



SoapBox



Fluency: Understanding every phoneme

December, 2021

By Agape Deng

Computational Linguist

Speech Technology Department

Introduction

By now, the pedagogical advantages of SoapBox Labs' speech technology for assessing and improving kids' reading outcomes are well-attested by users and well-supported by evidence. A controlled study in 2019 showed that SoapBox's voice-powered reading tools bring not only motivation and encouragement but also improved reading ability in young children.¹ More recently, an extensive test over a six-month period by Amplify, a leading U.S. curriculum and assessment company, showed that SoapBox's automatic assessment correlated with human scoring at 96%, a correlation comparable to that of human-to-human scoring and that exceeds market standards.² Prescient education companies like Amplify are integrating SoapBox's technology to enable large-scale, remote evaluations for good reason.

This article will focus on SoapBox's Fluency solution in detail and explain why it provides more value than traditional, observed reading assessments. This article will also introduce a brand-new Fluency feature SoapBox is launching this quarter, which education companies have affirmed will bring significant additional value to educators.

Oral reading fluency

Oral reading – or the ability to read connected text quickly, accurately, and with appropriate expression – has 30 years of evidence to back up its use as one of the most reliable and efficient indicators of student reading comprehension.³ Because of its strong evidence base, repeatability and brevity, oral reading fluency (ORF henceforth) tests are used for universal screening for early intervention across grades 1 through 8 in America.⁴ They are typically conducted three times per school year, often even more frequently for younger students, to monitor reading progress.

¹ Deng, A. (2019). *Fostering Literacy with Speech Recognition: A Pilot Study*. Available for download at <https://www.soapboxlabs.com/white-paper/>

² Popoff, E. (2021). *Can Speech Recognition Help Children Learn to Read? A primer on the applications, risks and aspirations of speech recognition technology in the classroom*. Available for download at <https://www.soapboxlabs.com/educate-whitepaper/>

³ Reschly, A., Busch, T., Betts, J., Deno, S., & Long, J. D. (2009). Curriculum-based measurement oral reading as an indicator of reading achievement: A meta analysis of the correlational evidence. *Journal of School Psychology, 47*, 427-269. doi: 10.1016/j.jsp.2009.07.001

⁴ Reschly, A., Busch, T., Betts, J., Deno, S., & Long, J. D. (2009). Curriculum-based measurement oral reading as an indicator of reading achievement: A meta analysis of the correlational evidence. *Journal of School Psychology, 47*, 427-269. doi: 10.1016/j.jsp.2009.07.001

Until very recently, ORF assessments had to be carried out by a human examiner, who sat down with each student with a timer and pencil in hand. The examiner listened as the student read a grade-level passage aloud for one minute and noted down three points at the bottom of the page:

Total words read: ____ Total Errors: ____ Total words correct: ____
--

Figure 1: Points scored by a human examiner in an ORF assessment

Scores were then compared to the national fluency norms to determine whether the student is on target. This is currently how all the most widely used measures of ORF are carried out.⁵ We will refer to them as "observed" ORFs.

Though efficient compared to many other assessments, there is still a non-trivial time cost involved in administering these tests. To put it into perspective, there are roughly 35.5 million public K - 8 students in the United States,⁶ and an ORF takes about two minutes per student. Conducted three times a year, the total time spent on administering ORFs adds up to 213,000,000 minutes, or 3,550,000 hours, or 147,916 days. This does not include time spent comparing and reporting each student's performance against national norms.

This is just one of many assessments teachers are required to carry out as a result of policies such as Every Student Succeeds, Reading First, and other state-wide and district initiatives. Coupled with the demands of actual instruction and the cry from parents that the purpose of schools is to teach, not to test, teachers are left teeter-tottering on a tightrope, trying to balance between instructional time and assessment while getting pulled from both sides. With all these competing demands, lack of support, and unprecedented levels of stress due to Covid-19, it's no wonder that one in four burnt-out teachers are considering quitting after this past year.⁷ The message is loud and clear: the education of the next generation is at stake, and if teachers are to continue, now more than ever, they need and deserve all the support they can get.

⁵ e.g., DIBELS®8th Edition, easyCBM, Gray Oral Reading Test Fifth Edition (GORT-5). In GORT-5, even fewer data points- only time and total errors- are recorded.

⁶ Hanson, M. "K-12 School Enrollment & Student Population Statistics" EducationData.org, September 19, 2021, <https://educationdata.org/k12-enrollment-statistic>

⁷ Hess, A. "I felt like I was being experimented on: 1 in 4 teachers are considering quitting after this past year" <https://www.cnbc.com/2021/06/24/1-in-4-teachers-are-considering-quitting-after-this-past-year.html>

SoapBox Fluency solution

As part of SoapBox's suite of education solutions, Fluency responds to requirements for literacy practice and assessment from Grades 1 through 8. The solution enables remote and automated assessment of a student's ORF, delivering highly valuable, rich data to teachers via client dashboard reporting systems. Teachers have anytime access to data, information, and recordings of students' reading practice and assessment. This enables teachers to focus on designing lesson plans and intervention for a diverse cohort of students, reducing time spent on data administration and focusing on areas of higher value to both teacher and student. In this way, SoapBox Fluency supports and empowers teachers. The system, built using state-of-art AI technology, has been shown to correlate highly with human assessors.⁸ Just like an observed ORF, SoapBox Fluency can return the traditional "Total Words Read," "Total Errors," and "Total Words Correct," but it can also do much more than that at no additional time cost to the educator; this is where its key value lies.⁹

In an observed ORF, whether errors are mispronunciations, substitutions, or omissions is not marked. Inserted and repeated words are also not marked, since they do not count as errors. The examiner is simply instructed to "put a slash through the error." Given the time constraints and the number of students and assessments a teacher needs to plow through, this brevity makes sense. But it sacrifices a lot of the big picture, a picture that SoapBox Fluency can help complete.

There is another diagnostic tool developed by Ken Goodman, called a miscue analysis, that does take into account the types of errors (or miscues) that a student makes.¹⁰ Unlike the ORF which is a screening tool that gives a "quick and dirty" estimate of a student's reading proficiency, miscue analyses help identify specific reading weaknesses to help teachers design appropriate interventions. The trouble is miscue analyses are even more time consuming to administer than ORFs, taking 10 to 15 minutes per student and "should be done every 6 to 8 weeks to give a sense if reading interventions are addressing the student's needs."¹¹ But 15

⁸ See footnote 2

⁹ See more on SoapBox Fluency including a demo at <https://www.soapboxlabs.com/fluency/>

¹⁰ Goodman, K. (1973). Miscues: "Windows on the reading process." In F. Gollasch (Ed.) *Language and literacy: The selected writings of Kenneth Goodman* (pp. 93–102). Vol. I. Boston: Routledge & Kegan Paul.

¹¹ Watson, S. (2019). Miscue Analysis for Diagnosing Reading Skills. Retrieved October 22, 2021, from <https://www.thoughtco.com/miscue-analysis-for-diagnosing-reading-difficulties-3111062>

minutes of individual time per child in a class of 25 would take over more than six hours of a teacher's day, every day. Because of their impracticality in the realities of a busy classroom, miscue analyses fell out of favor with many teachers. In the 1978 journal article titled "Is Miscue Analysis Practical for Teachers?" the answer was a hearty "no."¹² The time cost then was simply insurmountable. Today, with solutions like SoapBox Fluency, the answer has transformed into a resounding "yes."

SoapBox Fluency is like an automated-ORF-plus-miscue-analysis-in-one solution. It provides more insight than an observed ORF by showing the breakdown of total errors and removes the barrier with miscue analysis by providing near real-time feedback. It is well known that not only the number of errors but the types and frequencies of the errors are important in assessing reading behaviors.¹³ Knowing whether substitutions, omissions, and mispronunciations are regularly occurring can offer much more pinpointed and differentiated insight into a child's literacy journey than a blanket statement like "Total Errors" can.

In an acute example, if the majority of a child's errors are omissions with whole words or phrases often skipped, in conjunction with other signs, this may actually indicate a visual processing deficit or attention deficit rather than a reading problem.¹⁴ Similarly, if a child's errors are mostly mispronunciations, and the mispronunciations are systematic, such as regular substitutions of /th/ for /s/ as in "thing" instead of "sing," then it may be more indicative of an articulation disorder.¹⁵ Though SoapBox Fluency is by no means a diagnostic and does not claim to be one, what it can do is provide data points to aid teachers in making the most informed decisions about what instruction or intervention to provide. From just a few examples, it's easy to see that the types and frequencies of the errors matter and hold a wealth of insight that could be used to provide more individualized instruction.

¹² Hood, J. (1978). Is Miscue Analysis Practical for Teachers? *The Reading Teacher*, 32(3), 260–266. <http://www.jstor.org/stable/20194750>

¹³ Brown, J. Goodman, K. & Marek, A. (Eds.) (1996) *Studies in miscue analysis: An annotated bibliography*. Newark, DE: International Reading Association.

¹⁴ Kibby M., Dyer S., Vadnais S., Jagger A., Casher G., Stacy M. (2015). Visual processing in reading disorders and attention-deficit/hyperactivity disorder and its contribution to basic reading ability. *Frontiers in Psychology*, 6, p1635. doi: 10.3389/fpsyg.2015.01635

¹⁵ American Speech-Language-Hearing Association (n.d.) *Speech Sound Disorders: Articulation and Phonology*. Retrieved October, 22, 2021, from www.asha.org/Practice-Portal/Clinical-Topics/Articulation-and-Phonology/

New feature: Phoneme-level scores

In SoapBox Fluency, each word in a prompt is returned as "correct," "substitution," "deletion," "insertion," or "repetition," and accompanied with a score to indicate the level of confidence with which the model made the prediction. The more clearly a child pronounces a word, the more confident the model is in its prediction, and the higher the score. Previously, this score was only available at the word level. So for example, if a child misreads "tutter" instead of "butter," the system would return a lowered word score (e.g., 55%), but previously, where exactly within the word the error occurred is not indicated.

Now, SoapBox Fluency is offering phoneme-level scores in addition to word-level scores to provide educators with finer granularity as to where errors occur word-internally. Though this feature is already available in some of SoapBox's other products (powering voice-enabled assessments of letter sounds and shorter sentences), it is now available for reading passages of any length in Fluency. Thus, in addition to the 55% word score, phoneme scores for each constituent part of the word "butter" (/b/, /ah/, /t/, /er/) are provided, so now the teacher can see that the /b/ phoneme received a low score since the child had substituted with a /t/. Identifying the mistake at the phoneme level allows the educator to zero in on reading weaknesses and address them early on.

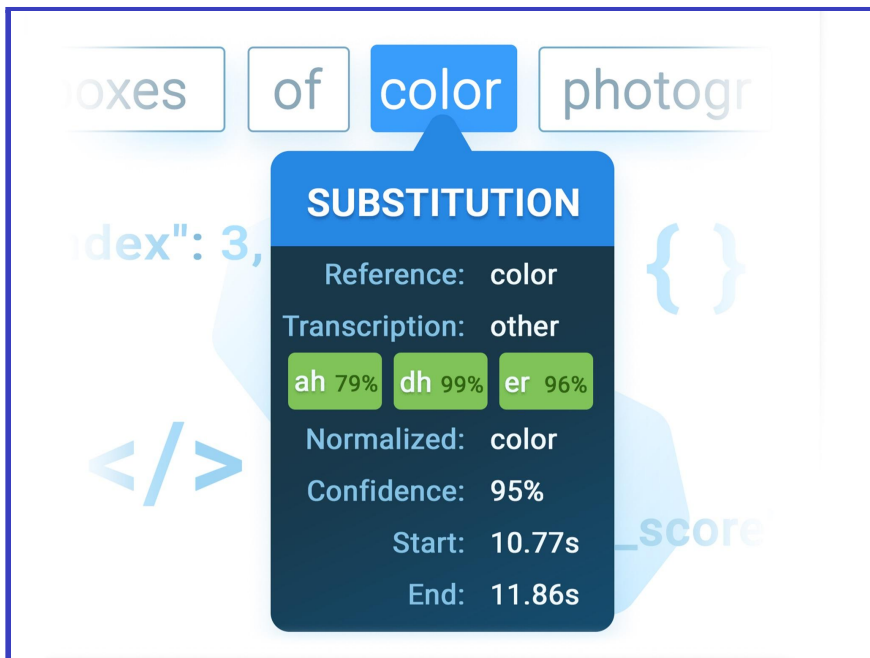


Figure 2: Phoneme breakdown of a transcribed word "other"

Let us consider another example: if the reference sentence is “The color is blue” but the child utters, “The other is blue,” the engine will return “other” as a substitution with the reference word “color.” From this, the educator can immediately see that the child has confidently, albeit wrongly, made a substitution and can address this with the child. An example JSON of this substitution output from the system would look like the below, where the reference_word is “color,” but the transcription_word is “other.” Note that the phoneme scores provided for the phonemes /ah/, /dh/, and /er/ are constituents of the substituted word “other.”¹⁶

```
{
  "normalised_word": "color",
  "reference_word": "color",
},
"transcription_details": {
  "confidence": 95.325660705566,
  "phones": [
    {
      "phone": "ah",
      "score": 79.470123291016,
    },
    {
      "phone": "dh",
      "score": 99.071838378906,
    },
    {
      "phone": "er",
      "score": 95.694442749023,
    },
  ]
},
"transcription_word": "other",
}
```

Figure 3: Example JSON response of the reference word “color” and transcribed word “other”

Educators will also be able to see the phoneme scores when a child inserts additional words or repeats words. Repetitions and insertions are not marked in traditional ORFs because they do not count as errors; however, that does not mean they hold no value. While it is a common and positive sign when a student goes back to repeat words or phrases as a self-correction, excessive repetitions beyond the norm may indicate an underlying reading disorder. Though uncommon, teachers have reported cases where students failed to achieve reading fluency goals due to

¹⁶ This is a simplified JSON to highlight the relevant information; in reality, richer data than this is returned, including start and end times, pause durations, and more.

seemingly involuntary repetitions of words and phrases.¹⁷ More likely, it may mean that the text is too difficult, in which case the student should be given a lower levelled text until he or she is able to progress. On the other hand, insertions may mean that the reader is reading too fast, and if occurring too frequently, can detract from the comprehension of the text.¹⁸ In this case, reading speed may need to be addressed by the teacher.

Instead of ignoring repetitions and insertions and blanketing substitutions, mispronunciations, and omissions as "errors," as is done in observed ORFs, SoapBox Fluency returns these data points to clients who may then choose to analyze, discard, or use them as they see fit.

A key philosophy at SoapBox Labs is to build solutions that empower educators to make decisions, not make decisions for them. SoapBox Labs supplies educators with insights so that they can make best-practice, data-informed decisions about their lessons.

For more information about the SoapBox Fluency solution, visit our [website](#) or contact us at hello@soapboxlabs.com.

¹⁷ A to Z Teacher Stuff. (2011). Repeating everything in fluency test. Retrieved October, 22, 2021, from <https://forums.atozteacherstuff.com/index.php?threads/repeating-everything-in-fluency-test.153967/>

¹⁸ Watson, S. (2019). Miscue Analysis for Diagnosing Reading Skills. Retrieved October, 22, 2021, from <https://www.thoughtco.com/miscue-analysis-for-diagnosing-reading-difficulties-3111062>

About the author

Named finalist two years running for Rising Star of the Year by Women in Tech Excellence and Diversity in Tech, Agape Deng works as a computational linguist at SoapBox Labs, where she looks after data, natural language processing, and evaluation improvement efforts. Agape has served as a reading teacher and speech pathologist aide across Texas and California and draws from her classroom and clinical experience to inform her work. She holds a B.A. in Applied Linguistics from UCLA and a Masters in Speech Processing from Trinity College.

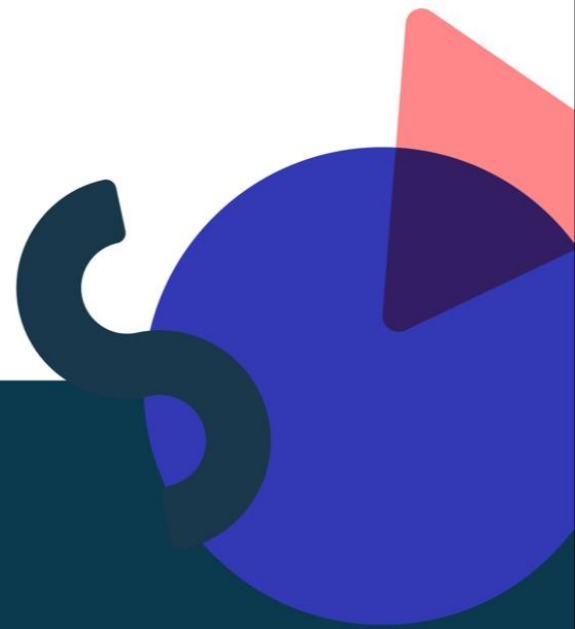
About SoapBox Labs

SoapBox Labs makes kids' unique voices heard in the digital world. Our speech technology is proprietary and built from the ground up to empower 2 to 12 year old kids of every accent and stage of development to have joyful digital experiences using their voices.

SoapBox voice-enables learning and play experiences for education, games, and entertainment companies. We are a privacy-first company and will never reuse or sell our partners' voice data for marketing or advertising purposes.



SoapBox



For more, visit **our website**
or send us an email at
hello@soapboxlabs.com