



USAID
FROM THE AMERICAN PEOPLE

Harnessing AI Speech Recognition Technology for Educational Reading Assessments amid the COVID-19 Pandemic in the Philippines

Presenters:

Dr. Carmen Strigel, Sr. Director, Education Technology, RTI International, cstrigel@rti.org

Anthony Udeh, Sr. EdTech Advisor, RTI International, audeh@rti.org

Mitch Rakusin, Sr. Research Education Analyst, RTI International mrakusin@rti.org

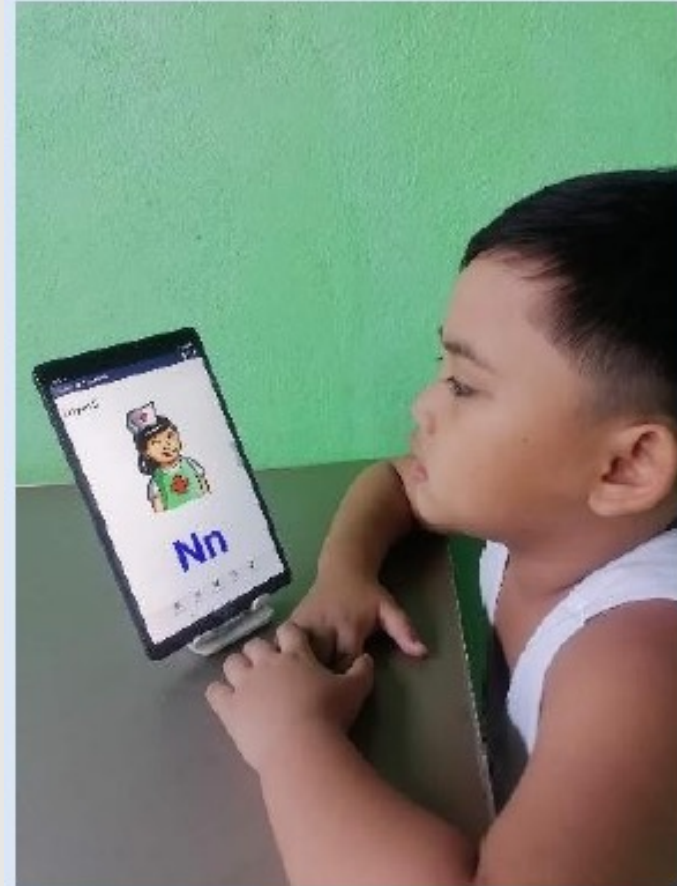
Introduction

COVID-19 pandemic has brought unprecedented challenges to educational assessment, particularly in low- and middle-income environments.

The Computer-based Reading Assessment (CoBRA) implemented with the Philippines Department of Education (DepEd) under the USAID-supported All Children Reading Asia (ACR Asia) initiative.

CoBRA was a 'Proof of Concept' design that explored the parameters of a self-administered and automatically scored version of the Philippines Informal Reading Inventory (Phil-IRI) in English and Filipino for Grades 4-6.

CoBRA built on existing platforms used by DepEd (Moodle) adding a customized AI plug-in (Poodll).



Potential Benefits of AI Speech Recognition Technology

- **Efficiency:** Reading fluency and comprehension can be efficiently measured (to a degree) with limited need for human inputs.
- **Instant feedback:** The technology provides quick guidance by assessing intonation, pronunciation, and tempo while students speak.
- **Personalized learning:** Students receive tailored feedback to enhance their reading abilities, even in the absence of in-person teacher interactions.
- **Scalability:** AI-backed evaluations enable educators to collect extensive data on reading patterns and address common issues among students.
- **Adaptability:** The platform allows teachers to make corrections to mis-evaluated words and provide the AI model with the opportunity to learn from the teachers' correction. Teachers also can provide alternative words that sound similar in their specific context in support of mother-tongue based multilingual education.

Research Questions

- How feasible are the logistics of a computer-based reading assessment in the context of the Philippines for students in the two grades?
- How well does the speech recognition technology manage the variety of English spoken by the participating children in the respective grades?
- How reliable are students' results on the computer-based reading assessment, especially the speech recognition technology-supported tasks, compared to the assessor-administered approach of the assessment?
- What are considerations and recommendations regarding the future development of similar technology for other languages relevant to the Philippines?



CoBRA User Testing



- In April 2022, the team conducted user testing with 10 learners from Grades 4-6.
- Teachers were also observed accessing the reports and reviewing and re-grading the learners' scores.
- This experience provided valuable feedback to the development team, which focused on how to improve the user interface, key functionality and navigation of the system to make it more user friendly.

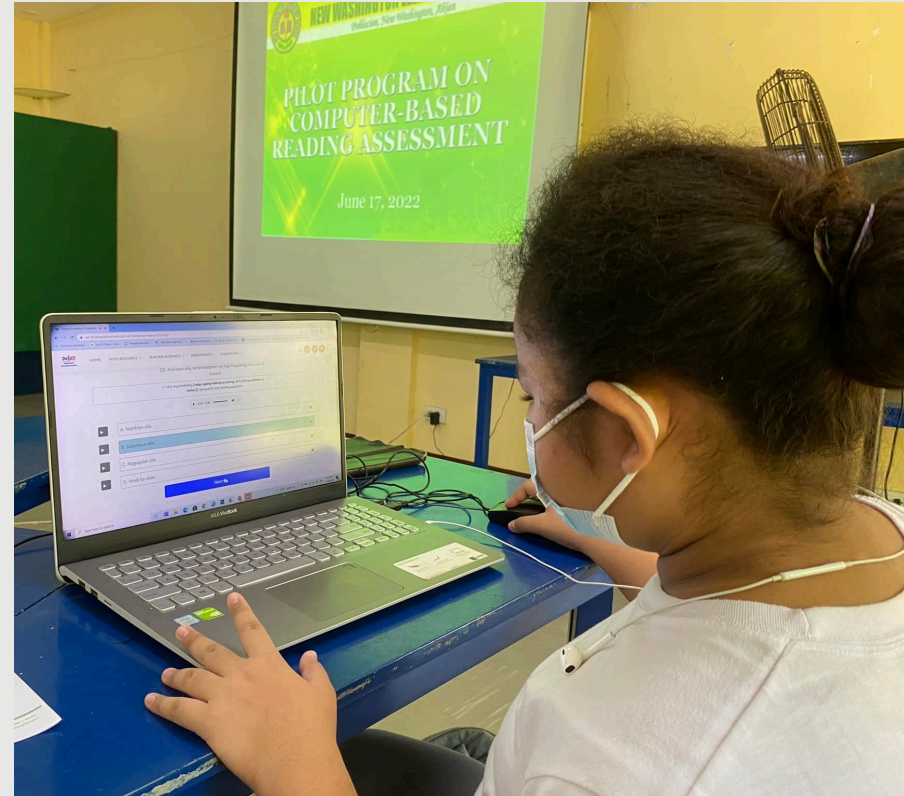
CoBRA Pilot Testing

- In June 2022, the ACR team oriented 67 school and District ICT Coordinators to help prepare them to support the pilot.
- The team then introduced the CB Reading Assessment to 145 teachers and educators across the 42 schools.
- Following the trainings, smaller check-ins were held with teacher groups to ensure they understood the tasks, requirements, and features to facilitate their students to take the test.



CoBRA Pilot Testing – cont'd

- 1,063 learners took the Grade 4-6 tests in English
- 933 learners took the G4-6 tests in Filipino
- Teachers re-graded 5 of their students, leaving an additional 5 students per grade not re-graded.
- Three independent assessors reviewed the AI-scores of 346 English assessments and 345 Filipino assessments which teachers had not already re-graded for variances between AI-generated scores and human-generated scores.



CoBRA – Passage Reading Student View

GRADE 4: SELF-ADMINISTERED READING ASSES...

Home / MyEResource / CBRAG4A / Grade 4 SET ... / English Oral ...

English Oral Reading Test "GET UP, JACKY!"

In this activity, you will read a passage out loud. You may be required to use a headset/microphone. To start click the record button on the audio recorder icon to your left to show the reading passage.
Start reading the passage aloud as clearly as you can. Remember to end the recording as soon as the learner has finished reading the passage.

"Ring! Ring!" rang the clock.

But Jacky did not get up.

"Wake up, Jacky! Time for school," yelled Mom.

And yet Jacky did not get up.

"Beep! Beep!" honked the horn of the bus.

Jacky still laid snug on the bed.

Suddenly, a rooster crowed out loud

and sat on the window sill.

Jacky got up and said with cheer,

"I will get up now. I will!"



CoBRA - Re-Grading View

GRADE 4: SELF-ADMINISTERED READING ASSES...

Home / EResources / Reading Asse... / Grade 4: Co... / CBRAG4C / Grade 4 SET ... / English Oral ...

English Oral Reading Test. "ANANSI'S WEB"

View

Grading

Reports

Model Audio

Administrator

Showing attempt for: Student Grade 4C Amado T Reyes Elementary School

0:50 / 0:50

WPM	Acc(%)	Grade(%)	Mistakes
75	77	64	18

Anansi

 was

tired

of

 her web. So one day, she

said

"I

will

go

live

with

 the ant."

Now, the ant lived in

a

 small hill. Once in the hill Anansi cried, "This place is too dark! I will go live with

the

 bees."

When she got

to

 the

beehive

,

Anansi

 cried,

"This

 place

is

 too hot and sticky! I will go live

with

 the

beetle

."

But on her way to beetle's home she saw her web. "Maybe a

web

 is the best place after all."

Grading Mode

Spot Check Mode

Transcript Check Mode

Clear all markers

Cancel

Save changes

Results from the Concurrent Validity and Analysis

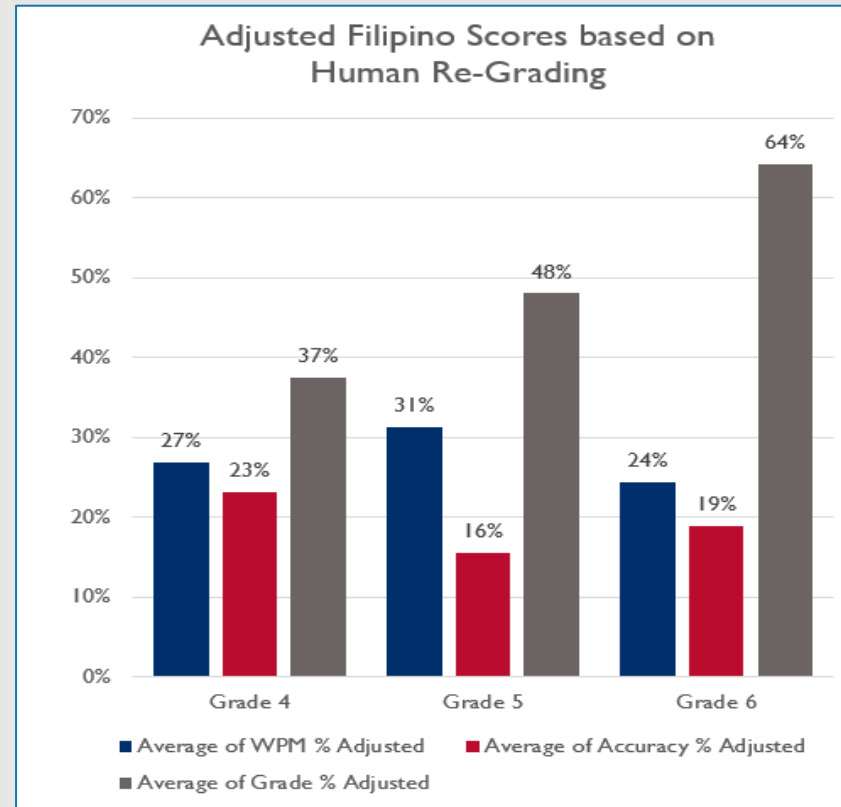
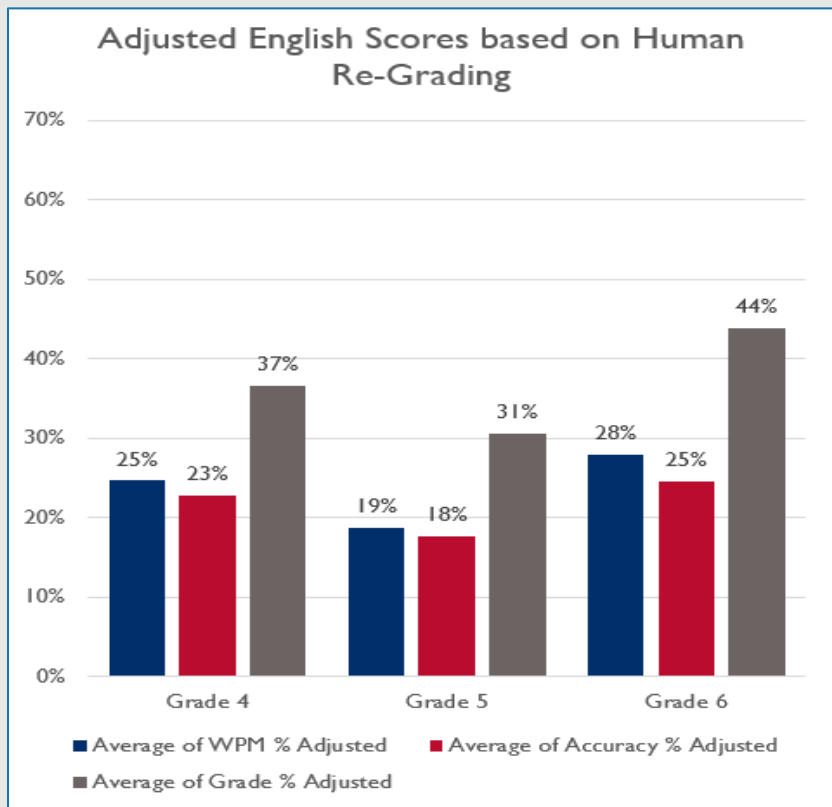
English Words Per Minute Variance between AI and Human Scoring			
Grade	Average of Original AI WPM	Average of Regraded Human WPM	WPM Difference (Human - AI)
Grade 4	68.7	87.2	18.5
Grade 5	86.8	105.3	18.5
Grade 6	74.1	100.8	26.7
Overall	76.2	98.0	21.8

Filipino Words Per Minute Variance between AI and Human Scoring			
Grade level	Average of Original AI WPM	Average of Regraded Human WPM	WPM Difference (Human - AI)
Grade 4	53.0	71.5	18.5
Grade 5	63.6	105.4	41.8
Grade 6	63.1	81.5	18.3
Grand Total	60.2	85.3	25.1

- Most teachers and students were able to cope with the technology with little disruption due to connectivity or technological issues.
- The concurrent validity results indicate that the **AI scored on average 25.1 fewer correct words per minute in Filipino and 21.8 fewer correct words per minute in English** than those same tests re-graded by human assessors*

**data from the independent assessors, not the teachers*

Results from the Concurrent Validity and Analysis - cont'd



- **Overall, the AI incorrectly flagged or skipped 21% of all English text and 60% of all Filipino text read by students.** The higher rate of error for the Filipino text was due more to the incorrect skipping of text—that is the AI did not recognize the student reading blocks of passage—more so than the incorrectly recording of mispronounced words.

Results from the Concurrent Validity and Analysis - cont'd

English Passage words most frequently scored incorrect by AI

English Word	Total count	No. Errors	Error rate
falls	41	33	80%
Mmm	34	27	79%
I'll	34	23	68%
hour	147	94	64%
meteoroid	205	128	62%
300	21	13	62%
600	21	13	62%
waves	21	13	62%
meteor	82	50	61%
Please	111	67	60%

Filipino Passage words most frequently scored incorrect by AI

Filipino Word	Total count	No. Errors	Error rate
nagsaka	201	60	30%
drowning	201	50	25%
taga-Lucban	133	27	20%
Ana	130	25	19%
isandaang	133	21	16%
eskultura	201	31	15%
patubigan	201	30	15%
nakapagpasigla	171	24	14%
ring	301	40	13%
di	130	17	13%

- The lower rate of error of Filipino words compared to English was offset by the higher rate of incorrectly skipped words (as noted earlier).
- **The AI performed much better in recognizing the Filipino spoken word than the English spoken word.** This is most likely due to the ways children pronounced the words.

Post Pilot Survey Results – Teachers

- 90% of teachers noted that it was easy for students to take test and teachers to review and regrade results
- 82% indicated that their Internet and equipment was stable enough to conduct the assessment online.
- 90% agreed or strongly agreed that the automated scoring was generally accurate, and the comments were overwhelmingly positive.
- According to teachers, parents and school directors showed high interest in expanding the CoBRA pilot to all the student in the school for the pretest and posttest in the coming school year



Post Pilot Survey Results - Students

- Participating students had high phone ownership (50%); mobile phone access (90%); and functional usage of mobile phones and computers (50%)
- Most students liked the instant feedback on several of the reading tasks
- Students wanted more passages on trendy topics and more elements of gamification.



“It saved weeks of calculating and counting words, and at same time it saved the school from printing papers” (Teacher)

Pilot - Conclusions

- The AI-supported reading assessment showed promise in offering teachers a more efficient method for assessing students' reading and listening comprehension skills;
- The AI technology used was not (yet) accurate or reliable enough to serve as a reliable, independent means for measuring students' oral reading fluency skills, and required thorough review and re-grading.
- For scaling up the pilot, the licensing costs were deemed unsustainable. Developing new speech models or connecting to other open-source language models might suit a larger context.
- Additional research is needed to better understand how AI can benefit education especially for large-scale assessment.

Next Steps

- Work with other AI speech recognition service providers to expand the use and scope of the language models
- Explore offline integration on mobile devices
- Design and conduct qualitative research how the re-grading the AI's scoring may affect teachers' intentions and enactment of instructional changes in their classroom

Thank You!

Carmen Strigel | cstrigel@rti.org

Anthony Udeh | audeh@rti.org

Mitch Rakusin | mrakusin@rti.org

Learn more about RTI's work in
International Education:



www.rti.org/idg_education

<https://shared.rti.org/content/computer-based-reading-assessment-pilot-report>

