HOWARD FORGE PRESS

Methodology and Research

The inspiration for this readability program came from my dyslexic daughter. She read an early draft of our historical novel, *William and Hannah*, and when she finished, she said it was a 'super sweet, easy read.' I was floored. Having helped her through years of school, from high school into college, I had seen firsthand how she struggled with nearly every assigned reading. I had never heard her describe a book that way before. So I thought to myself: 'Maybe this book really is Dyslexia-Friendly. Let's find out why.'

Building the Model

From that inspiration, we began to explore which features of text might contribute to readability for struggling or dyslexic readers. We focused on elements that were both meaningful and measurable — things a computer could detect and compare reliably.

Key features included:

- Sentence Count
- Word Count
- Average Sentence Length
- Average Word Length
- Average Syllables per Word
- Passive Sentence Count
- Rare or Abstract Words
- Sensory Words

Scoring Algorithm

We created a scoring system based on these features and trained it using known examples of accessible and challenging prose. We tested early versions of the model against multiple text samples:

Dyslexia-Friendly Text Samples:

- *The Lightning Thief* by Rick Riordan
- *Lost Ross* by Catherine Parr, a novelist who specializes in dyslexia-friendly writing
- An early version of *William and Hannah*

Challenging Text Samples:

- *Great Expectations* by Charles Dickens
- *Moby-Dick* by Herman Melville
- *The Scarlet Letter* by Nathaniel Hawthorne

Validation and Visualization

To better understand our results, we developed a validation tool that creates charts and heat maps of readability scores across full novels. This allowed us to test whether our algorithm could meaningfully distinguish easy and difficult passages — not just between books, but between chapters.

Understanding the Z-Score Heat Map

As part of our validation tool, we use **Z-score heat maps** to visualize how each chapter or section of a book compares to the others in terms of readability.

What's a Z-Score?

A **Z-score** tells you how far a value is from the average, measured in standard deviations. In our tool, we calculate the Z-score for each readability feature in each chapter:

- A **Z-score of 0** means the feature is exactly average for the book
- A **positive Z-score** means it's higher than average (more complex)
- A **negative Z-score** means it's lower than average (easier or simpler)

For example:

- A chapter with a **+2.0 Z-score for sentence length** has sentences much longer than the book's average
- A **−1.5 Z-score for word complexity** means the words are significantly simpler than usual

What Do the Colors Mean?

Color	Interpretation
🔵 Deep Blue	Very low Z-score (easy/readable)
🟦 Light Blue	Slightly below average
⚪ Neutral	Around the average
🟥 Light Red	Above average (potentially harder)
🔴 Deep Red	Very high Z-score (harder to decode)

These colors are mapped **per feature**, not per score — so you can visually track which features (like sentence length or passive voice) are spiking or dipping in each chapter.

Why It's Useful

This system allows educators, parents, and authors to:

- **Pinpoint chapters or passages** that may present decoding challenges
- **Compare patterns across books** (e.g., is Chapter 1 always the hardest?)
- **Visualize which features are most responsible** for a drop in accessibility

The heat map doesn't just tell you a chapter is "harder" — it tells you **why**, in terms that can be addressed, revised, or supported through scaffolding.

Current Status and Future Work

We are now in the process of refining the model and expanding its features. One area of active development is measuring cognitive load — not just decoding difficulty, but the mental effort required to follow complex ideas, symbolism, or layered narratives. These improvements are still experimental, but we hope to share updates soon.