An Audio Visual Corpus for Emergent Leader Analysis

Dairazalia Sanchez-Cortes[1,2], Oya Aran[1], and Daniel Gatica-Perez[1,2]
[1] Idiap Research Institute, Martigny, Switzerland
[2] Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland
(dscortes,oaran,gatica)@idiap.ch

ABSTRACT

In this paper we discuss our experience designing and collecting a data corpus called ELEA (Emergent LEader Analysis), and describe the use of a light portable scenario to record small group meetings. The corpus was gathered with the aim of analyzing emergent leadership, as a social phenomenon that occurs in newly formed groups. For each group in the corpus, the participants performed the winter survival task. To date, the annotations of the corpus include personality, concepts related to leadership, and participants' performance in the survival task. In addition, several features have been extracted automatically from audio and video.

Categories and Subject Descriptors

H.3.1 [**Information Storage and Retrieval**]: Content Analysis and Indexing

General Terms

Human Factors

Keywords

small groups, emergent leadership

1. INTRODUCTION

Human interactions are rich, ranging from courtship to family, working in teams and building communities. Psychologists and sociologists have long studied these interactions of varying scale, to understand behavior, motivation, and emergence of interaction patterns [27, 29].

From the viewpoint of social computing research, results traditionally obtained in psychology can now be revisited with significant developments in recording, automatic analysis, and machine learning [13]. One recent advantage of this trend is sharing not only the research results, but also the data and therefore allowing the possibility of reproducing the results by other research groups.

ICMI-MLMI'11, Multimodal Corpora for Machine Learning: Taking Stock and Road mapping the Future. November 14-18, 2011, Alicante, Spain.

A number of behavioral data corpora with multimodal data, certain annotations, and experimental results are currently available. Affect and behavioral cues like facial expression, prosody, turn-taking patterns, head pose, and gestures have been studied [13]. Furthermore, manual and automatic versions of behavioral cues have been used to infer social constructs like influence, performance, and cohesion. The recording solutions have also varied, from wearable devices, i.e., fully portable sensing, to fixed infrastructure-based sensors. While wearable sensors are flexible, fixed sensors are hassle-free for the people interacting.

In this paper, we discuss our experience in creating a corpus of small group conversational interaction. Our work has two novelties. First, we introduce a portable audio and video recording solution for small group recordings. Second, to our knowledge, ours is the first corpus explicitly collected with the aim of studying the emergent leadership phenomenon in social computing. The scenario used in the recordings has been specifically designed to study the possible emergence of leaders, and it contains a variety of annotated data. The definition of emergent leader involves a person that arises from the group and has her/his base of power from followers, rather than from a higher authority [32]. This is a problem that has not been analyzed with automatic means, either as an isolated concept or in relation to other constructs like personality or performance.

We present related work in Section 2. In Section 3, we present the instruments used to collect the corpus. Section 4 explains the coding scheme. Section 5 presents a brief analysis on the data on leadership and personality, as an illustration of the type of research questions that can be addressed. Finally we present our conclusions in Section 6.

2. RELATED WORK

Several corpora have been collected to study behavior in small groups; most of them have centered their attention on meeting scenarios where realistic rich interacting patterns can emerge. The corpora vary among sensors used to record, manual annotations, and the way to promote the interaction (real scenarios vs scripted scenarios). Table 1 summarizes the available corpora focused on small group interactions, described in this section.

With the aim of understanding the structure in meetings where mission objectives are defined, roles and hierarchy are known, and the planning activity is present, the VACE meeting corpus has been recorded in real-world scenarios (wargames and military exercises) at the Air Force Institute of Technology [10].

Capturing as well natural weekly discussions from a research group where roles and hierarchy are known, the ISL corpus recorded several meetings at ICSI's conference room [15]. The goal of this corpus is to offer resources to improve automatic speech recognition, transcription, prosody and dialog modeling.

Another corpus collected real and scripted meetings, considering project planning, military exercises, games, chatting and discussion [7]. The aim of this corpus is to distinguish between different kinds of meetings taking into account speaking styles.

The AMI-12 corpus was collected at the IDIAP smart meeting room [17]. In the scenario proposed, participants have roles assigned and follow a script. Apart from audio and video resources, manual annotations that involve verbal, nonverbal and contextual features are available.

Another behavior that emerges in meetings is related to dominance and influence [17, 10, 20]. The DOME corpus is a subset of the AMI corpus, containing 10 hours of meetings recorded at IDIAP smart room, and includes dominance annotations on small group conversations [4]. To analyze participant's influence in project scenario meetings, a part of the AMI corpus was analyzed, containing 40 meetings recorded at TNO-Soesterberg [28]. Several manual annotations are available for this corpus, mostly derived from the audio channel.

Participant's involvement has also been analyzed in small business meetings. Campbell et al., gathered a corpus named ATR in which monthly sessions from a real group project meeting were recorded [8]. The main goal of this corpus is to identify the type of participation and the flow of the discourse.

Another approach in the meeting area research detects social interactions (including dominance) and promotes group collaboration (trough real time feedback), by using a wearable sociometer that gathers nonverbal signals and proximity data from short distance transmitters. The available corpus was recorded in two scenarios, brainstorming and problem solving [20]. For this corpus nonverbal features and self-reported dominance annotations are available.

In [25] a small corpus called NTT was presented with the aim of inferring the structure of the meeting and the participants roles. This corpus recorded discussion scenarios in which no roles were assigned. The data collected includes audio, video, and head directions extracted from sensors.

Close to our work, the multimodal corpora Mission Survival (corpus MSC-1 and MSC-2) have been recently collected [21, 26]. The data comprises small groups performing the winter survival task. The MSC-1 focuses on the individual behavior during the decision making process; it includes audio and video recordings of four participants and annotations of functional roles. The MSC-2 focuses on analyzing performance, group cohesion, and personality, and used the same video recording resources used in MSC-1; in addition they performed an online 3D multi-person tracking during the interaction. For audio recording they reduced the number of sensors to 4 close-talk microphones and one omnidirectional microphone placed on the top of the table. The MSC recordings differ from our corpus in terms of participants, given that participants at MSC-1 knew each other (all worked at FBK). In terms of settings, both corpora (MSC-1 and MSC-2) used a static setup, i.e. recorded all the meetings in a static location in the smart CHIL room at FBK.

Although real scenarios have been recorded and some behaviors that emerge in small group interactions have been analyzed, the emergent leadership phenomenon has not been considered. An emergent leader is defined as the one who arises from an interacting group and has her/his base of power from followers rather than from a higher authority [32]. The emergent leadership phenomenon arises from group interactions in which participants do not have roles assigned. Since this appears mostly in newly formed groups, the behavior of a participant during this short interaction makes him/her succeed (or fail) as a leader, without considering past information of competence, related task performance or friendship. On the other hand, personal traits might have an impact on leadership skills; this include Big-Five personality and dominance [19]. The richness of the proposed collected scenario in our corpus ELEA and preliminary findings could be of interest to researchers in social-psychology, and social computing, as well as organizations and assessment centers.

3. EMERGENT LEADERSHIP CORPUS

The corpus to analyze emergence of leadership consists of 40 meetings, 27 were completely recorded with a portable audio-visual setup, and 10 with a static setup, and 3 meetings recorded with only audio (the portable video recordings were not successfully recorded and thus discarded from the corpus). For the group interactions, three or four people are seated around a table, and audio and video is recorded, while the participants perform a winter survival task. Before and after the task, the participants fill some questionnaires that are used as ground truth in the analysis of emergent leadership and related concepts. The total duration of the corpus is approximately 10 hours. An initial part of the data was originally studied in [30]. Here, we describe the full corpus in more detail, and discuss our practical experience with its design and implementation.

Sensing infrastructure: To collect the audio and visual data, we used two setups, one static and one portable. The static setup includes the IDIAP smart meeting video resources [23], composed of four closeup views, two side views and a center view recording at 25 fps. For the audio we used Dev-Audio's microcone, a commercial portable microphone array designed to record group focus interactions [1]. This device directly outputs speaker segmentation for each participant (assuming that people do not change seats during the interaction).

The portable video setup uses two wide-angle webcameras (Logitech Webcam PRO 9000), with frame size 640x480 pixels, at 30 fps. The design of this portable system was chosen such that it is easy to obtain and replicate in diverse settings, and allows adequate resolution and frame rate for our analysis purposes. Although spherical camera systems (either with a single 360 degree lens or with multiple lenses) provide a larger camera view, these cameras are in a higher price range and few of them meet our resolution and frame rate criteria. Among the portable video recording systems used in social computing research, in [8], a spherical lens with a frame rate of 12 fps is used. The resolution is low and does not allow the analysis of fine details of participants' movements. In [24], two omnidirectional cameras with fish eye lenses are used. The system provides high resolution and 30 fps frame rate. In comparison to these video recording systems, our system uses commercial webcams and provide a cheap and easy-to-obtain solution for small group video

Corpus	Audio/video/room setting	Questionnaires/annotations
VACE [10]	up to 8 EWM, OTMs, 1 OD and 1 FC 10 VC	conversation transcripts, dominant speaker, language metadata (e.g. floor control change), gesture
ICSI [15]	4 to 8 CTM	involvement
ISL [7]	3 to 9 LAM 3 VCs ISL meeting room	word tokens, turns, question/non-question, disfluency
AMI-12 [17]	4 CTM, 4 LAM, 1 ARM 4 CU and 3 VC IDIAP Smart meeting room	conversation transcript, addresses, gaze direction, adjacency pairs (question-answer, statement-agreement)
AMI-40 [28]	1 ARM 4 CU and 3 VC TNO-Soesterberg meeting room	influence ranking (inter-ranking) dominance
AMI [9]	same as AMI-12 and AMI-40 same as AMI-12 and AMI-40 IDIAP, TNO-Soesterberg and Edinburgh room	same as AMI-12 and AMI-40, hand and head gestures
DOME [4]	same as AMI-12 same as AMI-12 IDIAP room	same as AMI-12, dominance annotations
M4 [22]	12 microphones (ARM and LAM) 3 VC	conversation transcript, word segmentation, interest level
NIST [12]	3 to 9 CT, LAM and OTMs 5 VC	conversation transcript, speaker segmentation
ATR [8]	1 ARM 1 C360, 1-6 VCs NAIST and ATR	none
MIT [20]	4 SBM	dominance, questions and ideas, team performance
NTT [25]	4 LAM 3 VC	regime estimates (class + directionality) head direction (from magnetic sensors 6-DOF)
MSC-1 [26]	4 CTM, 6 TTM and 7 ARM 5 VC (4 corners, 1 ceiling), 4 WC (walls) CHIL room	functional relational roles (task area and socio-emotional)
MSC-2 [21]	4 CTM, 1 ODM same as MSC-1	personality LCB and E-BFMS, group cohesion, individual and group performance

Table 1: Corpora available for small-group interaction study. The audio sensors/microphones include CTM-close-talk, EWM-earset wireless, TTM-tabletop, LAM-lapel, SBM-sociometer badge, ARM-microphone array, ODM-omnidirectional, FCM-four-channel cardiod, OTM-Other distantly placed microphones. Video sensors include CU-close-up, VC-video camera, WC-webcamera, C360-360 degree camera. The personality annotations correspond to LCB-Craig's Locus of Control of Behavior scale and E-BFMS-Extroversion part of the Big Marker Five Scales.

recordings with sufficient resolution and frame rate.

The portable setting requires two laptops, one for the microcone with an USB port and one for the video with two USB ports. Since audio and video were recorded separately, the synchronization was done manually by clapping once in the center of the table. Figure 1 and 2 shows a snapshot from the portable recording scenario and the capture devices respectively.

Figure 1: A snapshot from the ELEA corpus, portable setting. The webcam is circled in red and the Microcone is circled in blue.

Subjects: Potential volunteers were invited to participate in a study on casual social interactions, the invitations were posted in English and French offering a monetary compensation for their participation. Advertisements were placed in two universities, a research center and a business management school in French-speaking Switzerland. After participants contacted us by phone or email, they were informed

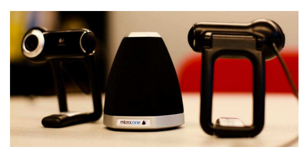

Figure 2: Capture devices, the webcam and the Microcone, used in the ELEA recordings.

of the process and, if they agree to participate, cellphone number and email were requested. Since the participants were not supposed to have previous partnership or work relationship, ad-hoc groups were formed and participants were requested to attend the recordings.

We recruited 148 participants, from which we have 48 females and 100 males in mixed teams. 28 teams are four-person and 12 teams are three-person. Average age is 25.4 years, with standard deviation 5.5.

Trust agreement: On arrival, participants signed a trust agreement. The agreement explained the process of the study, and exposes that audio and video recorded will be used only for research purposes and their identity will be anonymized. The agreement emphasizes the participants' right to quit the study at any time. Participants were provided with a copy of the signed agreement, including our

complete names and email addresses for their own records.
Survival task: There are several tasks that promote group discussion and decision-making. After reviewing the tasks most often used for training in assessments centers, we chose the winter survival task, given that it promotes interactions among the participants in the group. The participants in the task are supposed to be survivors of an airplane crash. They have 12 items that they have to rank in order of their importance, giving 1 to the item considered the most important to survive as a group, 2 to the second most important, and so on. The task is performed first individually (5 min) and then we asked them to come up with the group ranking (max 15 min). Considering that not all the participants could be familiar with the items, we provide them with slides containing a picture and the definition of the item. The slides were consulted only during the individual ranking, to avoid the occlusion of the cameras during the group discussion.
Questionnaires: Four well structured questionnaires were applied, with the aim of getting ground truth for several variables from the participants in the group. For each participant, we obtained three or four questionnaire outputs, which reflected the participants' perception. The averaged outputs are considered as the ground truth.

First we administered OCEAN [11], which is a well known measure of the Big Five personality traits: Openness to Experience, Conscientiousness, Extraversion, Agreeableness, and Neuroticism (OCEAN). We used the self-reported long version of the instrument composed of 60 items, each item has a score from 1 to 5 ('Disagree totally' to 'Total agreement').

This questionnaire was followed by the Personality Research Form (PRF) [14]. This questionnaire yields scores for personality traits relevant to the functioning of individuals in power dominance and leadership. It consists of 16 true-false items. After the personality tests, we recorded the survival task.

After the task, participants filled out a Perceived Interaction Score, that captures perceptions from participants during the interaction, in which they score every participant in the group through four items related to the following concepts: perceived leadership (PLead), perceived dominance (PDom), perceived competence (PCom) and perceived liking (PLike). The 16-item questionnaire can be scored from 1 to 5 ('Not at all' to 'Frequently if not always', respectively). Afterwards they provide a dominance ranking (RDom), i.e., participants were asked to rank the group, given 1 to the most dominant participant, and 3 or 4 for the less dominant, such that they have to include themselves in the ranking, similarly to previous work in dominance annotation [16].

Finally, participants were asked to provide additional information including age, and experience in practicing outdoor activities and winter sports in a scale from 1-5 ('Not at all'-'Frequently, if not always'). It was optional to provide additional comments to express their feelings during the interaction and about the process.

4. CODING

This section describes the coding used to process the collected data and the results of analyzing the questionnaire data.

To keep their identity anonymized, participants chose a letter K, L, M, or N and to link them with their respective questionnaires and audio/video files, the final identifier is defined as: number of group, participant letter, day and month of recording and a letter indicating the gender. Below we describe the computations done from each of the questionnaires.

OCEAN: From this questionnaire, we compute mean values over the items that correspond to each of the big five traits, taking into account that some items needed to be reversed. For each person we have a vector of five real values between 1.0-5.0. For the full population, mean values for the big five traits are: O=3.52, C=3.82, E=3.63, A=3.68, and N=2.52.

PRF: Since this questionnaire is of the form true-false, we mapped the values to 1-0, such that we accumulated the number of items corresponding to power or dominance. In the data set we have two values, one corresponding to the number of items related to leadership and dominance, and a second value that represents the mean value.

Perceived interaction scores: For this questionnaire we calculated mean values for each of the perceived variables PLead, PDom, PCom, PLike, using the judgment about other participants (i.e., not herself/himself). Since the participants filled this questionnaire about themselves, we have mean values for self-reported perception of the variables as: SLead, SDom, SCom and SLike. We consider as ground truth the annotations from the perceived interactions, such that the emergent leader in the group is the participant with the highest mean value of perceived leadership, and similarly for the related concepts.

Ranking Dominance: We calculated the ranking dominance value per participant as the mean value of the rank assigned from the other participants. Additionally we have the self dominance ranking value $SRDom_i$.

Survival task performance: Although there is no unique solution for the winter survival task, there is a ranking provided by experts, that justify the item rank order with more chances to survive. We used the survival experts' ranking list to code some variables, and we proposed another coding as follows:

AIS: The Absolute Individual Scores are calculated based on the absolute difference between the individual ranking list and the survival experts ranking list. The smaller the score, the better the answer.

AGS: The Absolute Group Score is calculated based on the absolute difference between the group ranking list and the survival experts ranking list. The smaller the score, the better the answer.

AII: The Absolute Individual Influence is calculated accumulating the absolute difference between the individual and the group ranking list. The smaller the score, the higher the influence of the individual in the group.

NTII: Number of Top Individual Items in the top group items, which counts the number of items in the individual list that also appear in the group list top items.

DTII : Absolute Distance of the Top items in the top individual list with respect to the position of the same item in the top group list. If one item is not in the top rank, it is assigned with the maximum distance + 1.

Automatic nonverbal features: In addition to manual coding, our corpus includes a number of automatically extracted features. Given that the emphasis of the paper is on the corpus and not on automatic recognition, we discuss these features only briefly. The speaking turn features available (speaking time, interruptions, turns, etc.) are described in [30]. With respect to prosody we extracted energy

	O	C	E	A	N
PLead	0.12	0.03	0.07	-0.09	-0.13
PDom	0.08	-0.03	0.03	-0.20	-0.08
RDom	0.28^*	-0.04	0.13	-0.11	-0.10

Table 2: Correlation values between perceived variables and self reported personality. Significance values $+: p < 0.05$, $*: p < 0.02$.

and pitch, using the opensource software Wavesurfer [2]. Regarding visual features, we have computed measures of body and head activity (total visual activity, total activity turns, average visual activity, etc) extracted as described in [3]. Additionally we extracted measures of head and body motion, through variations of well known works in vision, more specifically motion energy image (MEI) and motion history image (MHI) as proposed by [6]. More details can be found in [31].

5. AN EXAMPLE OF RESEARCH QUESTIONS STUDIED IN ELEA

In the previous section we discussed the main aspects of the corpus. An in-depth analysis of the use of automatically annotated nonverbal features to predict emergent leadership has been reported in [31]. In this section, we present an initial analysis of the relation between leadership and personality traits as gathered from the various instruments.

5.1 Leadership and Big Five correlation

One question that our corpus allows to address is whether emergent leadership is associated with specific personality traits, as suggested in [19]. To investigate this, we performed a correlation analysis between the perceived interaction variables and the self-reported personality traits (N = 148). In this analysis we do not consider the competence and liking variables. From Table 2 we can observe that the only significant correlation is between RDom and openness to experience (0.28^*). This indicates that high dominance is associated to more open people.

Our findings show similarities with Kickul and Neuman [19] regarding self reported personality tests. Individuals that posses dominance tend to emerge as leaders, and emergent leaders score higher in openness to experience than followers. This trait is characteristic of individuals who are curious, broad-minded, creative, and imaginative. On the other hand, our correlations with extroversion are not statistically significant as compared with [19, 18].

We also explored the correlations among two classes, i.e. the personality traits corresponding to the Emergent-Leaders and the Non-Emergent leaders. Although the data is not balanced, it can provide an idea of the relationship between the classes. The correlations for the non-emergent leadership class (N=108) showed more significant correlations than the entire corpus. For PLead the correlations are: extroversion (0.20^+) and agreeableness (-0.23^*), for PDom there are significant correlations as well with: extroversion (0.21^+) and agreeableness (-0.32^*). On the other hand correlations among the class emergent leaders, became less significant due to the number of examples (N=40). Nevertheless there is significant correlation between PLead and agreeableness (-0.38^*), and PDom and neuroticism (-0.32^+). According to [18] the correlation with agreeableness is ambiguous, in one hand leaders should be more agreeable, but agreeable individuals are more likely to be modest, and leaders tend not to be excessively modest [5].

5.2 Multiple regression on personality

Perceived variables and OCEAN data can be represented as a model of the form:

$$y_{calc} = a_0 + a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4 + a_5x_5 \quad (1)$$

where x_i corresponds to each of the big five traits $x_1 = O, x_2 = C, x_3 = E, x_4 = A$ and $x_5 = N$. The multiple regression solves for unknown coefficients a_0, a_1, a_2, a_3, a_4 and a_5 by minimizing the sum of the squares of the deviations of the data from the model (least-squares fit).

$PLead = 3.63 + 0.16x_1 - 0.02x2 + 0.14x_3 - 0.25x_4 - 0.18x_5$
$PDom = 3.88 + 0.16x_1 - 0.21x2 + 0.27x_3 - 0.33x_4 - 0.19x_5$
$RDom = 0.67 + 0.05x_1 - 0.02x2 + 0.02x_3 - 0.01x_4 - 0.03x_5$
with max error: 1.56, 1.69, 1.08, 2.10, 0.30.

The max error is calculated as:

$$Error_{max} = \arg\max |y_{calc} - y|, \quad (2)$$

where y is our ground truth.

To have a better understanding of how well the predictions can be done, we calculated the coefficient of determination, r^2, which is defined as:

$$r^2 = 1 - \frac{\sum(y - y_{calc})^2}{\sum(y - \overline{y})} \quad (3)$$

For the above model, we obtained $\overrightarrow{r^2} = 0.107$, 0.139, 0.066, 0.062, 0.088. These results show that for PLead the most influential personality traits are openness to experience (0.16), followed by extroversion (0.14). On the other hand, we can observe that for PLead the maximum distance in estimation is 1.556.

Considering the three models: PLead, PDom, and RDom we can observe the same sign trend in the coefficients ($a_0 to a_5$). As we can observe, the coefficients are noticeable close for PLead and PDom with respect to Openness to experience and Neuroticism. Finally, for PDom, the trait of Extroversion it is more important than Openness to experience.

6. CONCLUSIONS

We described our experience in creating a new data corpus, named ELEA, collected with the aim of analyzing emergent leadership in small groups. The novelty of our corpus is a portable recording solution, and a detailed set of questionnaires related to perceived leadership, personality, and performance collected from the participants in each group.

The annotations available for every group includes the big five personality scores, scores on dominance and leadership, and scores from perceived and self reported leadership, dominance, competence and likeness. The corpus also includes individual and group outcomes from the performed survival task, coded as individual performance, group performance, individual influence, and number and distance of top individual items in top group item list. Finally, the corpus also includes several automatically extracted features such as binary speaking segmentation, speaking turn features (speaking time, interruptions, turn duration, speaking turn matrix) and prosodic features (pitch and energy) from audio; body and head activity, and body and head motion from video.

As an illustration of research questions that can be addressed with this corpus, we presented a brief analysis of the relation between leadership, dominance and personality variables.

We plan to work on three directions. First, we plan to extract more automatic state-of-the-art features in small group interactions, including floor patterns, gaze, visual focus of attention and emotional states. Second, for manual annotations, since we only have inter-rate annotations, we are considering to collect annotations from external observers too. Finally, we plan to investigate the use of nonverbal features to detect leadership patterns and personality traits as a joint process.

Acknowledgments: This research was supported by CONACYT (Mexico) through a doctoral scholarship, and partially supported by the projects NOVICOM (EU-FP7-IEF) and SONVB (Swiss NSF). We would like to thank Marianne Schmid Mast (University of Neuchatel) for her valuable help in the selection of the applied tests and her feedback on the design of the recorded scenario. Also, thanks to Iain McCowan (dev-audio) for technical support; Denise Frauendorfer and Pilar Lorente (University of Neuchatel), Radu-Andrei Negoescu (Idiap) for valuable help during the collection and data processing, and all the participants in the recordings.

7. REFERENCES

[1] `http://www.dev-audio.com/`.

[2] `http://www.speech.kth.se/wavesurfer/`.

[3] O. Aran and D. Gatica-Perez. Fusing audio visual nonverbal cues to detect dominance in small group conversations. In *ICPR*, 2010.

[4] Oya Aran, Hayley Hung, and Daniel Gatica-Perez. A multimodal corpus for studying dominance in small group conversations. In *Workshop LREC*, 2010.

[5] B. M. Bass. *Bass and Stogdill's handbook of leadership. Theory, research, and managerial applications.* Free Press, 1990.

[6] Aaron Bobick and James Davis. The recognition of human movement using temporal templates. *PAMI*, 23:257–267, 2001.

[7] S. Burger, V. MacLaren, and H. Yu. The isl meeting corpus: the impact of meeting type on speech style. In *ICSLP*, 2002.

[8] N. Campbell, T. Sadanobu, M. Imura, N. Iwahashi, S. Noriko, and D. Douxchamps. A multimedia database of meeting and informal interactions for tracking participant involvement and discourse flow. In *LREC*, 2006.

[9] J. Carletta, S. Ashby, S. Bourban, M. Flynn, M. Guillemot, T. Hain, V. Karaiskos J. Kadlec, W. Kraaij, M. Kronenthal, G. Lathoud, M. Lincoln, A. Lisowska, I. McCowan, W. Post, D. Reidsma, and P. Wellner. The ami meeting corpus: A pre-announcement. In *Workshop MLMI*, 2005.

[10] L. Chen, T. R. Rose, F. Parrill, X. Han, J. Tu, Z. Huang, M. Harper, F. Quek, D. McNeill, R. Tuttle, and T. Huang. Vace multimodal meeting corpus. In *Workshop MLMI*, 2005.

[11] P. Costa and R. McCrea. *NEO PI-R profesional manual.* 1992.

[12] I. Garofolo, M. Michel, C. Laprun, V. Stanford, and E. Tabassi. The nist meeting room pilot. In *LREC*, 2004.

[13] D. Gatica-Perez. Automatic nonverbal analysis of social interaction in small groups: a review. *Image and Vision Computing*, 1(12), Dec 2009.

[14] D. N. Jackson. *Personality research form manual.* Research Psychologists Press, 1967.

[15] A. Janin, D. Baron, J. Edwards, D. Ellis, D. Gelbart, N. Morgan, B. Peskin, T. Pfau, E. Shriberg, A. Stolcke, and C. Wooters. The icsi meeting corpus. In *ICASSP*, 2003.

[16] D. Jayagopi, H. Hung, C. Yeo, and D. Gatica-Perez. Modeling dominance in group conversations using nonverbal activity cues. *TASLP*, 17(3), 2009.

[17] N. Jovanovic, R. op den Akke, and A. Nijholt. A corpus for studying addressing behavior in multi-party dialogues. In *Workshop DD*, 2005.

[18] Timothy A. Judge, Joyce E. Bono, Remus Ilies, and Megan W. Gerhardt. Personality and leadership: A qualitative and quantitative review. *Journal of Applied Psychology*, 87(4):765–780, 2002.

[19] J. Kickul and G. Neuman. Emergent leadership behaviours: The function of personality and cognitive ability in determining teamwork performance and ksas. *Journal of Business and Psychology*, 15(1), 2000.

[20] T. Kim, A. Chang, L. Holland, and A. Pentland. Meeting mediator: enhancing group collaboration with sociometric feedback. In *CSCW*, 2008.

[21] Nadia Mana, Bruno Lepri, Paul Chippendale, Alessandro Cappelletti, Fabio Pianesi, Piergiorgio Svaizer, and Massimo Zancanaro. Multimodal corpus of multi-party meetings for automatic social behavior analysis and personality traits detection. In *Workshop TMR*, 2007.

[22] I. McCowan, D. Gatica-Perez, S. Bengio, G. Lathoud, M. Barnard, and D. Zhang. Automatic analysis of multimodal group actions in meetings. *PAMI*, 27(3):305–317, 2005.

[23] Darren C. Moore. The idiap smart meeting room. Technical report, Idiap Research Institute, 2002.

[24] K. Otsuka, S. Araki, K. Ishizuka, M. Fujimoto, M. Heinrich, and J. Yamato. A realtime multimodal system for analyzing group meetings by combining face pose tracking and speaker diarization. In *ICMI*, 2008.

[25] K. Otsuka, Y. Takemae, J. Yamato, and H. Murase. Probabilistic inference of multiparty-conversation structure based on markov-switching models of gaze patterns and head directions and and utterances. In *ICMI*, 2005.

[26] Fabio Pianesi, Massimo Zancanaro, Bruno Lepri, and Alessandro Cappelletti. A multimodal annotated corpus of consensus decision making meetings. *Language Resources and Evaluation*, 41:409–429, 2007.

[27] M. S. Poole, Andrea B. Holligshead, Joseph E. McGrath, Richard L. Moreland, and John Rohrbaugh. Interdisciplinary perspectives on small groups. *Small Group Research*, 35(1):3–16, 2004.

[28] Rutger Rienks, Dong Zhang, Daniel Gatica-Perez, and Wifried Post. Detection and application of influence rankings in small group meetings. In *ICMI*, 2006.

[29] E. Salas, D. E. Sims, and C. S. Burke. Is there a big five in teamwork. *Small group research*, 36(5):555–599, 2005.

[30] D. Sanchez-Cortes, O. Aran, M. Schmid Mast, and D. Gatica-Perez. Identifying emergent leadership in small groups using nonverbal communicative cues. In *ICMI*, 2010.

[31] D. Sanchez-Cortes, O. Aran, M. Schmid Mast, and D. Gatica-Perez. Detecting emergent leaders in small groups using nonverbal behavior. Technical report, Idiap Research Institute, 2011.

[32] R. T. Stein. Identifying emergent leaders from verbal and nonverbal communications. *Personality and Social Psychology*, 32(1):125–135, 1975.