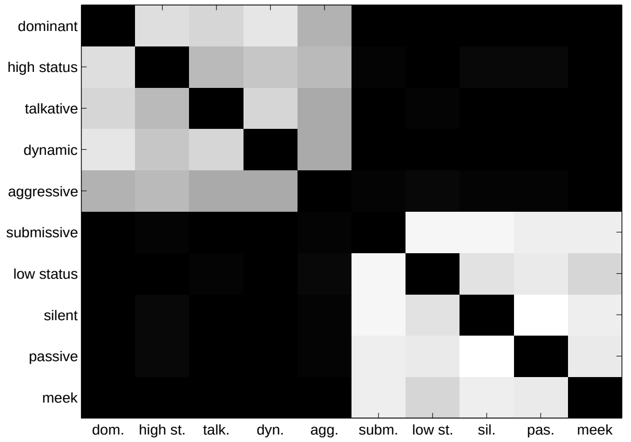

Figure 7: Person matching accuracy of the estimations based on participant evaluations across behaviors. The matrix is symmetric and in grayscale (black:0 and white:1).

For the first analysis, we computed the average of the evaluation scores for each participant by the three annotators for each of the five questions. For each question, based on the evaluation score, we can define two behavior types, one being the extreme opposite of the other. Then we select the most representative participant for each behavior, by choosing the participant with the highest or lowest average evaluation score of the related question. Taking the highest or the lowest value solely depends on which part of the seven-point scale that behavior is placed within the questionnaire. For example, in the question that asks for the dynamism of the participants, if the evaluation score is close to one, it indicates that the person is very dynamic, on the other hand if it is close to seven, it indicates passiveness.

For each behavior and for each dominance task, we count how often the person selected by each behavior was also labeled as the most or least dominant person. Figure 6 shows the person matching accuracies of the estimations based on participant evaluations with respect to the most and least dominant participants. We see that people highly scored as dominant, high status, talkative, dynamic and aggressive are more likely to be selected as most dominant, whereas people scored as submissive, low status, silent, passive and meek are more likely to be selected as the least dominant. Furthermore, Figure 7 shows a pairwise analysis across annotated behaviors. For each pair, we count how often the person selected by one behavior matches the person selected by the other behavior and calculated the person matching accuracy. In general, parallel behaviors highly match each other and low profile behaviors (e.g. silent/passive) have higher accuracies than high profile ones (e.g. dominant/dynamic). In addition, contrasting behaviors do not match at all, with accuracies very close to zero.

For the second analysis, we computed the Pearson correlation of the scores given in the five questions for the project manager and the most/least dominant person (Figure 8). In

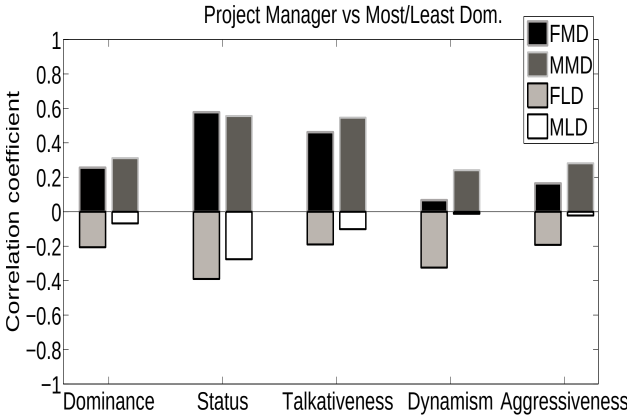

Figure 8: Pearson correlation of the scores of project manager and most/least dominant person for FMD, FLD, MMD, and MLD sets.

general, there is a positive correlation with the most dominant person and a negative correlation with the least dominant person. The highest correlation is observed for status and talkativeness.

6. Conclusion

We have described a multimodal corpus for analyzing dominance in meetings. To our knowledge, this is the first publicly available dataset that combines multiple annotations, rich sensors and multiple annotated tasks. The analysis of the annotations indicate that the annotators are quite consistent within themselves and with each other. Furthermore, the analysis results are consistent with the social psychology findings on dominance.

We believe the presented corpus provides a good testbed and benchmark for the automatic dominance analysis in small group conversations. The AMI meeting corpus is rich in terms of sensors and allows extraction of multimodal nonverbal cues for each participant in the meeting. On the other hand, the new dominance annotations and the identified datasets provide the perceived dominance of the meeting participants, as agreed by multiple annotators. This dataset allows researchers to study the links between multimodal nonverbal cues and dominance perception as well as to assess the performance of the computational models that can be used to estimate dominant and submissive behavior.

The database is available in the following address:

www.idiap.ch/scientific-research/resources/dome/

Acknowledgments: This work was supported by EU FP7 Marie Curie Intra-European Fellowship project Automatic Analysis of Group Conversations via Visual Cues in Non-Verbal Communication (NOVICOM), EU project Augmented Multiparty Interaction with Distant Access (AMIDA), and Swiss National Center for Competence in Research on Interactive Multimodal Information Management (IM2) project. We thank Dinesh B. Jayagopi (Idiap) and the annotators at Bogazici University and at Idiap for their help.

7. References

N. Ambady and R. Rosenthal. 1992. Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis. *Psychological Bulletin*, 111:256–274.

J. K. Burgoon and N. E. Dunbar. 2006. Nonverbal expressions of dominance and power in human relationships. In V. Manusov and M. Patterson, editors, *The Sage Handbook of Nonverbal Communication*. Sage.

J. Carletta, S. Ashby, S. Bourban, M. Flynn, M. Guillemot, T. Hain, J. Kadlec, V. Karaiskos, W. Kraaij, M. Kronenthal, G. Lathoud, M. Lincoln, A. Lisowska, I. McCowan, W. Post, D. Reidsma, and P. Wellner. 2005. The AMI meeting corpus: A pre-announcement. In *Workshop on Machine Learning for Multimodal Interaction (MLMI'05) Edinburgh, U.K.*, July.

J. F. Dovidio and S. L. Ellyson. 1982. Decoding visual dominance: Attributions of power based on relative percentages of looking while speaking and looking while listening. *Social Psychology Quarterly*, 45(2):106–113, June.

N. E. Dunbar and J. K. Burgoon. 2005a. Measuring nonverbal dominance. In V. Manusov, editor, *The sourcebook of nonverbal measures: Going beyond words*. Erlbaum.

N. E. Dunbar and J. K. Burgoon. 2005b. Perceptions of power and interactional dominance in interpersonal relationships. *Journal of Social and Personal Relationships*, 22(2):207–233.

D. Gatica-Perez. 2009. Automatic nonverbal analysis of social interaction in small groups: a review. *Image and Vision Computing, Special Issue on Human Behavior*, 27(12):1775–1787, December.

J. A. Hall, E. J. Coats, and L. Smith LeBeau. 2005. Nonverbal behavior and the vertical dimension of social relations: A meta-analysis. *Psychological Bulletin*, 1313(6):898–924.

H. Hung, D. B. Jayagopi, S. Ba, and and D. Gatica-Perez J.-M. Odobez. 2008. Investigating automatic dominance estimation in groups from visual attention and speaking activity. In *International Conference on Multimodal Interfaces (ICMI'08), Chania, Greece*, October.

D. B. Jayagopi, S. Ba, J.-M. Odobez, and D. Gatica-Perez. 2008. Predicting two facets of social verticality in meetings from five-minute time slices and nonverbal cues. In *International Conference on Multimodal Interfaces (ICMI'08), Chania, Greece*, October.

D. B. Jayagopi, H. Hung, C. Yeo, and D. Gatica-Perez. 2009. Modeling dominance in group conversations from nonverbal activity cues. *IEEE Transactions on Audio, Speech, and Language Processing, Special Issue on Multimodal Processing for Speech-based Interactions,*, 17(3):501–513, March.

M. L. Knapp and J. A. Hall. 2009. *Nonverbal Communication in Human Interaction*. Wadsworth Publishing, 7 edition.

A. Pentland. 2005. Socially aware computation and communication. *IEEE Computer*, 38(3):33–40.

R.J. Rienks and D.K.J Heylen. 2005. Dominance detection in meetings using easily obtainable features. In *Workshop on Machine Learning for Multimodal Interaction (MLMI'05) Edinburgh, U.K.*, July.

M. Schmid-Mast. 2002. Dominance as expressed and inferred through speaking time: A meta-analysis. *Human Communication Research*, 28(3):420–450, July.

L. Smith-Lovin and C. Brody. 1989. Interruptions in group discussions: The effects of gender and group composition. *American Sociological Review*, 54(3):424–435, June.

A. Vinciarelli, M. Pantic, and H. Bourlard. 2009. Social signal processing: Survey of an emerging domain. *Image and Vision Computing*, 27(12):1743–1759, November.