Understanding Linguistically Diverse Students' Development of Vocabulary During Science Learning Using Semantic Network Analysis

Yinuo Hu, Computer Science & Sociology
Faculty Advisor: Kihyun "Kelly" Ryoo, School of Education

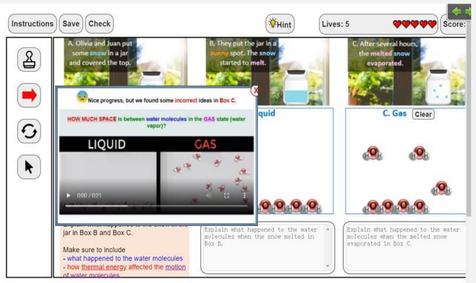
THE UNIVERSITY of NORTH CAROLINA at CHAPEL HILL

Background and Purpose

- Engaging linguistically diverse students in **developing, revising, and explaining scientific models** in pairs can promote their understanding of unobservable scientific phenomena through discourse-rich practices.
- Exploring linguistically diverse students' conversations may reveal patterns of their **vocabulary development** during scientific modeling practices.
- Previous work showed that **Semantic Network Analysis (SNA)** has the potential to visualize the structure and relationships among the use of vocabulary.
- Given the limited research in applying SNA on pairs' discourse during scientific modeling practices, the study explored the following research questions (RQs):

RQ1: What's the structure of linguistically diverse students' vocabulary and the relationships among the words?
RQ2: How does students' vocabulary development change over time?

Analysis Methods

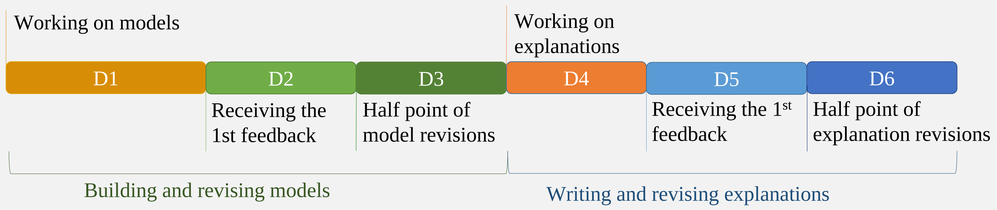


- As part of a larger NSF project that explores how visualizations can improve 8th-grade linguistically diverse students' science learning, students worked in pairs to develop models and write explanations of how thermal energy affects the state of water molecules during a phase change.
- Pairs used automated feedback to revise models and explanations.

- This SURF project used **16 transcripts** from the video files that were collected and analyzed as part of the larger NSF project.
- This resulted in 5649 talk turns of pairs.

Divisions of Modeling Activities

- Students' talk turns were divided into six divisions based on the following:



Semantic Network Analysis (SNA)

- Data was cleaned up by removing stop words (e.g., the, a) and lemmatization.

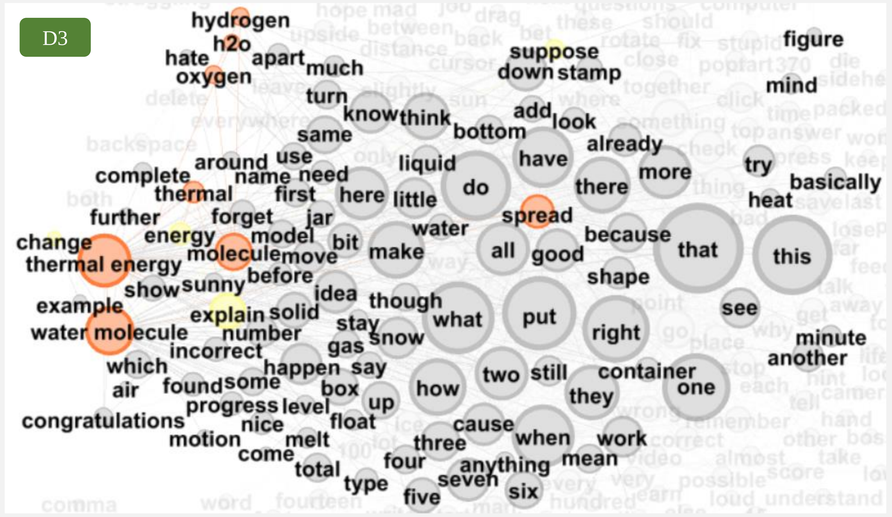
changes, changed, changing → changes

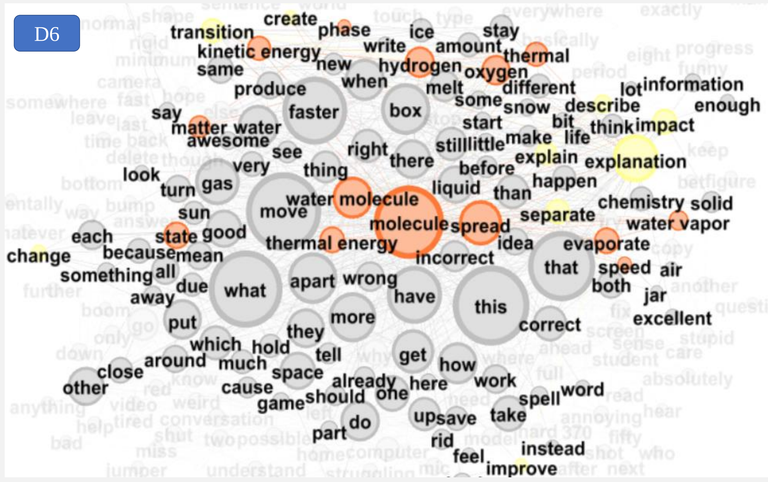
- Any words that were used more than three times were included in a word list.
- Building on the existing list from the larger project, this SURF project added additional words from the analysis to revise the Three Tier Model (Beck et al., 2002) through iterative processes.

Tier 1: Basic Words	**Tier 2**: General Academic Words	**Tier 3**: Specific Content Words
e.g., ice, water	e.g., describe, explain	e.g., water molecule, state

- The project used KBDeX, R, and Gephi to visualize the semantic networks.

Results





Tier 1 Tier 2 Tier 3

- Tiers 1, 2, and 3 words were mixed in D1.
- Basic words were the center of the D1 network.
- Tiers 2 and 3 words were weakly connected.

- A cluster of Tiers 2 and 3 words showed better connections of key concepts.
- Tiers 2 and 3 words were still not the center of discussions yet.

- Some Tier 3 words (e.g., water molecule, thermal energy) were at the center of the network in D6.
- Tiers 2 and 3 words moved closer and formed stronger connections among each other.

Conclusion

- With SNA, the study explored the structure and development of students' discussions as they gradually centralized around several key scientific concepts.
- Students made stronger connections of academic and content-specific words through discourse-rich collaborative modeling practices.
- SNA can be more widely used in the future to understand the vocabulary development of students in different collaboration settings or in other disciplines.

Acknowledgement: This material is based upon work supported by the National Science Foundation under Grant No. #1552114. The project is supported by UNC Summer Award of Undergraduate Research (SURF).