Towards Multimodal Analysis of Dialogic Moments in Storytelling-based Discourse

Ngoc N. T. Doan[1], Andrius Penkauskas[1], Ecaterina Grigoriev[1]
Lisa E. Rombout[2], Paris Mavromoustakos-Blom[2], Martin Atzmueller[2], Maryam Alimardani[2]
Tilburg University
Contact: n.t.n.doan@uvt.nl or m.alimardani@uvt.nl
[1]Students, [2]Project supervisors

ABSTRACT

The key to an effective communication method arguably lies in its ability to facilitate moments of high mutual understanding (dialogic moments). It would, therefore, be useful to identify these moments and perhaps facilitate them. In this pilot study, we present a multimodal analysis of dialogic moments in storytelling-based discussions. We collected skin conductance, heartrate, speaking turns, relative body position, conversation transcripts, and subjective experience. This multimodal data corpus enables the computational study of these highly subjective moments and the potential creation of digital communication aids. Preliminary results show that there might be subcategories of dialogic moments that were previously unidentified.

Keywords

Storytelling, multimodal interaction, conversation analysis, dialogic moment

INTRODUCTION AND RELATED WORKS

Storytelling has long been used to improve teaching quality (1,2) and to strengthen partnerships within and between organizations (3,4). A storytelling conversation can elicit moments of high mutual understanding, called dialogic moments, more efficiently than an issue-oriented one (5-7). Dialogic moments occur when each participant strives to acknowledge and respond to others' experienced truth, while remaining faithful to their own. This mutual understanding fosters a more impartial consideration of different opinions. Generally, dialogic moments are defined by the real 'Self' and 'Other' being in direct contact, enabling the inclusion of 'Other' in the 'Self'. This has 2 specific steps:

- The participant places themselves at a distance from the 'crowd', and becomes their independent, autonomous self, achieving personal unity. They feel safe enough to permit conflicts to rise, manifested by responding and acting in an authentic way (5-9).
- Then, they enter in a relationship with others. They explore these others' selves and the ways in which they are tied to them, and develop a sense of belonging to the group.

Even though dialogic moments do not guarantee unanimity, they lay its foundation and prompt an overall meaningful conversation (5-9). Promoting the occurrence of dialogic moments can therefore be useful for mediating real-life collaborations and conflict situations.

SRC 2019, December 5, 2019, The Netherlands.

Storytelling is generally an effective way of triggering dialogic moments: it can revive a storyteller's experiences in detail. This allows listeners to take something from the experience and treat it, on some level, as their own (8,7). However, storytelling and dialogic moments have not yet been extensively studied, especially from a computational perspective. Traditional dialogue researchers have often inspected these moments by ruminating on their own dialogic encounters (9), on the lives of highly dialogical individuals (11), or on the content of such moments (4,7). Their research results have shown the importance of dialogic moments but give us few indications on how to predict or replicate a dialogical moment (7,12). For this, a multimodal approach might be the answer.

Multimodal approaches to the study of social interaction are becoming more common. Data is obtained from participants using wearable sensors, transcripts and recordings, and analyzed with sophisticated computational modeling techniques (13, 14). Most prominently in this research direction, Okada, Hang and Nitta (15) studied the correlation between an evaluation of the storytelling by external observers, and the multimodal features of the participants. Yet, they seemed to be interested more in the effects on an individual level. What makes a conversation effective at group level remains largely unknown.

We theorized that dialogic moments should leave certain traces in a participants' multimodal data. For this pilot study, we therefore recorded a wide variety of data while participants were engaged in a storytelling-based discussion - data such as speaking turns, relative body position (indicating face-to-face interaction), skin conductance, and heartrate. The moderator, a storytelling expert, then indicated the critical moments where participants seemed to come closer together, and potentially all share a mutual understanding. Thus, we create a multimodal data corpus that could be used for the modeling and predicting of dialogic moments in social interactions. Secondly, we evaluated the different types of sensors used in this pilot for their effectiveness in the circumstances of a discussion. This paper will elaborate on this pilot's methodology and its initial results.

METHODOLOGY

Participants

The participants (n=4, 2 female) were all university students. They received no incentives for their participation and gave informed consent. At the time of the experiment, none of them had any experience in professional acting or storytelling. Also, none reported being in close relationship with another. The participants sat around a square table in an empty room and remained sedentary during the experiment.

Storytelling

The experiment included 4 discussion sessions of 10 minutes each, with breaks of 2 minutes. All discussions were moderated by a professional storytelling expert. Each session was moderated to have either a positive or a negative valence. The order was Negative – Positive – Negative - Positive.

The storytelling mimicked an existing performance by the art collective SPACE. In the beginning, the moderator presented the following fictional scenario to the group: the Earth had become inhabitable, and the group was chosen as the first to seek asylum on Mars. Valence was moderated by focusing on dilemmas (negative) or opportunities (positive). Examples are the existential choice between our safety and our planet, the possible dangers in the environment on Mars, the possibility of creating a more just society, etc. The moderator could also choose to direct attention to a less active member of the group, or ask a direct question. Such basic acts of promoting respect, responsibilities and equal speaking chances are argued to harbor the conception of dialogic interaction (16).

Afterwards, the strongest dialogic moments were heuristically chosen by the moderator. Additionally, one external observer checked these against the 2 requirements of a dialogic moment mentioned in the introduction.

Data collection

Objective measurements

During the experiment, participants' heartrate (HR) and galvanic skin response (skin conductance, GSR) were recorded using Shimmer GSR+ wearable sensors, which are widely used in previous works (17,18). These data, especially GSR, indicate arousal (19,20). Arousal can be calculated by subtracting the highest GSR in the period of interest and the average GSR in the 30 seconds prior to it.

In addition, speaking volume and physical proximity were recorded using the Rhythm Badge (20). To determine the proximity to other badges, each badge, worn around the neck, scans for nearby Bluetooth devices every 60 seconds and records their Received Strength Signal Indicator (RSSI) values (21). Based on these values, face-to-face interaction was extracted (22,23). A face-to-face contact between two participants was logged if the RSSI signals between their badges was greater than their baseline. The baseline was the mean RSSI between a pair of badges when participants were in a neutral sitting posture. The termination of a contact was empirically determined if the signal did not exceed the baseline for 40 seconds.

Lastly, we transcribed the content of participants' speech from video recordings, and manually recorded their speaking turns. The recording of speaking turns assumes one speaker at a time. In a period of overlapping speech, a participant who spoke the longest was recorded as taking the turn.

Subjective measurement

An Inclusion of Other in the Self (ISO) questionnaire (24) was distributed to each participant to fill out during each 2-minute break. In this break, they were asked to momentarily exit the fictional world and individually report the perceived level of inclusion in the group in the preceding session.

RESULTS

Due the nature of this study and the small participant group, results of the pilot were not statistically analyzed. Any reflections on the data brought forwards in the following section should be regarded as exploratory only. Additionally, heartrate and speaking volume data were too noisy and incomplete to be analyzed. We will reflect further on measurement choices in the discussion.

In total, 6 discussion moments were selected as dialogic moments. Content-wise, moment 4 was when participants engaged in a heated dispute about the ethical way to choose which people on Earth were to be evacuated first. Conversely, the other 5 moments evolved around rather non-provocative sharing such as during Moment 1:

> **P4**: My favorite place on Earth is… uh… the mountains near where I lived. [...] Eh, yeah, it's really beautiful. Uhm. It's like mostly forested and you can look across. There's this really… really giant lake that just stretches for like… like a hundred kilometers… And across there are more mountains and the sun sets over the mountains. And I love to just like… you can watch every evening as the sun like sets over the lake over the mountains. And they all like blast with color, and then it gets dark.

Or during Moment 6:

> **Mod:** So, who can we afford to lose? […]
>
> **P2:** Can't we all step out? **[**…]
>
> **Mod:** Like go together, die together?
>
> **P2:** Well, I don't know because like it it's been nine months right, so like... if theoretically we could... like have gotten close and, I don't know, it's like.... for me personally it would be really hard to just point out someone and just go and die. Like... come on, this is hard [...] Imagine that they die… so… like I would feel super guilty.

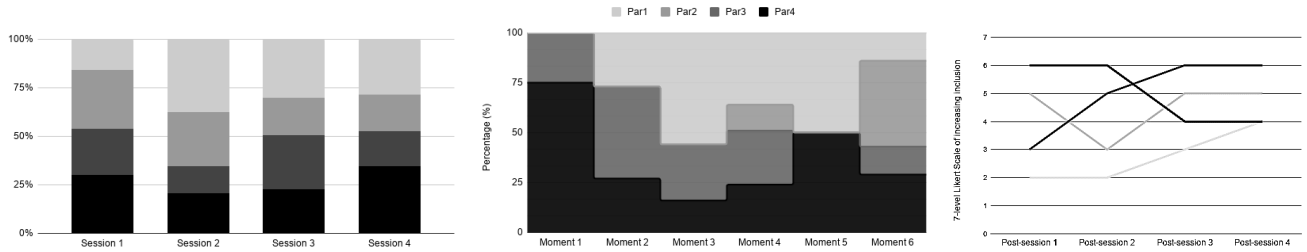
During the 6 moments, we observed no significant changes in face-to-face contacts, nor balance of speaking contribution (**Figure 1, 2**). The only moment with rather even share of speaking turns is Moment 4 in session 3. Moment 4 also was the moment almost started by everyone, while the other 5 were started by a single member (a chief storyteller), who also took the most speaking turns (**Figure 2**).

Additionally, Moment 4 witnessed a great fluctuation in the GSR signals of each participant and an overall varied trend among participants, while the opposite was seen in the other 5 (**Figure 4)**. In for example Moment 1 and Moment 6, the chief storyteller, by expressing their highly personal perspectives (or experiences), seemed to galvanize the others into a less intense state: their GSR all appear to similarly diminish (**Figure 4: moment 1 and 6**). Note that changes in GSR only signify the level of emotional intensity, and do not imply any specific type of emotion.

Note also that Moment 4 was the only one happened during a session were negative valence was promoted by the moderator (session 3), whereas the other 5 happened all in session 2 and 4 of positive valence. Interestingly, Moment 4, as session 3's only critical moment and its ending moment, could possibly be the reason for a positive change in participants' ISO scores. Specifically, 3 out of 4 participants reported a sudden increase in ISO **(Figure 3).** Whereas right after session 2, which had positive valence and 3 dialogic moments, 3 out 4 participants reported the same or decreasing ISO score.

DISCUSSION

The preliminary descriptive analysis of the data has shown great potential in studying dialogic moments computationally. Although all 6 moments in the pilot arguably fell into the category of dialogic moments, we observed little agreements between their resulting multimodal data, especially in the case of Moment 4. This implies that there could exist subcategories or a spectrum of dialogic moments that previous literature is unaware of. This may be revealed in larger datasets by further analysis of the



From left to right. ***Figure 1.*** Share (%) of speaking turns by each participant (PAR) per session. ***Figure 2.*** Share (%) of speaking turns by each participant per moment. Moment 1 to 3 belong to session 2; Moment 4 belong to session 3; Moment 5 to 6 belong to session 4. ***Figure 3.*** Participants' self-report scores of "Inclusion of Others in the Self" (ISO) on the scale of 7 after every session.

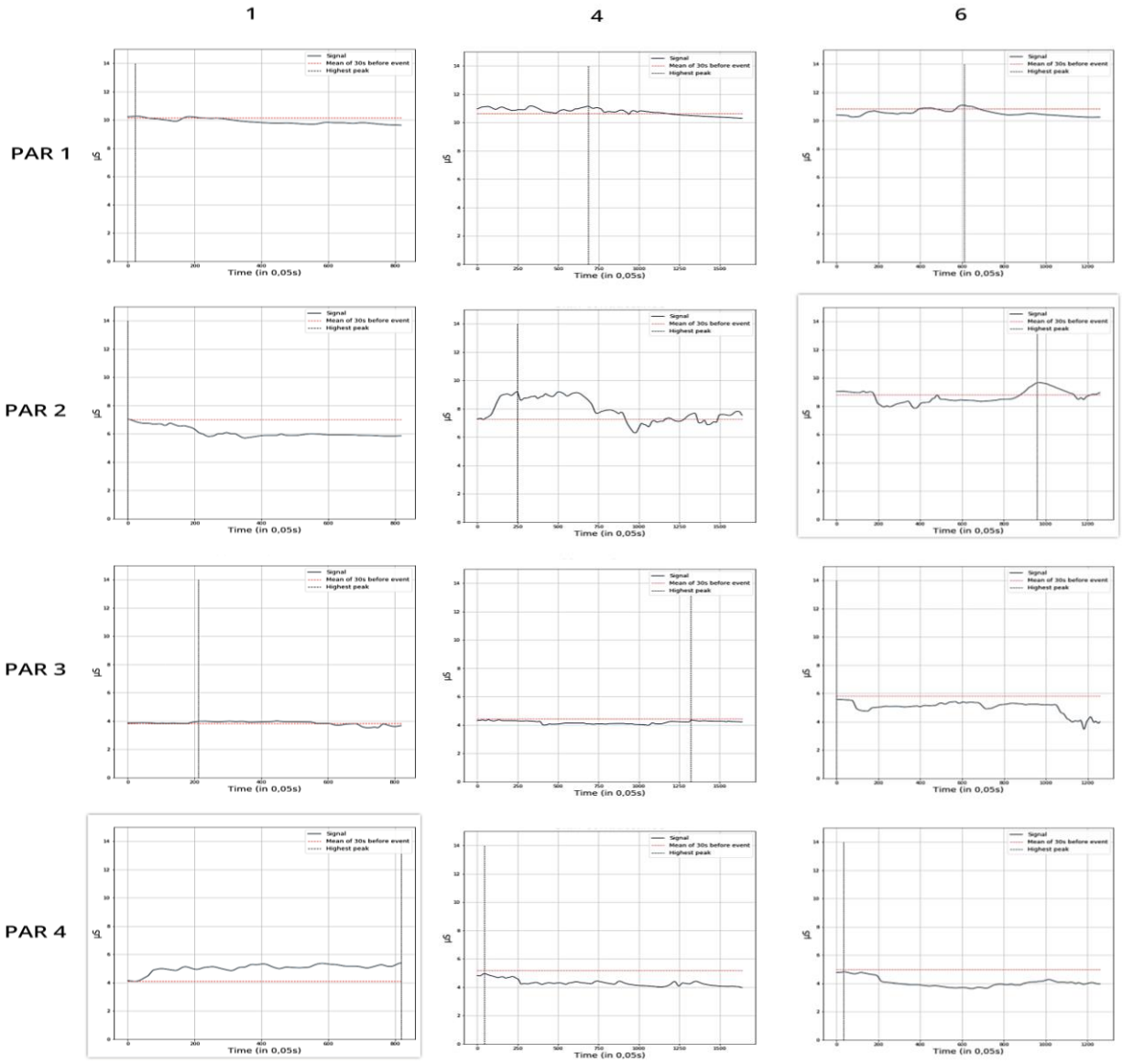


Figure 4. Galvanic Skin Response (GSR) signals of each participant (PAR), from left to right, in Moment 1, Moment 4, Moment 6.

The moments in Figure 1-4 are Moment 1: Par4 shared memory of Lake Vermont, Moment 2: Par3 argued for "kindness" as a moral norm on Mars, Moment 3: Par1 shared the desire for a Plato's utopia-like society on Mars, Moment 4: All negotiated for a moral evacuating scheme for the people "who are still on Earth", Moment 5: Par1 volunteered to step out of the spaceship first regardless of the potential dangers, and Moment 6: Par2 shared affection toward others.

variables entailed in a moment: its position in the timeline, the valence of content, participants' turn-taking behavior, etc. Very roughly, it seems like there might be a difference between dialogic moments arising from resolved conflict, and those arising from more benign sharing of experiences.

Within each moment, similarities were often observed in the multimodal data of not all, but some participants. This implies that though from an external observer's standpoint a dialogic moment includes every participant, the level of connectedness among participants may vary. These are fine-grained details that can be further analyzed in the future when there is more data from similar-structured experiments and counterbalanced stimuli (i.e. an alternative order of Negative-Positive valence).

Heartrate and speaking volume data were too noisy and incomplete to be analyzed. This shows that the

corresponding sensors are very susceptible to external factors. Although more sophisticated firmware for the sensors and data processing techniques are being developed, using wearables in a dynamic environment comes with specific challenges. Movement being the main culprit for data-loss, we recommend choosing sensor-locations away from extremities such as hands or feet.

A more robust method of selecting dialogic moments should be developed for future research, for instance by using several observers. Additionally, the selected moments should be compared to random moments to avoid confirmation bias. Alternatively, the moderator could be supplied with a continuous measuring device, were the intensity of connection can be indicated on a sliding scale and throughout the experiment.

Future research in the direction initialized by this pilot can help to create digital storytelling technology that has the ability to identify dialogic moments in conversation, quantifying an otherwise very subjective phenomenon. Potentially, understanding dialogic moments can even help us induce them in otherwise unproductive conversations, assisting with difficult discussions and negotiations.

ROLE OF THE STUDENTS

Ngoc N.T. Doan, Andrius Penkauskas, and Ecaterina Grigoriev are undergraduate students. The research direction was proposed by the 4 other authors. Research design was a collaborative effort. Data collection and data processing was performed by the students. This paper was primarily written by Ngoc N.T. Doan, with help from the supervisors.

ACKNOWLEDGMENTS

We want to thank the artists of SPACE, Petra Ardai and Rinske Bosma, for their help in setting up and executing this study, and transcribing the discussions.

REFERENCES

1. Davidhizar R, Lonser G. Storytelling as a teaching technique. Nurse Educator. 2003 Sep 1;28(5):217-21.
2. Robin B. The power of digital storytelling to support teaching and learning. Digital Education Review. 2016 Dec 15(30):17-29.
3. Barker RT, Gower K. Strategic application of storytelling in organizations: Toward effective communication in a diverse world. Int. J. Bus. Commun. (1973). 2010 Jul;47(3):295-312.
4. Orr K, Bennett M. Relational leadership, storytelling, and narratives: Practices of local government chief executives. Public Adm. Rev.. 2017 Jul;77(4):515-27.
5. Cissna KN, Anderson R. Theorizing about dialogic moments: The Buber-Rogers position and postmodern themes. Commun. Theory. 1998 Feb;8(1):63-104
6. Cissna KN, Anderson R. Moments of meeting: Buber, Rogers, and the potential for public dialogue. Albany: State University of New York Press. 2002.
7. Black LW. Deliberation, storytelling, and dialogic moments. Commun. Theory. 2008 Jan 15;18(1):93-116.
8. Bauman R. Story, performance, and event: Contextual studies of oral narrative. Cambridge University Press; 1986 Sep 26.
9. Czubaroff J. Dialogical rhetoric: An application of Martin Buber's philosophy of dialogue.2000:168-89.
10. Hammond SC, Anderson R, Cissna KN. The problematics of dialogue and power. Annals of the International Communication Association. 2003 Jan 1;27(1):125-57.
11. Arnett RC. Toward a phenomenological dialogue. West. J. Commun. 1981 Dec 1;45(3):201-12.
12. Johannesen RL. The emerging concept of communication as dialogue. 1971:373-82.
13. Gaskin J, Jenkins J, Meservy T, Steffen J, Payne K. Using wearable devices for non-invasive, inexpensive physiological data collection. In Proc. of the 50th Hawaii International Conference on System Sciences 2017 Jan 4.
14. Rombout LE, Atzmüller M, Postma-Nilsenová M. Towards Estimating Collective Motor Behavior: Aware of Self vs. Aware of the Other. AfCAI 2018.
15. Okada S, Hang M, Nitta K. Predicting performance of collaborative storytelling using multimodal analysis. IEICE Trans. Inf.& Syst. . 2016 Jun 1;99(6):1462-73.
16. Gastil J. Democracy in small groups: Participation, decision making, and communication. John Gastil; 1993.
17. Burns A, Greene BR, McGrath MJ, O'Shea TJ, Kuris B, Ayer SM, Stroiescu F, Cionca V. SHIMMER™–A wireless sensor platform for noninvasive biomedical research. IEEE Sens. J. 2010 Jun 7;10(9):1527-34.
18. Mavromoustakos-Blom P, Bakkes S, Spronck P. Personalized crisis management training on a tablet. In Proceedings of the 13th Int. Conf. on the Foundations of Digital Games 2018 Aug 7 (p. 33). ACM.
19. Critchley HD. Electrodermal responses: what happens in the brain. The Neuroscientist. 2002 Apr;8(2):132-42.
20. Boucsein W. Electrodermal activity. Springer Science & Business Media; 2012 Feb 2.
21. Lederman O, Mohan A, Calacci D, Pentland AS. Rhythm: A unified measurement platform for human organizations. IEEE MultiMedia. 2018 Jan 16;25(1):26-38.
22. Isella L, Romano M, Barrat A, Cattuto C, Colizza V, Van den Broeck W, et al. Close encounters in a pediatric ward: measuring face-to-face proximity and mixing patterns with wearable sensors. PloS one. 2011 Feb 28;6(2):e17144.
23. Scholz C, Atzmueller M, Barrat A, Cattuto C, Stumme G. New insights and methods for predicting face-to-face contacts. In 7th Int. AAAI conf. on weblogs and social media 2013 Jun 28.
24. Aron A, Aron EN, Tudor M, Nelson G. Close relationships as including other in the self. J. Pers. Soc. Psychol. . 1991 Feb;60(2):241.