Text and Word Level Differences in Children's Books by Genre

Ana De Freitas[3], Leo Raden[4], Rebecca M. Vasile[1,2]

School of Teacher Education, Anne Spencer Daves College of Education, Health, and Human Sciences, Florida State University[1]
Florida Center for Reading Research, Florida State University[2]
College of Nursing, Florida State University[3]
College of Arts and Sciences, Florida State University[4]

INTRODUCTION

Prior research shows that language used in children's picture books differs from everyday speech, often containing more diverse and advanced vocabulary that is important for language and later reading development (Dawson et al., 2021; Flack et al., 2018; Montag et al., 2015).

Children's books vary based on their genre. Researchers and advocates alike have pushed a more balanced diet of narrative and informational texts given the benefits of both (Duke, 2000; Wright et al., 2022).

Examination of informational and narrative texts in early childhood and fourth grade settings have found differences between the genres (Green and Keogh, 2024; Hiebert and Cervetti, 2011).

The poster compares **informational and narrative texts** read in early childhood education classrooms and explores **whether they differ based on text and word level** characteristics.

METHOD

Analytic Plan:

This poster utilizes corpus-based methods to analyze the language of informational and narrative texts at the text and word levels.

Classrooms and Teachers:

Early childhood education classroom observational data were collected across three cohorts of teachers from 2015-2018. Classrooms were in three Southeastern US states and served children ranging from 3 to 5 years olds.

- 3-to-4-year-old classrooms - 19 (22%)
- 4-to-5-year-old classrooms - 31 (36%)
- Mixed age group classrooms - 36 (42%)

The classroom observations were intended to represent typical classroom activities. 86 teachers' classrooms were observed.

Books:

During those daily observations, teachers were recorded during any shared book reading interactions. Every title was noted and then condensed into a list (n = 597).

Currently, (***n* = 451**) books from this list have been transcribed into a corpus with **266 narrative and 175 informational texts**.

Genre coding was completed using the Reading Log Coding Guide (Pentimonti et al., 2018).

Table 1: Genre Examples

GENRE	EXAMPLE TITLE
Narrative	How the Grinch Stole Christmas by Dr. Seuss
Informational	All About the Human Body by Tish Rabe
Mixed	The Three Little Pigs by Paul Galdone
Other	Alphabet ABC by Aurora Chien

RESULTS

Frequency of Book Language in Child Directed Speech
Using the CHILDES corpus of child-directed speech (MacWhinney, 2000), we examined the frequency of the words in the texts by genre.

In this sample, the words in informational texts appear more in child-directed speech than the words in narrative texts.

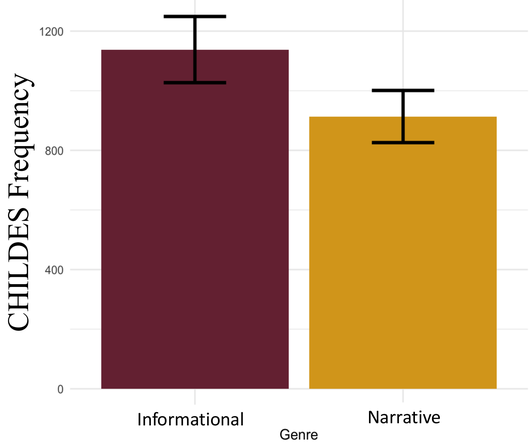

Figure 1: Bar graph of informational and narrative text frequency in child-directed speech. Bars are standard error of the mean.

Broad Descriptives: Type Token Analysis:

The full corpus (n = 451) has a total of 11,113 unique words (i.e., types) to over 100,000 total words (i.e., tokens).

The distribution of the words in the corpus follow a Zipf distribution, which is seen in other language corpora (i.e., the most common type in the corpus is the word "the" [n = 9,615] followed by "and" [n = 5,604]).

Overlap of Book Language in Informational and Narrative

In this sample, only 35% (n = 3,887) of the unique words in the total corpus overlap across genres.

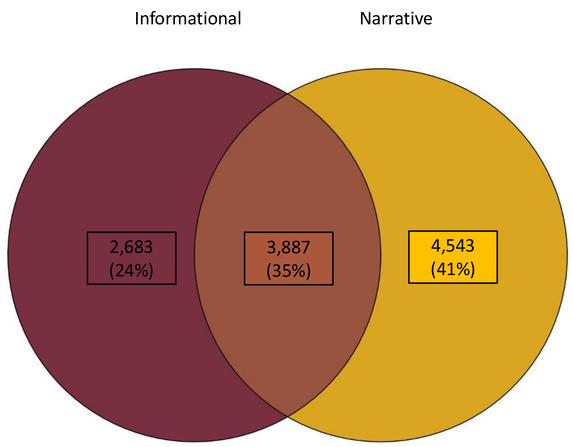

Figure 2: Venn Diagram displaying the distribution of unique words in informational and narrative texts.

RESULTS - *TEXTS*

Frequency of Questions in Informational versus Narrative texts
In this sample, there are more questions in informational texts than in narratives.

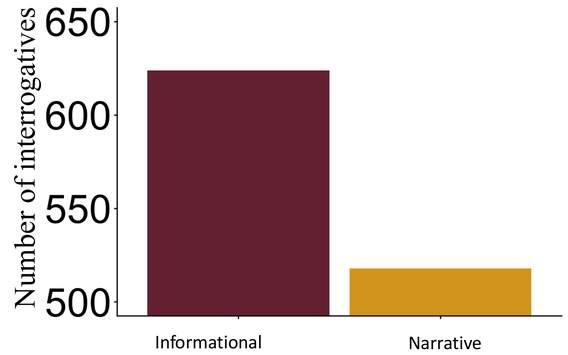

Figure 3: Bar graph of the number of questions in informational and narrative texts. Bars represent total observations per level of genre. Sampling procedure controlled for difference in quantity of each genre in the sample.

Length of Utterance in Informational versus Narrative texts
The length of utterances here are measured by words per sentence.

In this sample, there is a small numerical difference between length of utterance for informational and narrative texts.

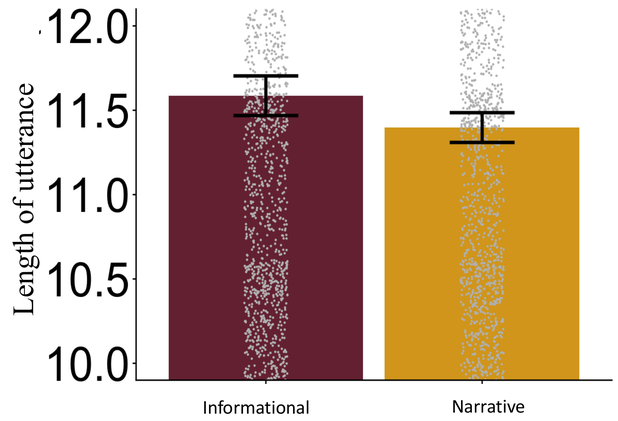

Figure 4: Bar graph displaying the mean length of utterances of informational and narrative texts. Points are sentences within texts. Bars are standard error of the mean.

RESULTS - *TEXTS*

Concreteness of Book Language by Genre
To examine concreteness in informational and narrative texts, ratings from Brysbaert et al., 2014 were used.

The rating scale is a 5-point-scale (i.e., 1 = abstract, 5 = concrete).

The language of informational books are more concrete than narrative book language.

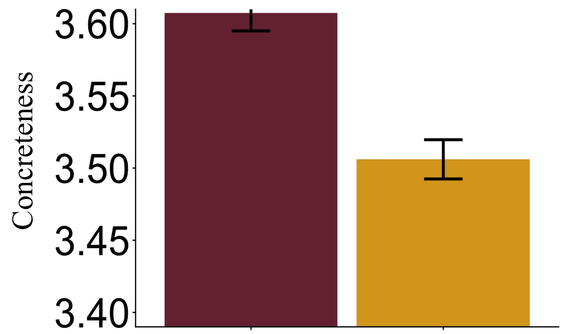

Figure 5: Bar graph of the mean concreteness of words across informational and narrative texts. Bars are standard error of the mean.

Number of Phonemes in Words by Genre
There is a small numerical difference in the number of phonemes in informational and narrative texts.

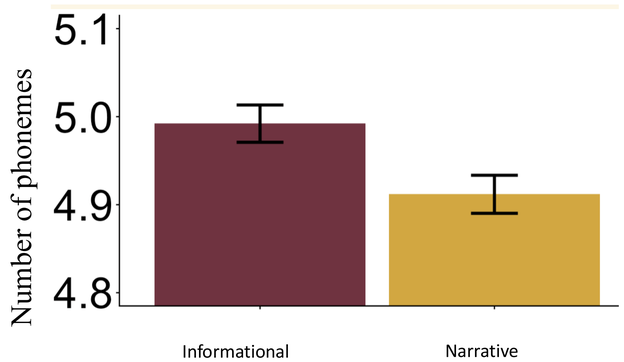

Figure 6: Bar graph displaying mean number of phonemes in the words in informational and narrative texts. Bars are standard error of the mean.

CONCLUSIONS

The distributions of the variables examined here between informational and narrative texts appear to be different based on descriptive analyses.

- Informational texts use words more common in everyday child-directed speech.
- Only 35% of unique words appeared in both types of texts.
- Informational texts have more questions than narratives texts.
- Informational text language is more concrete than the language of narrative books.

Overall, these findings indicate variability in the language children are exposed to through narrative versus informational texts. The results suggest that early childhood educators should provide children with both narrative and informational texts to create balanced language learning experiences.

REFERENCES

The work here was also funded by grant number R305A2204420 from the Institute of Education Sciences (PI Beth Phillips). The work was also funded by grant number H325D210062 from the U.S. Department of Education, Office of Special Education and Rehabilitative Services Preparation of Special Education, Early Intervention, and Related Services Leadership Personnel (PI Laura Steacy). Lastly, the work here was also funded by the Spencer Foundation (PI Beth Phillips Lyle Spencer Award). No authors have any relevant nonfinancial relationships to disclose. The opinions expressed are those of the authors and do not represent views of the funding agencies or the university.