Text Messages: Examining Different Estimates of Text Complexity

Joseph F. T. Nese
Julie Alonzo
Gina Biancarosa
Akihito Kamata
Josh Kahn

February, 2017
Poster presented at the Pacific Coast Research Conference annual meeting
Text Messages: Examining Different Estimates of Text Complexity

Joseph F. T. Nese, Julie Alonzo, Gina Biancarosa, Akihito Kamata, Josh Kahn

Behavioral Research and Teaching, University of Oregon, 2017

Purpose
Text complexity has received increased attention due to the CCSS5, which call for students to comprehend increasingly more complex texts as they progress through grades. Quantitative text complexity (or readability) indices are based on text attributes (e.g., sentence length, and lexical, syntactic, & semantic difficulty), quantified by mathematical formulas to estimate text complexity. Researchers have warned that texts should not be created to reach particular readability scores2-3, a frequent practice of “cooking” ORF passages. Issues we encountered were: estimates often differ; estimates do not account for student variance; and many require tests of a certain length. The purpose of this study was to provide comparisons of quantitative text complexity estimates based on text features to estimates of “text difficulty” based on student performance. Specifically, these comparisons are situated in the context of curriculum-based measurement (CBM) assessment of oral reading fluency (ORF), where the passages range from 20 to 105 words.

Method

Participants
P10 students (Grade 2 = 259, Grade 3 = 329, Grade 4 = 342) were assessed online, using laptops in one-to-one administration, during which each student read approximately 3 long, 5 medium, and 10 short passages.

Passages
We administered 330 ORF passages – 110 at Grades 2, 3, 4. Each passage was an original work of narrative, and targeted readability at the mid-year level for each grade.

Text Complexity

Formality5 Many measures inform 5 principal components: narrativity, syntactic simplicity, word concreteness, referential cohesion, deep cohesion (referential cohesion + deep cohesion – narrativity - syntactic simplicity - word concreteness) / 5.

Automated Readability Index (ARI): based on letter count – 4.73 (letters/words) + 0.5 (words/sentences) – 21.43

Flesch–Kincaid (grade): based on syllable count (less reliable than letter count) – 0.39 (average sentence length) + 11.8 (average syllables per word) – 15.59

WCPM (words correct per minute): Recorded audio files scored by trained assessors (using the same scoring rules as traditional ORF procedures). – (total words read – words read incorrectly) / (60 / computer recorded sec duration)

Latent fluency: Bi-factor CFA model included a general factor (fluency) and two group factors (speed and accuracy). Observed scores were at the sentence level for each reading; accuracy = percent of words read correctly; speed = average sec per word. Models estimated separately by grade.

Analysis

Spearman’s rank correlations were computed between the text complexity estimates. All analyses conducted with R:

Text Complexity Estimate Not Used

Estimation methods – Based on texts of 100 words.

Spearman's rank correlation coefficient: 0.06

Table:

<table>
<thead>
<tr>
<th>Grade</th>
<th>Length</th>
<th>Formality</th>
<th>ARI</th>
<th>Latent fluency</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>2.0</td>
<td>0.65</td>
<td>1.07</td>
<td>0.56</td>
</tr>
<tr>
<td>3</td>
<td>3.0</td>
<td>0.62</td>
<td>1.32</td>
<td>0.58</td>
</tr>
<tr>
<td>4</td>
<td>4.0</td>
<td>0.51</td>
<td>1.90</td>
<td>0.67</td>
</tr>
<tr>
<td>Short</td>
<td>0.27</td>
<td>0.75</td>
<td>2.30</td>
<td>0.82</td>
</tr>
<tr>
<td>Medium</td>
<td>0.53</td>
<td>1.34</td>
<td>3.63</td>
<td>0.67</td>
</tr>
<tr>
<td>Long</td>
<td>0.72</td>
<td>0.22</td>
<td>5.00</td>
<td>0.72</td>
</tr>
</tbody>
</table>

Conclusions

FK aligns closest to grade level as an artifact of the passage development

WCPM and Latent fluency generally increase across lengths and grades

All estimates increase across grades except Formality, which remains fairly stable across Grades 3 & 4

FK and ARI FLUENCY remains stable across lengths

In our study, it was not possible to examine the Latent fluency correlations between grades because they were modeled independently. Also, WCPM and Latent fluency functioned differently by length; the former with bimodal distributions and the latter with near-normal, which was likely a function of the model choice.

Moderate correlation among Latent fluency and WCPM, and the only consistently moderate relation within grades

The measure with the lowest correlations among all measures was Formality = Perhaps a different dimension of text complexity

Highest correlations observed among ARI, FK, and WCPM by length (which increases by length) = Partly an artifact of passage development, and partly a similarity between the formulas of ARI and FK, and the only consistently moderate correlation within grades

But, based on within grade correlations and previous research (FK), this is a spurious relation resulting from the “developmental” nature of the scales.

Discussion

- Does CCSS conflate text complexity with text difficulty? If quantitative text complexity is a function of text features, can ORF (wcpm, latent fluency) define the interaction between text and reader (“reader-task considerations”)?

- Text complexity estimates are based on entire text, but students most often do not read entire ORF passage in 60 sec.

- What is a better reflection of “comprehension” – measure of text or student ORF performance?