

Technical Report # 2302

**The Development of Oral Reading Fluency Assessments with
Automatic Speech Recognition in CBMSkills**

Aaron Glasgow

Gerald Tindal

Joseph F. T. Nese

University of Oregon



behavioral research & teaching

Published by

Behavioral Research and Teaching
University of Oregon • 175 Education
5262 University of Oregon • Eugene, OR 97403-5262
Phone: 541-346-3535 • Fax: 541-346-5689
<http://brt.uoregon.edu>

Copyright © 2023. Behavioral Research and Teaching. All rights reserved. This publication, or parts thereof, may not be used or reproduced in any manner without written permission.

APA Reference: Nese, J. F. T., Glasgow, A., & Tindal, G. (2023). *The Development of Oral Reading Fluency Assessments with Automatic Speech Recognition for Use in CBMSkills*. University of Oregon: Behavioral Research and Teaching. Technical Report Number 2302.

In this technical report, we describe the development of an oral reading fluency assessment system that uses automatic speech recognition with immediate feedback provided on words read correct, incorrect, and words correct per minute (WCPM).

The University of Oregon is committed to the policy that all persons shall have equal access to its programs, facilities, and employment without regard to race, color, creed, religion, national origin, sex, age, marital status, disability, public assistance status, veteran status, or sexual orientation. This document is available in alternative formats upon request.

Abstract

This technical report describes the development of oral reading fluency passages for use in the CBMSkills assessment system. Seventeen passages were created for each of Grades 1 through 5. Passages were designed to target the fluency component of a developmental model of reading. All passages were written with the specification to use language and structure common for the grade level with similar cultural sensitivity and lack of bias to modern classroom materials. Each fictional passage was between 100 to 110 words in length. All passages were reviewed and edited by two experts in reading assessment for errors (e.g., format and grammatical), and bias (e.g., gender, cultural, religious, and geographical).

The Development of Oral Reading Fluency Passages for Use in CBMSkills

CBMSkills is a diagnostic system designed by researchers working with teachers to identify gaps students' skills and tailor classroom instruction to meet the needs of every student. The system includes reading and mathematics assessments aligned to curriculum standards from the Kindergarten through Grade 5. In this technical report, we describe the development of the Oral Reading Fluency (ORF) assessment and initial research on its accuracy and its diagnostic focus. We begin with a brief overview of ORF as a construct and an assessment. We then provide information on the specifications for passage development. Finally, we describe the ORF assessment application in CBMSkills with sample results from a student.

ORF Overview

Strong theoretical and empirical research supports ORF as an essential component of reading proficiency. In 2000, the National Reading Panel (NRP) concluded that reading fluency was one of five areas deemed central to learning to read, and that instruction promoting fluent reading is critical through elementary school, and that ORF is a malleable factor, amenable to instructional influence (NICHD, 2000). Reading with fluency decreases the cognitive demand of identifying and grouping words and allows for the simultaneous cognitive processing that underlies comprehension of text (Rasinski, 2006). Literacy experts consider fluency to be a critical component of reading development (Rasinski et al., 2006; Samuels & Farstrup, 2006).

For over three decades, researchers have documented the strong association between ORF and reading proficiency (see Tindal, 2013) and its assessment has become ubiquitous in most in elementary classrooms (Shapiro et al., 2006; Speece et al., 2003; Wood, 2006). Often, it is used to universally screen for students at risk of low reading proficiency to (a) ensure they are meeting teacher expectations, (b) monitor progress of those receiving reading intervention or identified as having a disability, and (c) predict their future reading outcomes (e.g., Baker et al.,

2008; Kim et al., 2010; Speece & Ritchey, 2005). Thus, it is critical to assess students' ORF to ensure adequate growth toward reading proficiency.

Passage Development

The general plan on passage development was to create short fictional stories of approximately 100 words (like those used in informal reading inventories) that could be assessed with automatic speech recognition (ASR). Under a contract with SoapBox Labs, these passages were trained to provide the following information: words correct per minute (WCPM), substitutions (when a student says a different word than expected in the reference text), omissions/deletions (when a student omits words present in the reference text), insertions (when a student's audio file contains extra words not present in the reference text), repetitions (when a student repeats a word from the reference text), pauses (when a student pauses for 1 to 3 seconds while reading), hesitations (when a student pauses for three seconds or more while reading), and self-corrections (when a student corrects a word they have just read within 3 seconds). Only substitutions and omissions were recorded as incorrect, which is consistent with the original research conducted by Deno, Mirkin, and Chiang (1982), based on the thesis that insertions and repetitions reduced the number of WCPM (and therefore counting them as incorrect doubles the penalty). Because this measure is self-administered by the computer, hesitations and pauses were addressed as needed by the teacher in either initial training of students or in monitoring them as they read.

Three grade school educators were recruited to write 17 passages for each of Grades 1 through 5 (85 passages total). All passages were written with the specification to use language and structure common for the grade level and have similar cultural sensitivity and lack of bias to modern classroom materials. Each passage was written to be over 100 words in length, but not more than 110 words. The fictional passage could be any topic of the author's choosing and did

not require a title. All passages were reviewed and edited by two experts in reading assessment for errors (e.g., format and grammatical), and bias (e.g., gender, cultural, religious, and geographical).

A readability score was computed for each passage to estimate the passage's 'difficulty' or grade-equivalence using attributes such as word count and word lengths, sentence lengths, and syllable counts. The Spache index (Spache, 1953) was used to estimate readability for passages in Grades 1 through 3, and the Dale-Chall index (Dale, Chall, 1948) was used to estimate readability for passages in Grades 4 and 5 (<https://readabilityformulas.com>). The Spache formula calculates the grade level of a text sample based on sentence length and number of unfamiliar words (i.e., words that are not in the Spache Revised list of words that are familiar to most third-grade students). The Dale-Chall formula calculates the grade level of a text sample based on sentence length and the number of 'hard' (i.e., words that do not appear on the Dale-Chall list of common words that are familiar to most fourth-grade students). Table 1 below shows text features and readability estimates for each of the 85 passages across Grades 1 to 5.

Table 1

Passage Text Features and Readability Estimates

Passage ID	Unique Words	Repeated Words	Words per Sentence	Double Syllable Words	Single Syllable Words	Spache Level	Spache Grade	Dale-Chall Grade	Dale-Chall Level		
Grade 1											
27625	112	55	57	8	14	15	97	2.3	2	--	--
27626	104	58	46	8	13	4	95	2	2	--	--
27627	109	55	54	8	14	15	94	1.8	2	--	--
27628	97	49	48	8	12	9	84	1.8	2	--	--
27629	101	59	42	9	11	10	90	1.9	2	--	--
27630	114	76	38	6	18	23	89	1.6	2	--	--
27631	114	61	53	8	14	11	102	1.8	2	--	--
27632	116	57	59	6	21	8	102	2.1	2	--	--
27633	109	77	32	11	10	16	92	2.6	3	--	--
27634	108	72	36	11	10	14	89	2.2	2	--	--
27635	106	68	38	5	20	12	94	2.1	2	--	--
27636	110	65	45	8	14	14	95	2.2	2	--	--
27637	112	67	45	9	13	20	90	2.4	2	--	--
27638	116	79	37	10	12	18	97	2.2	2	--	--
27639	123	71	52	9	13	17	106	2	2	--	--
27640	105	58	47	10	11	5	99	1.8	2	--	--
27641	105	66	39	8	14	11	92	1.6	2	--	--

Passage ID	Unique Words	Repeated Words	Words per Sentence	Double Syllable Words	Single Syllable Words	Spache	Spache Grade Level	Dale-Chall	Dale-Chall Grade Level		
Grade 2											
27642	125	80	45	7	19	23	102	2.5	3	--	--
27643	118	72	46	10	12	7	107	2.3	2	--	--
27644	109	74	35	8	14	28	81	2.3	2	--	--
27645	114	77	37	9	13	21	91	2.5	3	--	--
27646	108	64	44	7	15	9	98	2.3	2	--	--
27647	105	74	31	8	14	23	81	2.1	2	--	--
27648	100	60	40	8	12	15	84	2.1	2	--	--
27649	107	73	34	8	14	13	92	1.7	2	--	--
27650	103	71	32	10	10	12	84	2.6	3	--	--
27651	103	74	29	9	11	19	79	2.8	3	--	--
27652	107	79	28	7	16	24	82	2.3	2	--	--
27653	115	83	32	10	12	24	89	2.9	3	--	--
27654	111	78	33	12	9	16	93	2.9	3	--	--
27655	112	75	37	11	10	22	90	2.4	2	--	--
27656	101	71	30	8	13	9	81	1.6	2	--	--
27657	99	65	34	10	10	12	83	2.3	2	--	--
27658	112	63	49	11	10	15	97	2.3	2	--	--

Passage ID	Unique Words	Repeated Words	Words per Sentence	Double Syllable Words	Single Syllable Words	Spache	Spache Grade Level	Dale-Chall	Dale-Chall Grade Level		
Grade 3											
27659	108	75	33	10	11	22	85	2.7	3	--	--
27660	118	76	42	11	11	18	98	2.4	2	--	--
27661	109	65	44	12	9	18	88	3	3	--	--
27662	106	60	46	11	10	17	86	2.5	3	--	--
27663	108	83	25	9	12	28	78	3	3	--	--
27664	106	75	31	9	12	21	81	2.2	2	--	--
27665	111	82	29	7	15	17	90	2.2	2	--	--
27666	103	69	34	9	11	15	80	2.2	2	--	--
27667	107	76	31	9	12	20	86	2.1	2	--	--
27668	107	77	30	8	14	26	79	2.6	3	--	--
27669	105	75	30	10	11	9	92	2.1	2	--	--
27670	112	79	33	7	16	19	90	2.2	2	--	--
27671	102	75	27	10	10	19	78	2.7	3	--	--
27672	101	75	26	8	12	13	88	1.8	2	--	--
27673	102	62	40	11	9	19	83	2.7	3	--	--
27674	103	73	30	13	8	15	87	2.9	3	--	--
27675	102	80	22	11	9	14	84	2.8	3	--	--

Passage ID	Unique Words	Repeated Words	Words per Sentence	Double Syllable Words	Single Syllable Words	Spache Level	Dale-Chall Grade Level				
Grade 4											
27676	110	80	30	12	9	17	91	--	--	5.8	5 to 6
27677	111	69	42	11	10	21	84	--	--	6	7 to 8
27678	111	64	47	11	10	26	83	--	--	5	5 to 6
27679	104	75	29	10	10	22	74	--	--	5.4	5 to 6
27680	108	76	32	11	10	14	93	--	--	5.6	5 to 6
27681	111	84	27	10	11	17	89	--	--	5.8	5 to 6
27682	101	77	24	13	8	19	68	--	--	6.3	7 to 8
27683	113	79	34	9	13	16	90	--	--	5.5	5 to 6
27684	103	70	33	11	9	17	78	--	--	5.3	5 to 6
27685	109	81	28	10	11	17	89	--	--	6.2	7 to 8
27686	130	84	46	13	10	22	104	--	--	5.5	5 to 6
27687	110	78	32	12	9	20	87	--	--	6.7	7 to 8
27688	104	72	32	10	10	22	75	--	--	5.8	5 to 6
27689	113	79	34	11	10	22	81	--	--	5.9	5 to 6
27690	100	72	28	11	9	25	72	--	--	5.5	5 to 6
27691	106	81	25	11	10	16	79	--	--	5.2	5 to 6
27692	105	79	26	10	11	16	77	--	--	4.9	4 to 5

Passage ID	Unique Words	Repeated Words	Words per Sentence	Double Syllable Words	Single Syllable Words	Spache Level	Spache Grade	Dale-Chall Grade	Dale-Chall Level		
Grade 5											
27693	112	83	29	14	8	23	80	--	--	6	7 to 8
27694	108	75	33	12	9	30	72	--	--	5.5	5 to 6
27695	106	75	31	12	9	19	76	--	--	6.9	7 to 8
27696	110	70	40	16	7	20	85	--	--	6.4	7 to 8
27697	115	79	36	16	7	32	78	--	--	6.5	7 to 8
27698	109	82	27	14	8	19	78	--	--	5.6	5 to 6
27699	105	77	28	12	9	25	77	--	--	6.2	7 to 8
27700	115	90	25	6	18	16	89	--	--	6.4	7 to 8
27701	106	78	28	12	9	26	74	--	--	7.8	9 to 10
27702	117	78	39	15	8	19	92	--	--	5.4	5 to 6
27703	108	80	28	15	7	17	82	--	--	6.3	7 to 8
27704	111	84	27	14	8	29	71	--	--	6.7	7 to 8
27705	111	77	34	14	8	23	80	--	--	5.9	5 to 6
27706	113	89	24	10	11	28	77	--	--	6.9	7 to 8
27707	112	82	30	11	10	12	87	--	--	6	7 to 8
27708	103	77	26	11	9	17	70	--	--	5.3	5 to 6
27709	106	67	39	11	10	24	75	--	--	5.7	5 to 6

ORF Application

In traditional ORF assessment, students are given one minute to read as many words as possible in a grade-level text while a trained assessor follows along and indicates on a scoring protocol each word the student reads incorrectly (Wayman et al., 2007). If a student pauses for more than three seconds, the assessor prompts the student to continue and marks the word as read incorrectly. Student self-corrections are not marked as errors, but omissions are considered incorrect. After one minute, the assessor calculates words correct per minute (WCPM) by subtracting the number of incorrectly read words from the total number of words read.

The CBMSkills fluency assessment is a computer-based assessment (CBA). A student must sign into the CBSkills platform and select their teacher-assigned module. The student is greeted with a welcome screen, told to say ‘Yes’ when they were ready to begin (at which time a timer displaying a countdown of three seconds appeared and at 3, 2, 1, 0) and a voice states ‘Begin reading now’ as the reading passage appears. The student then reads the passage aloud. The student either reads the passage in its entirety, or the session is automatically ended after 60 seconds, and the screen indicates that the student is finished.

CBMSkills uses automatic speech recognition (ASR) software to score the reading and returns that score immediately to the teacher interface. The ASR system scores each word as read correctly, incorrectly, or not reached; marks the last word read, and records the precise time the passage was read. These data are used to automatically calculate the student’s words read correctly per minute (WCPM). In addition, CBMSkills awards badges for students’ fluency performance. Table 2 below shows the WCPM score ranges for each badge by grade level which is based on the norms from the Hasbrouck and Tindal (2017) data.

Table 2

Words Correct per Minute (WCPM) Ranges in each Grade Level of CBMSkills for Medal Attainment



Grade	Bronze	Silver	Gold
1	34 – 59	60 – 90	91+
2	72 – 99	100 – 123	124+
3	91 – 111	112 – 138	139+
4	105 – 132	133 – 150	151+
5	122 – 145	146 – 172	173+

ORF Results: Sample Student

In reviewing the results for individual students, teachers first select a group and student. In Figure 1 below, student1 has completed an ORF and is receiving a ‘participation’ medal (no color) for having taken the measure but performing at 44 WCPM which is below the minimum for grade 3 (91 WCPM). Note this performance level reflects the highest attained by the student.

Figure 1

Report Access for Review of ORF Results

Student	1: 1.1 Reading Fluency	2: 1.2 Reading Fluency	3: 1.3 Reading Fluency	4: 1.4 Reading Fluency	5: 1.5 Reading Fluency
student1	—	—	 WCPM .44	—	—
student2	—	—	—		—

Teachers can then click on that icon to view the history of the performance level and review the running record which displays in this example: correct words (42), incorrect words (5), reading time (56 seconds), and words correct per minute (44 WCPM). The teachers also can play the

recording to listen to the student reading, and if warranted, score the reading manually (see upper right corner). See Figure 2.

Figure 2

Running Record of ORF Results

The screenshot displays the 'Fluency Scoring' interface. At the top right, there is a toggle switch labeled 'manual' which is currently turned on. Below this is a reading passage with several words highlighted in red to indicate errors: 'coming', 'worrying', 'important', 'looking', 'thunder', 'each', and 'have'. Below the passage is a table with four columns: 'Correct', 'Incorrect', 'Reading Time (seconds)', and 'WCPM'. The values in the table are: Correct: 42, Incorrect: 5, Reading Time: 56, and WCPM: 44. At the bottom of the interface, there is a playback control bar showing a progress bar from 0 to 0:58, a play button, and a 'Save' button.

Correct	Incorrect	Reading Time (seconds)	WCPM
42	5	56	44

Summary and Next Steps

This technical report provides the essential design and development of an automated oral reading fluency assessment system as part of CBMSkills. Although the original research on ORF assessments was thought to be sufficiently brief (one minute) for frequent administration, the time to systematically deploy it for 25-30 students is likely prohibitive. Furthermore, other measures (letter names or oral word reading measures) are often either unavailable or present the same time constraints as ORF. In CBMSkills, use of a CBA mitigates the first issue and with 10 different early reading measures also available (also taken independently by students to mitigate the second issue), teachers can now have a complete diagnostic assessment system available for targeting specific reading skills that need remediation.

Such a system is very different than the original general outcome measurement (GOM) that guided the original research, which was designed to privilege time-series data for monitoring

progress over time. With both systems, however, teachers can diagnose specific error patterns, target their instruction to remediate them, and then monitor their interventions by using easyCBM, a complementary assessment system built with GOM in mind.

As CBMSkills is used with sufficient students, the early reading diagnostic modules are designed to reflect learning trajectories for individual students. Each of the 10 modules has a common group of 20 items that allow them to be scaled across all items in the pool. These scores can then be analyzed with ORF performance as the outcome. At that time, further technical research can be completed on the necessary developmental correlates in learning to read fluently. Furthermore, as students read successive passages, analyses can be conducted on typical error patterns, given that all data are collected into a related database. Errors made in ORF can be cross-tabulated with early reading measures to identify commonalities among the domains at each grade level.

References

- Baker, S. K., Smolkowski, K., Katz, R., Fien, H., Seeley, J. R., Kame'enui, E. J., & Beck, C. T. (2008). Reading fluency as a predictor of reading proficiency in low-performing, high-poverty schools. *School Psychology Review, 37*, 18–37.
- Dale, E., & Chall, J. (1948). A formula for predicting readability: Instructions. *Educational Research Bulletin, 27*, 37– 54.
- Deno, S. L., Mirkin, P. K., & Chiang, B. (1982). Identifying valid measures of reading. *Exceptional Children 49*(1), 36-47.
- Hasbrouck, J., & Tindal, G. (2017). An Update to Compiled ORF Norms. Technical Report #1702. *Behavioral Research and Teaching*.
- Kim, Y-S., Petscher, Y., Schatschneider, C., & Foorman, B. (2010). Does growth rate in oral reading fluency matter in predicting reading comprehension achievement? *Journal of Educational Psychology, 102*, 652-667.
- National Institute of Child Health and Human Development (2000). *Report of the National Reading Panel. Teaching children to read: an evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups* (NIH Publication No. 00-4754). Washington, DC: U.S. Government Printing Office.
- Rasinski, T.V. (2006). A brief history of reading fluency. In S.J. Samuels & A.E. Farstrup (Eds.), *What research has to say about fluency instruction* (pp. 4–23). Newark, DE: International Reading Association.
- Rasinski, T., Blachowicz, C., & Lems, K. (Eds.) (2006). *Fluency instruction. Research-based best practices*. New York: Guilford

- Samuels, S.J., & Farstrup, A.E. (Eds.). (2006). *What research has to say about fluency instruction*. Newark, DE: International Reading Association.
- Shapiro, E. S., Keller, M. A., Lutz, J. G., Santoro, L. E., & Hintze, J. M. (2006). Curriculum-based measures and performance on state assessment and standardized tests: Reading and math performance in Pennsylvania. *Journal of Psychoeducational Assessment, 24*, 19-35.
- Spache, G. (1953). A new readability formula for primary grade reading materials. *The Elementary School Journal, 53*, 410– 413.
- Speece, D. L., Case, L. P., & Molloy, D. E. (2003). Responsiveness to general education instruction as the first gate to learning disabilities identification. *Learning Disabilities Research & Practice, 18*, 147-156.
- Speece, D., & Ritchey, K. D. (2005). A longitudinal study of the development of oral reading fluency in young children at risk for reading failure. *Journal of Learning Disabilities, 38*, 387-399.
- Tindal, G. (2013). Curriculum-based measurement: A brief history of nearly everything from the 1970s to the present. *ISRN Education, 2013*.
- Wayman, M. M., Wallace, T., Wiley, H. I., Tichá, R., & Espin, C. A. (2007). Literature synthesis on curriculum-based measurement in reading. *The Journal of Special Education, 41*, 85-120.
- Wood, D. E. (2006). Modeling the relationship between oral reading fluency and performance on a statewide reading test. *Educational Assessment, 11*, 85-104.