Hybrid Science Discourse in Middle School Classrooms: A Mixed Methods NLP Exploration

Kamil Hankour

Rachel Niemira
*Virginia Commonwealth University*

Singith Perara
*Virginia Commonwealth University*

*See next page for additional authors*

Follow this and additional works at: [https://scholarscompass.vcu.edu/education_grad](https://scholarscompass.vcu.edu/education_grad)

Part of the [Educational Psychology Commons](https://scholarscompass.vcu.edu/education_grad)

Downloaded from [https://scholarscompass.vcu.edu/education_grad/3](https://scholarscompass.vcu.edu/education_grad/3)

This Presentation is brought to you for free and open access by the School of Education at VCU Scholars Compass. It has been accepted for inclusion in Education Graduate Presentations by an authorized administrator of VCU Scholars Compass. For more information, please contact [libcompass@vcu.edu](mailto:libcompass@vcu.edu).
Authors
Kamil Hankour, Rachel Niemira, Singith Perara, Kimberly Williamson, Morgan Debusk-Lane, and Christine Lee Bae

This presentation is available at VCU Scholars Compass: https://scholarscompass.vcu.edu/education_grad/3
Hybrid Science Discourse in Middle School Classrooms: A Mixed Methods NLP Exploration

Kamil Hankour1, Rachel Niemira1, Singith Perera1, Kimberly Williamson2, Morgan Debusk-Lane3, Ph.D., & Christine Lee Bae1, Ph.D.
1Virginia Commonwealth University
2Cornell University
3GALLUP Inc.

Purpose and Background
The purpose of this study is to explore applications of natural language processing (NLP), a computer-assisted analytical technique aimed to automatically process, analyze, and comprehend textual data (Lu & Cohen, 2021; Manning et al., 1999) for providing new insights into science classroom discourse processes, including how agency is distributed among the teachers and their students and how students engage in science discourse socially. This inquiry is part of a larger study to develop transformative spaces that center students’ home and community resources as assets to classroom discourse practices (Gutierrez et al., 1999; Moje et al., 2004). Our goal is to understand how middle school classrooms are creating hybrid discourse spaces that organically process, analyze, and comprehend textual data (Liu & Cohen, 2021; Manning et al., 1999) for providing new insights into science classroom discourse processes, including how agency is distributed among the teachers and their students and how students engage in science discourse socially.

Methodology
We used an explanatory sequential mixed methods design for this study (Creswell & Plano Clarke, 2018). This approach allows the researchers to begin with an exploratory qualitative design (Phase 1), followed by a more quantitative descriptive design (Phase 2), and concluding with an integrative phase (Phase 3) to present and utilize in middle school science classrooms? (quant, NLP descriptives)

Context and Research Questions
Data were collected as part of the Science Discourse Project, a multi-year collaboration between the Discourse and Learning lab and local Richmond and Chesterfield area middle school science teachers.

1. RQ1: How are discursive resources (analytic vs. social language, open vs. closed-ended questions) and agency (language targeting via ‘I’ and ‘You’ statements) present and utilized in middle school science classrooms? (quant, NLP descriptives)

2. RQ2: How do discursive resources and agency interact or overlap in hybrid versus academic spaces? (qual, classroom video analysis)

Joint Display of Mixed Methods Findings

Data Collection
Quant Analyses
Qual Analyses
Integration of Quant Results and Qual Findings

Procedure:
Recorded videos in urban middle school science classrooms
Segmented and labeled video transcripts, analyzed using LWC, topic modeling and LASSO regression
Examination of video segments in classrooms context/activity structures
Investigation of qualitative segments informed by NLP results

Product:
8 classroom videos, each from a different middle school science teacher
NLP outputs and frequencies for closed/opened questions, analytic/social language etc.
Qualitative themes related to student and teacher language and behavior in context
Illustrative examples and quotes linked to quant trends and findings, deeper understanding of relationships between qual and quant findings

References available on request: hankourk@vcu.edu

Discussion

- NLP and qualitative methods can be combined to better understand the broader dynamics of classroom discourse (in aggregate) and more in-depth variations in context
- The findings indicate that teachers’ open-and closed-ended questions are primarily analytic (or science-related) in nature; an implication of this is to consider ways to create more inclusive spaces that redistribute agency to their students as appropriate
- Findings are exploratory in nature; cannot infer relationships between talk indices and desired student outcomes
- Limitations:
  - Targeting analyses were limited to “I” and “You” statements and didn’t include variations (e.g., “Y’all”) and other pronouns (e.g., “We”)
  - Findings are exploratory in nature; cannot infer relationships between talk indices and desired student outcomes

Findings are exploratory in nature; cannot infer relationships between talk indices and desired student outcomes

Joint Display of Mixed Methods Findings

Table 1. Descriptives of closed versus open-ended questions by teacher

<table>
<thead>
<tr>
<th>Teacher</th>
<th>Closed</th>
<th>Open</th>
<th>Total Q</th>
<th>Closed %</th>
<th>Open %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teacher 1</td>
<td>120</td>
<td>90</td>
<td>210</td>
<td>57.14%</td>
<td>42.86%</td>
</tr>
<tr>
<td>Teacher 2</td>
<td>140</td>
<td>110</td>
<td>250</td>
<td>56.00%</td>
<td>44.00%</td>
</tr>
<tr>
<td>Teacher 3</td>
<td>180</td>
<td>150</td>
<td>330</td>
<td>54.55%</td>
<td>45.45%</td>
</tr>
<tr>
<td>Teacher 4</td>
<td>220</td>
<td>170</td>
<td>390</td>
<td>56.41%</td>
<td>43.59%</td>
</tr>
</tbody>
</table>

Table 1. Proportion of open and closed questions made by students (Y) and teachers (T) by classrooms

<table>
<thead>
<tr>
<th>Teacher</th>
<th>Open %</th>
<th>Qual %</th>
<th>NLP %</th>
<th>Total %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teacher 1</td>
<td>45.00%</td>
<td>55.00%</td>
<td>45.00%</td>
<td>55.00%</td>
</tr>
<tr>
<td>Teacher 2</td>
<td>40.00%</td>
<td>60.00%</td>
<td>40.00%</td>
<td>60.00%</td>
</tr>
<tr>
<td>Teacher 3</td>
<td>35.00%</td>
<td>65.00%</td>
<td>35.00%</td>
<td>65.00%</td>
</tr>
<tr>
<td>Teacher 4</td>
<td>30.00%</td>
<td>70.00%</td>
<td>30.00%</td>
<td>70.00%</td>
</tr>
</tbody>
</table>

References available on request: hankourk@vcu.edu

References available on request: hankourk@vcu.edu