Audio Analysis of Group Discussion Patterns in Noisy Classrooms Before, During, and After Teacher-Group Interactions

Chris Palaguachi, Cynthia D'Angelo, Elizabeth Dyer, Nessrine Machaka
cwp5@illinois.edu, cdangelo@illinois.edu, edyer@illinois.edu, machaka2@illinois.edu
University of Illinois Urbana-Champaign

Abstract: Teachers use a variety of teacher-group interactions to support, sustain, and increase small group work and group discussion. These strategies include conferring interactions like nudging and eliciting/probing that help students examine their previous or current conceptual understanding. Conversations found in these teacher-group interactions are complex and difficult to code for and interpret. In this study, researchers use qualitative methods to code teacher-group interactions, and computational methods to interpret the relationship between teacher-group interactions and group discussion before and after teacher-group interaction. The results of this study showed many of the teacher-group interactions increased in group discussion after teacher-group interaction occurred. When coding teacher-group interactions by conferring interactions, nudging had a higher percentage of interactions with increased group discussion after teacher-group interaction occurred compared to the other non-conferring interactions. In exploring teacher-group discussions, we observed how probing/eliciting and nudging change the types of conversations student have during and after teacher-group interactions.

Introduction and background

Today, teachers are expected to be superhuman. In many schools, teachers are often expected to multitask and play many roles in the classroom. They observe, assess, evaluate, and facilitate over 27 students at any given time. Additionally, teachers are expected to be able to listen to every conversation in the classroom and to provide meaningful feedback to support students' conceptual understanding of domain material. However, this can be quite challenging, especially when it is not uncommon for students in small groups to have loud and noisy conversations with one another (Sherin & Star, 2011). Furthermore, the urgency to jump from one group to another to provide equitable feedback to all students while confronting these noisy challenges can make it difficult for teachers to evaluate the effectiveness of their own interventions. In this study, we investigate the feasibility of using students' non-lexical speech data in small groups to explore the relationship between group discussion patterns and teacher interventions. In this study, we aim to answer the following research questions:

- RQ1 – In what ways can non-lexical speech analytics like voice activity detection be used to measure the effectiveness of small group discussion in naturalistic classroom settings?
- RQ2 – In what ways do different types of teacher-group interactions influence small group discussion before, during and after teacher-group interaction?

Theoretical framework and prior research

Educators are called on to craft interventions that help to support students' domain thinking and sensemaking through discursive dialogue. These types of interventions allow for educators to learn more about students' domain understanding and allows them to respond to their students' thinking. However, not all interventions are equally productive for promoting students' learning, building on students' ideas, or fostering students' agency through discussion (Munson, 2019). As previously mentioned, teachers are expected to respond to students by jumping from group to group. Although these responses can be seen as the teacher being responsive, a growing body of researchers are more interested in understanding teachers' responsiveness by further examining the quality of discussion they have with their students (Munson, 2019). Therefore, it is important for educators and researchers to have reliable tools like speech-based analytics that can help provide insights on the quality of teacher-group interactions on students' sense-making and group discussions.

Speech-based analytics, specifically lexical and non-lexical processing like Automatic Speech Recognition (ASR) and Voice Activity Detection (VAD) help in studying group dynamics and teacher-group interactions. Despite these affordances, challenges persist in using these tools in naturalistic settings. Currently, challenges include insufficient modeling of student speech and transcribing speech in noisy, multi-speaker environments. Between the two tools, the VAD within openSMILE offers a more adaptable solution. Aligned with the 12 core commitments of the new era of MMLA that focuses on intentional and privacy-conscious data

processing (Worsley et al., 2021), using non-lexical speech analytics like VAD can aid in evaluating small group work, collaborative problem solving, and the quality of teacher-group interventions.

Methods

The study presented in this paper uses data from a larger observational dataset of secondary mathematics lessons in the United States (Dyer, 2016). Middle school and high school students were asked to work in small groups on math problems ranging from algebra to trigonometry. Each classroom had up to 27 students. Students were split into 5-7 groups depending on classroom size. Students were then audio-video recorded. Researchers used a pair of Sony HDR-MV1 or Zoom Q4 cam-recorders positioned at the front corners of the classroom capturing footage at 1080p to collect video. Zoom H1 microphones were stationed at the center of each group's table and used to collect group audio. In total, 106 classroom sessions were recorded from 10 instructors.

For this study, we selected one 100-minute lesson from the primary dataset. This classroom session had 27 students split into 7 small groups, and groups consisted of 3 or more students. Qualitative coding of the larger dataset determined that the session we selected had a high amount of group work and teacher-group interaction, and groupwork was a common participation structure used in this class (Hudson et al., 2021). Additionally, content log and timestamp data of when teacher-group interactions occurred were also coded from the previous study.

Audio analysis and labeling data

For this study, we used openSMILE – an open-source audio processing program (Eyben et al., 2010) to process the audio. We used the Voice Activity Detection (VAD) feature within openSMILE to extract speech turn timestamps. In Palaguachi et al. (2022), researchers modified openSMILE's RMS threshold values to an auto-mode to account for the variability in students' pitch and loudness in group audio, and decreased the turn-segmenting threshold (nPre and nPost) values to detect shorter turns. Similar to Palaguachi et al. (2022), we merged the different group microphone data and categorized them by colors. A duration variable was created by subtracting the turns' end-time by its start time. The duration variable allowed us to note the time each turn of speech lasted whenever speech was detected by the VAD. Lastly, we filtered the dataset to get rid of any turns that were less than .5 seconds because those turns were unlikely to contain meaningful speech information.

To segment the audio before and after teacher group interactions, content log and turn data was merged. This included labelling turns with different group mic, teacher-group interaction, and interaction ID information. We used the content logs to create R scripts that allowed us to label the group discussion patterns detected 30 seconds before and after a teacher-group interactions occurred. We then labeled the type of teacher-group interaction based off our qualitative coding. Through our qualitative coding, we removed interaction 21 because the teacher did not have a verbal interaction with the group. After these changes, there were a total of 29 teacher-group interactions.

Qualitative coding for teacher-group interaction

We qualitatively coded the different types of teacher-group interactions, drawing from Munson's (2019) conferring interaction and Ehrenfeld and Horn's (2020) entry-exit frameworks. Our coding scheme included codes for nudging, eliciting/probing, funneling, and other (i.e., interactions not about student thinking or non-conferring interactions) teacher-group interactions. Interaction 2 was dramatically longer than all the other interactions, and for this reason, interaction 2 was removed from this analysis. For the 28 teacher-group interactions, the inter-rater reliability between two coders yielded a Kappa of 0.884, signifying a high level of agreement. Both coders had a discussion on the three interactions that they did not agree on and came to a consensus.
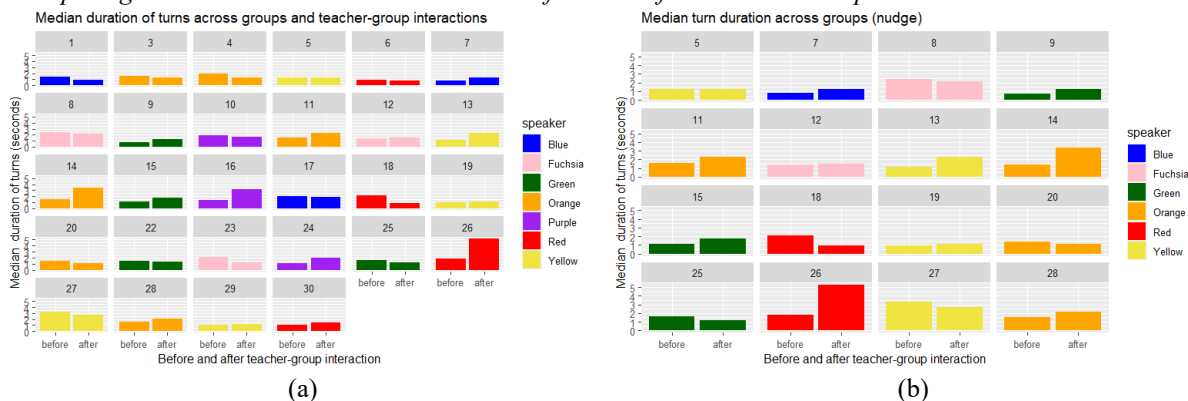
Findings

When examining the group discussion patterns before and after teacher-group interaction, the main metrics we looked at were the median length of turns, the total duration of turns, and the number of turns for each interaction. When comparing the group discussion patterns before and after teacher-group interactions, we found that for the median duration of turns there were 15 out of 28 (54%) interactions that had increased median duration of turns after the teacher-group interaction compared to the group discussion before (see Figure 1a). In further investigating the median duration of turns before and after teacher-group interaction, we found that for some interactions there were instances of a dramatically larger increase in median duration of turns after teacher-group interaction, however, in cases where there were decreases in median duration of turns after teacher-group interactions the decrease tended to be small or a minor decline. Additionally, while exploring

three different metrics to analyze group discussion patterns (median turn duration, total duration of turns, and number of turns) we found, in all three metrics, there were greater percentages of interactions that had increased group discussion after teacher-group interaction.

Figure 1

Comparing Median Turn Duration 30 Seconds Before and After Teacher-Group Interactions

(a)　　　　　　　　　　　　　　　　　　(b)

This is important to keep in mind as we further investigate the differences in conferring interactions and group discussion patterns before and after teacher-group interactions.

Group discussion patterns before and after conferring teacher-group interactions

When examining the group discussion patterns before and after conferring and non-conferring interactions, we used median duration of turns as our key metric. Since nudging involved teacher uptake and student sense-making, we were curious if these interactions were related to increased group discussion after teacher-group interaction occurred. In this analysis we used the median duration of turns metric because it accounted for the skew in the turn duration distribution. When comparing the differences in median duration of turns before and after the conferring interactions, we found that 11 out of 16 (69%) of the nudging interactions had a greater median turn duration after the teacher-group interaction compared to the group discussion before (see Figure 1b). In contrast, 1 out of 5 (20%) of eliciting/probing interactions and 3 out of 7 (43%) of other (non-conferring) interactions had a greater median turn duration after the teacher-group interaction. Since there were a greater number of nudging interactions detected, and since nudging had a greater percentage of increased median duration of turns after teacher-group interaction compared to all the other conferring interactions, we focused on the patterns across nudging interactions.

In further investigating the median duration of turns before and after nudging interactions, we found that the same pattern we observed in the previous analysis of teacher-group interaction was consistent with what we saw when observing the nudging interactions. In the nudging interactions we observed a majority of increased median duration of turns after teacher-group interaction occurred, and when median turn duration decreased after nudging, these decreases were minor. We also observed that the teacher did not use any nudging with the purple group, and only one nudging interaction with the blue group. In contrast, the orange group and yellow group both received four nudges. These findings suggest that when comparing group discussion patterns before and after teacher-group interactions, conferring interactions like nudging have a greater percentage of interactions with increased median turn duration after the teacher leaves the group.

Qualitative findings of conferring teacher-group interactions and group discussion

For eliciting/probing teacher interactions, we found examples of when the teacher-group interaction preceded students asking one another for clarification. In one example (interaction 16), a student shared their thinking about how to convert between radians and degrees as a response to the teacher eliciting their thinking on a problem (i.e., asking "what you got"). Quickly after the teacher steps away, another student in the group asked the student "can you explain that to me?" which began a conversation that was a shift from the silent, individual work the group engaged in before the teacher interaction. Thus, in this case, a student sharing some of the details behind their work to the teacher likely led to the other student asking them to explain their thinking.

Discussion and future work

By exploring the differences in small group discussion patterns before, during, and after teacher-group interaction through this exploratory approach, this study has provided insights on how non-lexical speech analytics can be used to better understand effective teacher-group interactions. Furthermore, our qualitative examples highlight the complex nature of the potential response that comes from quality teacher-group interactions. The implications of this study suggest that non-lexical tools like VAD can be used to explore the quality of teacher-group interactions and that quality teacher-group interactions can enhance student engagement and collaborative dialogue in small groups. In our ongoing work, we plan to expand how we code for teacher-group interactions to better account for the nuances and complexities found within some of the conferring interactions. Other approaches that include ASR will enable us to integrate both non-lexical and lexical processing methods to investigate moments of confusion, frustration, and uncertainty, which are instances where educators might be more likely to intervene. Through the knowledge gained from this study, we suggest further customizing the RMS thresholds values for individual groups (in this case detecting louder turns for quieter small groups) or using individual microphones as two potential solutions. In a future study, we hope to address how these changes to the VAD and how we collect audio data can potentially ameliorate these issues.

Conclusion

To conclude, we learned a lot about the benefits of combining additional qualitative codes to computational analysis of classroom audio. We learned that when working with non-lexical information like turns of speech, context information like content logs, interaction IDs, and conferring interactions are extremely helpful in accounting for how some of these types of interactions effect important turn metrics. By using qualitative methods to code teacher-group interactions and speech analytics to interpret the relationship between teacher-group interactions and group discussion, we gained a deeper understanding of group discussions. Further qualitative coding is necessary to parse through the complexities and nuances of group discussion.

References

Dyer, E. B. (2016). Learning through teaching: An exploration of teachers' use of everyday classroom experiences as feedback to develop responsive teaching in mathematics (Doctoral dissertation, Northwestern University).

Ehrenfeld, N., & Horn, I. S. (2020). Initiation-entry-focus-exit and participation: A framework for understanding teacher groupwork monitoring routines. *Educational Studies in Mathematics*, *103*(3), 251-272.

Eyben, F., Wöllmer, M., & Schuller, B. (2010). Opensmile: The Munich Versatile and Fast Open-Source Audio Feature Extractor. *Proceedings of the International Conference on Multimedia - MM '10*, 1459.

Hudson, H., Parr, E. D., Dyer, E. B. (2021). Factors influencing teachers' use of groupwork in secondary math classrooms. Poster presentation at TN STEM Education Research Conference.

Munson, J. (2019). After eliciting: Variation in elementary mathematics teachers' discursive pathways during collaborative problem solving. *The Journal of Mathematical Behavior*, *56*, 100736.

Palaguachi, C., Cox, E., & D'Angelo, C. (2022). Audio analysis of teacher interactions with small groups in classrooms. *General Proceedings of the 2nd Annual Meeting of the International Society of the Learning Sciences 2022*, (pp. 439 – 442). Hiroshima, Japan: International Society of the Learning Sciences.

Sherin, B., & Star, J. R. (2011). Reflections on the study of teacher noticing. *Mathematics teacher noticing: Seeing through teachers' eyes*, *66*.

Worsley, M., Martinez-Maldonado, R., & D'Angelo, C. (2021). A New Era in Multimodal Learning Analytics: Twelve Core Commitments to Ground and Grow MMLA. *Journal of Learning Analytics, 8*(3), 10-27.

Acknowledgements