Analysis of Teachers' Language Input Using Morpholex

Faith Collins, Ms. Audrey Hendrix, Dr. Sana Tibi, PhD, Dr. Carla Wood, PhD, CCC-SLP
Florida State University, College of Communication Sciences and Disorders

Background

Morphological complexity (MC) describes the intricacy with which words are formed using meaningful parts or morphemes, such as prefixes and suffixes when added to base words. Conversational language usually has low MC with a high frequency of grammatical markers or inflectional morphemes (i.e. -s, -ed). Academic language, which is often encountered in educational settings and therefore is important for students' success, **has higher MC due to a greater frequency of derivational morphemes** (i.e., -ly, -ism) which contributes to the complexity of academic words (Zhang et al., 2020).

Extensive research describes the role of robust adult language input in children's conversational language development. However, few research studies have examined the **MC of teachers' instructional language** and its implications in academic language development.

Research Questions

(a) What types and frequencies of derivational morphemes occur in teachers' instructional dialogues?
(b) Do morpheme types and frequencies differ by instructional subject?

Methods

- **This study analyzed 15-minute transcripts (n ~1500) of second-grade teachers' (n =40) instructional dialogues using Morpholex** (Cobb, 2023; Laufer & Cobb, 2020), an online affix profiler that describes the morphological complexity of written samples according to Bauer and Nation's (1993) Levels of Morphological Complexity. Bauer and Nation's Levels categorize derivational morphemes by frequency and regularity of spelling and pronunciation.
- Transcripts were collected during a previous study analyzing teachers' vocabulary during instructional dialogues (Wanzek et al., 2021). During data processing in the original study, transcripts were formatted according to the **Systematic Analysis of Language Transcripts (SALT)** conventions, which separate bound morphemes from base words using special conventions. In order for Morpholex to accurately describe morphological complexity, SALT formatting needed to be reverted to plain text. Given the number of transcripts, a SALT reformatting code was developed in Visual Basic for Application (VBA) for efficient data processing.
- High-frequency words from the BNC-COCA word family lists (Nation, 2017), Dolch's sight word list (Dolch, 1936), Fry's high-frequency word list (Fry, 1980), and Coxhead's Academic (2000) and science word (Coxhead & Hirsh, 2007) lists were combined and formatted according to SALT conventions. The ensuing wordlist was reformatted using the code. Patterns of errors and abnormalities in the reformatted wordlist were identified and addressed in the code in order to improve its accuracy.
- **A random sample of 40 transcripts, with ten from each educational subject (English Language Arts, math, science, and social science)**, was selected and processed by the code created. Preliminary data on morphological complexity was collected. The processed transcripts were then analyzed for any unaddressed spelling errors or formatting issues. The code was then modified to address these issues and the sample transcripts were run through the code again. This process was repeated until the error rate was as minimal as possible, given the time restraints and resources available.

Figure 1
Average Total Morpheme Counts

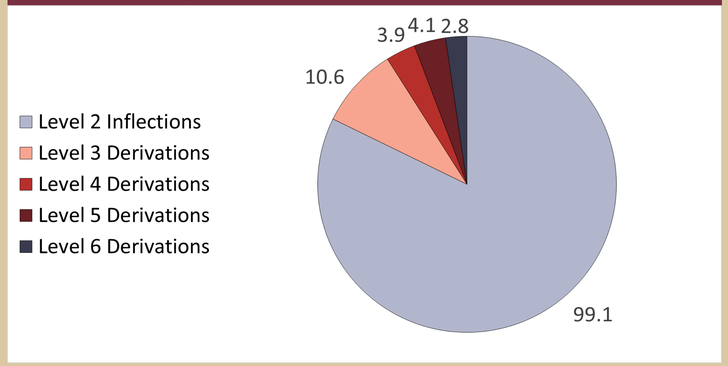


Figure 2
Total Token Counts

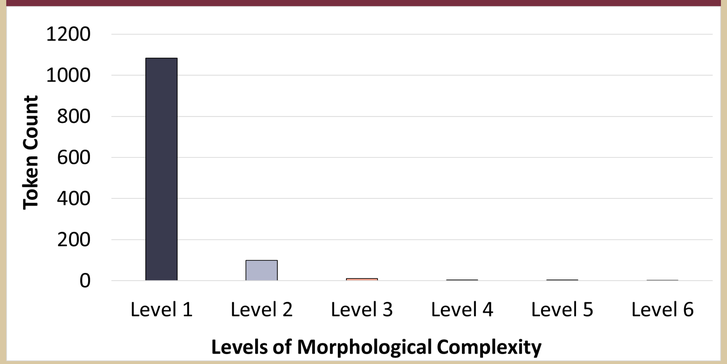
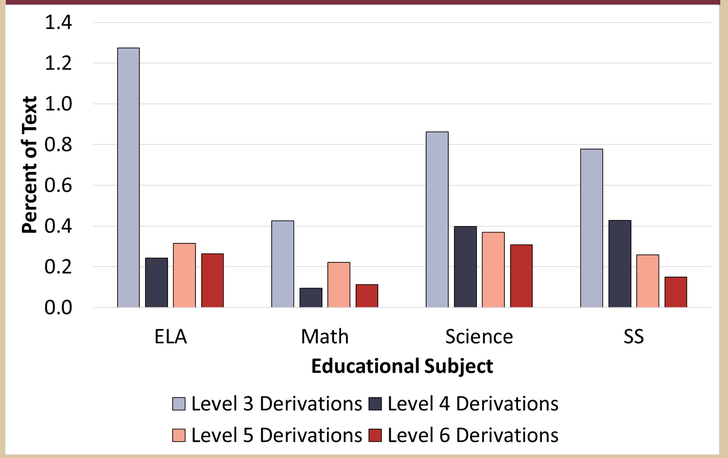


Figure 3
Frequencies of Derivational Morphemes by Subject



Preliminary Results

To address the first research question, the preliminary results indicated that the frequency of **derivational morphemes is low across the sample.** Base words (Level 1) made up an average of 88.9% of all individual words from sample text, while base words with inflectional morphemes (Level 2) made up an average of 8.1% of individual words. This means that only 3% of the text from the sample transcripts is composed of derivational morphemes (Levels 4-6).

To address the second research question, the preliminary results indicated that there are in fact **differences in morpheme level frequencies between the subjects.** For most educational subjects, but particularly science and social science, morpheme counts decrease as the Levels increase. While Level 3 morphemes occur most often in all subjects, English has the highest amount with 1.3% of words having a Level 3 affix. Math is the subject with the lowest percentages for every level, but it has a similar pattern to English. For these two subjects, the average number of Level 4 morphemes is less than the Level 5 and 6 averages for the respective subjects.

Preliminary Conclusion

The data collected thus far indicates that there is a **low frequency of derivational morphemes** present in the second-grade teachers' instructional dialogues across all subjects. When comparing the frequencies of morphemes in different Levels across the educational subjects, English and Math appear to share one pattern while Science and Social Studies share another. Statistical significance cannot be determined at this point due to a small sample size.

Discussion

- Previous research has examined the vocabulary (Wanzek et al., 2021) and syntax (Buchheit, 2023) of teacher instruction, but none has deeply examined the **variation in morphological complexity** across educational subjects.
- Preliminary results indicate differences across the subjects regarding the complexity of morphemes, indicating future training on morphological instruction may be tailored to each subject.
- Further research can examine different ways teachers can either incorporate more derivational morphemes into their teaching or can focus **on how teachers can bring attention to the morphemes currently present** in their instruction so children can have a better understanding of derivational morphemes.

Acknowledgements

I would first like to thank my research mentor Ms. Audrey Hendrix for all of her help and support throughout this project. I would also like to thank Dr. Sana Tibi for her guidance during my time as an assistant, and Dr. Carla Wood for allowing this project to use the transcripts gathered in a previous project.

Bauer and Nation's Levels

References