Technology-Enhanced SORT for Students (TESS): A Computerized Oral Reading Assessment System for incoming Grade 7 students

A Special Problem Proposal
Presented to
the Faculty of the Division of Physical Sciences and Mathematics
College of Arts and Sciences
University of the Philippines Visayas
Miag-ao, Iloilo

In Partial Fulfillment
of the Requirements for the Degree of
Bachelor of Science in Computer Science by

DIESTRO, Kimberly
SALIBIO, Alexa Therese

Francis DIMZON
Adviser

April 28, 2025

Abstract

Iloilo City National High School (ICNHS) is one of the few schools that has a dedicated remedial reading program. Students with reading difficulties face more challenges in their education than their peers, and so it is important for schools to intervene and provide support as early as possible. At the start of every school year at ICNHS, incoming Grade 7 students have their reading levels assessed in order to determine if they require remediation. Each student takes the Slosson Oral Reading Test (SORT); which is a test used to determine the student's reading level and word recognition skills and is manually conducted one on one with a teacher.

Every year, there are over 1200 Grade 7 students enrolled in ICNHS. However, the testing population was narrowed down to 300 by choosing to only test students with a grade of 83

The Computerized SORT called Technology-Enhanced SORT for Students (TESS) was developed in order to help and assist the Reading Remediation Program streamline their assessment process. TESS used Vosk for it's voice recognition software and Jellyfish to further improve Vosk's accuracy in transcribing non-native English speakers. TESS was tested on a testing population of 17 students and was evaluated on the following criteria: speed, accuracy, and ease of use.

It was found that in regards to speed, the students rated TESS at a 4 or "good" in the software evaluation form; additionally, TESS is able to assess a total of 17 students in under 3 hours with 2 students taking the test simultaneously. As for accuracy, using the Wilcoxon-Signed Rank Test we can determine that the difference between the results of a manual SORT and TESS is statistically significant. TESS consistently scores students lower than the teacher's do on a manual SORT and so TESS needs to be adjusted in order to be more reliable and accurate. However during the focus group discussions, teachers and students alike expressed excitement and strong preference to using TESS over manual SORT. Although some improvements and feature recommendations were provided in order to improve quality of life, accessibility, and efficacy, it can be concluded that TESS was considered easy to use. With some adjustments to accuracy and accessibility, TESS can prove to be a robust and reliable computerized SORT.

Contents

Chapter 1

Introduction

1.1 Overview

Iloilo City National High School (ICNHS) is one of the few schools that has a dedicated remedial reading program. It is a program that aims to address the learning gaps of students with reading difficulties. Students with such difficulties are frequently referred to as 'frustrated' or 'struggling' readers and is defined by a student's reading level falling short of their grade level. Estuya defines a frustrated reader as someone who "struggles with reading despite appropriate instruction" (Estuya). The reading level of a student is determined through testing, which the program does at the beginning of the year in order to identify students in need of remediation. Once identified, they are grouped together in a section or two so that remedial reading classes can be easily scheduled in hopes of improving their reading levels. Activities are also arranged to encourage students to persevere despite their circumstances.

Incoming students are assessed using the Slosson Oral Reading Test (SORT) to determine the grade level of their word recognition skills. SORT evaluates a student's word recognition level and grades them into 10 grade levels; wherein the first level P means primary, and the last grade is high school. Grades one to eight serve as the in-between reading levels. Students are asked to read from a set of 10 word lists that correspond to each of the reading levels starting from the lowest reading level and moving upward. A qualified teacher administers the test on a one-on-one and face-to-face basis. (Slosson Educational Publications, Inc.)

The test concludes when the student finishes all levels or incorrectly reads 10 words consecutively. Their reading level is then determined by the total number

of words read correctly divided by 20. Those students whose reading level is determined to be below level 2 are considered in need of intervention and are included in the program for remediation.

Every year, there are over 1200 Grade 7 students enrolled in ICNHS. However, they have narrowed down the testing population to incoming students who have an average grade of below 80% (over the pandemic they have increased it to 83% and below). Still, every year there are at least 300 students requiring assessment. In conducting this test, they identify over 60 students with reading difficulties every year. With this volume of students to be assessed, the process requires multiple teachers to volunteer their time. This is all on top of their other enrollment responsibilities. Due to this, only about 6 teachers volunteer every year in a span of a couple of days.

Students with reading difficulties face more challenges in their education than their peers. Thus, it is important for the school to recognize these issues early on and work with the students and parents to address these learning gaps.

1.2 Problem Statement

In order to better meet the testing requirements for the incoming Grade 7 population, the Committee of ICNHS' Reading Remediation Program is looking for technology that can help them streamline their testing process and improve the rate at which they can assess the population. The development of a Computerized SORT that meets their specifications would greatly help and improve their assessment process.

1.2.1 General Objective

To develop the Technology Enhanced SORT for Students (TESS) which can conduct the SORT and meets the testing requirements and specifications of ICNHS Remedial Reading Program. The program should also allow for the teachers to administer the test to a group of students.

1.2.2 Specific Objective

The specific objectives of this research is as follows:

1. To convert SORT into a working computerized assessment.
2. To develop an easy-to-use program that can be used by teachers as-is.
3. To develop a program that can be used offline and in a school computer lab.
4. To identify voice recognition software that can reliably transcribe words spoken by a non-english speaker in a school environment.

1.3 Scope and Limitations of the Research

TESS is predominantly a computerized testing tool and will only generate the result of the Computerized SORT for use by the program's faculty.

TESS will only be launched and tested within Iloilo CIty National High School (ICNHS). The test group will consist of a select group of students with differing reading levels from the school's Grade 7 population.

Chapter 2

Review of Related Literature

Reading is one of the basic and most essential literacy skills required to be successful in today's world of technology and information. Reading proficiency can affect not only a student's ability to perform in school, but also their self esteem and their attitude towards schooling. This in turn can severely impact the student's future prospects and social mobility. Therefore it is important to ensure that a student is able to achieve the reading milestones / achievements set for their grade.

DepEd's Program "Every Child a Reader", which was started S.Y. 2002-2003, mandates that every pupil shall not proceed to the next grade level unless they have achieved mastery over the basic literacy skills of their grade level; however, despite this and the many efforts of DepEd to encourage and develop the reading skills of the nation's learners, there are still slow and even non-readers in incoming 7th grade students. (Sancada (2022)). Additionally in the 2018 Programme for International Student Assessment (PISA), Philippines ranked last in the world in terms of reading and improved slightly in the 2022 PISA now ranking 6th lowest in the world instead. (PISA, 2018; PISA, 2022; EDCOMM II year one report, 2024)

There is much research and evidence that supports that early intervention, such as during elementary school, is effective at helping slow and frustrated readers improve, catch up, and even excel beyond the literacy milestones of their grade. (Lee et al. (2012)) Failure to intervene early can lead to the student developing poor reading skill which in turn leads to poor academic performance, creating consecutive experiences of failure that can further demotivate students to learn. (Fosudo, 2010 as cited in Sancada, 2022)

However, even if early intervention was not possible, remediation during their adolescence or in later grades, such as high school, is possible and effective (as an example, Rafanan & Raymundo, 2024; Lovett et al, 2021). As part of DepEd's efforts to improve the literacy of pupils and students in the Philippines, many schools have established Remedial Reading Programs in an effort to identify students in need of help. A study aimed to test the efficacy of remedial reading programs in helping struggling readers of Grade 7 students of Bolo Norte High School, found that intervention had a substantial impact on student's reading comprehension. Students were able to improve after a 5-month period of remedial classes, highlighting the positive impact of educational intervention. (Abergos et al. (2024))

Despite the efficacy and the potential of these reading programs, there can still be a lack of support for them. A study looking into the challenges and effective practices of Remedial Reading Programs in Iloilo found that one problem that these programs face is a lack of support from the Parents, School, and even the Local Government Units. A lack of support could mean not enough attention from the parents to support their children, a lack of funds to purchase books and equipment, a lack of interest from teachers to participate and proctor in the program and even a lack of interest from the school to support the program due to it not being a flagship program of the school, which can result in a lack of classroom space and no available schedules to conduct the program. It follows that schools with successful and more effective programs are ones that have the support needed to have proper implementation. (Sancada (2022))

The process of assessing students is the first step to providing support to students in need of remedial reading classes. If these schools cannot efficiently and effectively identify students in need, then the reading programs will not be able to perform to their full potential. Currently, reading assessments are conducted one-on-one with teachers' having to proctor and assess the student's results manually and all on their own.

For example, every year, as mentioned above there are over 1200 Grade 7 students enrolled in Iloilo City National High School. However, the school doesn't have the manpower to administer the test to every incoming student. Even after narrowing down potential students in need, they identify over 60 students with reading difficulties every year. With this volume of students to be assessed, the process requires multiple teachers to volunteer their time. This is all on top of their other enrollment responsibilities. Due to this, only about 6 teachers volunteer every year in a span of a couple of days.

This lack of available staff and resources proves to be a significant bottleneck in being able to identify students in need. One such solution created to streamline the

process of student assessment is to make use of current and emerging technology to create computerized assessments. In a 2023 study by Auphan et. al., they tested the efficacy of using computer-based assessments (CBA) for reading and found that such tools provide a low-cost and easy-to-use tool for professionals to assess reading skills. They even further suggested that CBA's have many advantages and developing tests to make use of the technological advantages of CBA can further improve the efficacy of a reading assessment.(Auphan et al. (2020))

In 2022, All Children Reading (ACR) Philippines, in coordination with the United States Agency for International Development (USAID) and DepEd, developed a Computer-Based Reading Assessment (CoBRA) Pilot and tested the program in 42 schools in the Philippines. CoBRA would allow students to independently take the Philippine Informal Reading Inventory (Phil-IRI) – a standardized reading assessment used as the primary means to determine the reading levels of students based on set benchmark scores. CoBRA would also provide them with instant feedback and assessment through an automated grading system and also used voice recognition software to facilitate the test – mainly Poodll and also Google Speech Recognition for the Filipino language.

The pilot itself was conducted on school premises and made use of the school's available equipment. Ultimately, the pilot was successful with positive reception from students and teachers alike. A teacher was quoted to be saying "it saved them weeks of calculating and counting words, and at same time saved the school from printing papers." CoBRA offered a more efficient and streamlined way to assess students – although its automated grading system and the voice recognition software used was not accurate and still required significant human intervention to be able to produce accurate results.

A planned second pilot for CoBRA would not proceed as planned for a few reasons. (1) The sponsoring unit with DepED ICTS was dissolved during a transition in leadership. (2) Licensing costs for the software escalated to what were deemed unsustainable levels.

One can learn a few things from the CoBRA pilot. (1) Secure an accurate voice recognition software that can operate in a school setting as well as detect accented english. (2) The program itself must be able to stand on its own or at least be developed using software that schools or the reading program can reasonably sustain.

In regards to voice recognition software, there are a few that come to mind. VOSK, Whisper, and Wav2Vec. All are open source voice recognition software that can be used to develop a computerized reading assessment system.

Chapter 3

Research Methodology

This chapter outlines the methodology and systems involved in the development of TESS as well as the process of testing and evaluating TESS. With these methodologies, the researchers should be able to develop a Computerized SORT and evaluate it in a way that its efficacy along the three main metrics of: speed, accuracy, and ease of use can be determined.

3.1 The Slosson Oral Reading Test

The Slosson Oral Reading Test (SORT) is a quick screening test used to determine a student's reading level. The SORT is composed of a reading list of 200 words grouped into 10 groups of 20. Each group is ordered according to 10 reading levels (P or primary, Grade 1 to 8, and High School). The word list was compiled from standardized school readers, and the reading level represents the median school achievement.

The test is administered one-on-one between the student and the teacher. The student is provided a copy of the word list and is instructed to read down each list loudly and clearly. A student is allowed to skip a word if they do not know how to read it. As the student reads, the teacher is tallying down the number of words read correctly and incorrectly. The goal of the test is to find the starting list, stopping list, as well as the raw score of the student.

The starting list is the reading level list that the student can read all words correctly, while the stopping list is the list the student will get all 20 of the words incorrectly. Lastly, the raw score is the total number of correctly-read

ILOILO CITY NATIONAL HIGH SCHOOL
Molo, Iloilo City
ENGLISH DEPARTMENT
SLOSSON ORAL READING TEST (SORT)

Reading Level: _____
School Grade level: ____

Name: _____ Age: _____ Date: _____
School: _____ Examiner: _____

LIST P (20)	LIST 1 (40)	LIST 2 (60)	LIST 3 (80)	LIST 4 (100)	LIST 5 (120)
1. see	1. with	1. game	1. safe	1. harness	1. cushion
2. look	2. friends	2. hide	2. against	2. price	2. generally
3. mother	3. came	3. grass	3. smash	3. flakes	3. extended
4. little	4. horse	4. across	4. reward	4. silence	4. custom
5. here	5. ride	5. around	5. evening	5. develop	5. tailor
6. can	6. under	6. breakfast	6. stream	6. promptly	6. haze
7. want	7. was	7. field	7. empty	7. serious	7. gracious
8. come	8. what	8. large	8. stone	8. courage	8. dignity
9. one	9. bump	9. better	9. grove	9. forehead	9. terrace
10. baby	10. live	10. suddenly	10. desire	10. distant	10. applause
11. three	11. very	11. happen	11. ocean	11. anger	11. jungle
12. run	12. puppy	12. farmer	12. bench	12. vacant	12. fragrant
13. jump	13. dark	13. river	13. damp	13. appearance	13. interfere
14. down	14. first	14. lunch	14. timid	14. speechless	14. marriage
15. is	15. wish	15. sheep	15. perform	15. region	15. profitable
16. up	16. basket	16. hope	16. destroy	16. slumber	16. define
17. make	17. food	17. forest	17. delicious	17. future	17. obedient
18. ball	18. road	18. stars	18. hunger	18. claimed	18. ambition
19. help	19. hill	19. heavy	19. excuse	19. common	19. presence
20. play	20. along	20. station	20. understood	20. dainty	20. merchant

_____ _____ _____ _____ _____ _____

LIST 6 (140)	LIST 7 (160)	LIST 8 (180)	HIGH SCHOOL (200)
1. installed	1. administer	1. prairies	1. traverse
2. importance	2. tremor	2. evident	2. affable
3. medicine	3. environment	3. nucleus	3. compressible
4. rebellion	4. counterfeit	4. antique	4. excruciating
5. infected	5. crisis	5. twilight	5. pandemonium
6. responsible	6. industrious	6. memorandum	6. scrupulous
7. liquid	7. approximate	7. whimsical	7. primordial
8. tremendous	8. society	8. proportional	8. chastisement
9. customary	9. architecture	9. intangible	9. sojourn
10. malicious	10. malignant	10. formulated	10. panorama
11. spectacular	11. pensive	11. articulate	11. facsimile
12. inventory	12. standardize	12. deprecate	12. auspicious
13. yearning	13. exhausted	13. remarkably	13. contraband
14. imaginary	14. reminiscence	14. contrasting	14. envisage
15. consequently	15. intricate	15. irrelevance	15. futility
16. excellence	16. contemporary	16. supplement	16. enamored
17. dungeon	17. attentively	17. inducement	17. gustatory
18. detained	18. compassionate	18. nonchalant	18. decipher
19. abundant	19. complexion	19. exuberant	19. inadequacy
20. compliments	20. continuously	20. grotesque	20. simultaneous

_____ _____ _____ _____

RAW SCORE:
List P _____
List 1 _____
List 2 _____
List 3 _____
List 4 _____
List 5 _____
List 6 _____
List 7 _____
List 8 _____
List H.S. _____

Figure 3.1: ICNHS SORT Test Paper.

words throughout the whole list. This raw score is divided by 20 to then get the student's reading level.

3.1.1 Reading Level

In addition to the reading level included in SORT, the reading level used in the Philippine Informal Reading Inventory (PHIL-IRI) was used to identify reading level groups for the testing of TESS.

The PHIL-IRI categorizes reading level as the following:

1. Independent - Readers are independent and can read with almost perfect oral reading and comprehension
2. Instructional - Readers in this level benefit the most from instruction and direction from the teacher.
3. Frustrated - Readers in this level find reading material so difficult that they cannot interact with the material.

(Slosson Educational Publications, Inc.)

3.2 Description and Functionality of TESS

Technology Enchanced SORT for Students (TESS) is a computerized reading assessment program that allows students to take the SORT with minimal to no assistance from a teacher. The following steps described were specifically requested and specified by the faculty-in-charge of ICNHS' reading remediation program.

In this computerized SORT, students read out loud from a randomized list of words flashed to them, word-by-word on the screen. This list of words is predetermined and is categorized by reading level. If the student was deemed correct, the program will proceed with the next word in the list. If the student is incorrect, then the program would remain at the current word. The student is allowed to attempt the word three times, with a timer of 10 seconds per attempt.

Additionally, the faculty wanted to decrease the amount of stress and frustration the examinee will experience during the assessment, so TESS is included with a built-in timer for each item in the test. Once eleven (11) seconds have passed an item is skipped, and once ten words in a row have been skipped, the test ends and the examinee's reading level is calculated.

A progress bar is visible in the bottom, so that students are able to track their progress through the assessment.

The program also evaluates the students and provides the teacher with an excel sheet for each student containing their results, which includes an itemized list of which words in the SORT they got right and wrong.

The test portion of the program is voice activated, and the examinee can make use of certain keywords to navigate through the test. Such as START to start the test; SKIP to skip a word; and STOP to end the test prematurely.

TESS is able to run on an offline computer with a headset.

3.2.1 Voice Functionality and Collected Voice Data

TESS' voice functionality is the most important part of the program. The speech-to-text program included should be able to function in a school environment and also be able to accurately assess accented English.

Student's voice will be recorded by the program, assessed, and then discarded. The program does not need to store data of any examinee's voice to function.

Other Collected Data

TESS takes note of the following information:

- Student Name
- Student's School
- Proctor's Name

This information is used to generate the excel report after the test is finished. TESS does not save this data after the completion of the test.

Recommended Equipment

For use, the recommended headset is Jabra Evolve 20. This headset was chosen for its affordability and efficacy.

3.3 Developing TESS

3.3.1 The Voice Recognition Software

TESS is developed using Python and a voice recognition software. The voice recognition software is responsible for transcribing the student's reading of SORT's word list. If the transcription matches the word being read, then that item is marked as 'PASS' and otherwise as 'FAIL'.

The three main choices for our voice recognition software are the following:

- Vosk,
- Wav2Vec, and
- Whisper

These software were chosen because all three are open source software and do not require any payment to keep it running.

In order to determine which of the three to use, a TESS prototype was developed that can evaluate the voice input using all three voice recognition software. An informal test was conducted to see which of the three is most effective.

However, Whisper was out of the running early on as it was unable to be run quickly enough to use in real time. This left Vosk and Wav2Vec as the viable choices.

After further research, the decision was to proceed with Vosk because of the following advantages:

1. Offline Capability,
2. Lightweight and Fast,
3. Customizable Models,
4. Open-Source, and
5. Cost-Effective.

Although, Vosk does have a few disadvantages. Most notably, it has difficulty dealing with the following:

1. Accuracy with Complex Speech,
2. Limited Support for Certain Accents, or Dialects
3. Limited Context Understanding.

In order to further improve Vosk's accuracy in dealing with complex speech as well as to deal with accented English, Jellyfish was used in the post-processing of the transcriptions. Jellyfish implements the Metaphone algorithm in Python for phonetic encoding. The Metaphone algorithm, originally developed by Lawrence Philips, converts words into phonetic representations to allow for approximate string matching, especially useful for handling spelling variations in names.

To calculate the phonetic similarity of the transcription to the given word, the phonetic representation of both is compared using Levenshtein Distance which measures the minimum number of single-character edits (insertions, deletions, or substitutions) required to change one string into another. To determine the percentage of similarity between the phonetic representations, this formula will be used:

Similarity Percentage = (1 - Leveshtein_Distance/Max_Length) * 100

A similarity percentage of 60% or above indicates that the transcription is phonetically similar enough to the given word; which is then marked as 'PASS'. Otherwise, it is marked as 'FAIL'.

3.3.2 Audio Post-processing

To accommodate the busy environment during testing, besides using recommended equipment, audio will also undergo post-processing as needed. Specifically the following:

1. Bandpass Filtering - removes frequencies outside a specified range, keeping only a target frequency band. This is useful in speech processing, music production, and noise reduction.
2. Normalization - adjusts the volume level of an audio file to a target value, typically making it as loud as possible without distortion.
3. Compression - reduces the dynamic range of an audio signal, making the quiet parts louder and the loud parts quieter.

3.4 GUI Mockup

The following are images of the current GUI of Tess.

Figure 3.2: Student Input Screen.

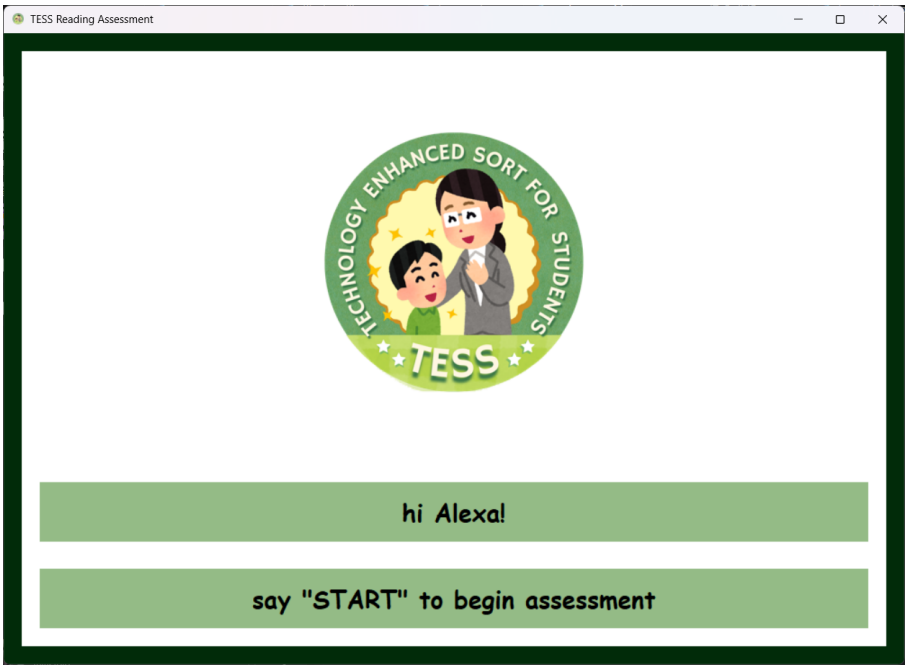

Figure 3.3: Pre-test Screen.

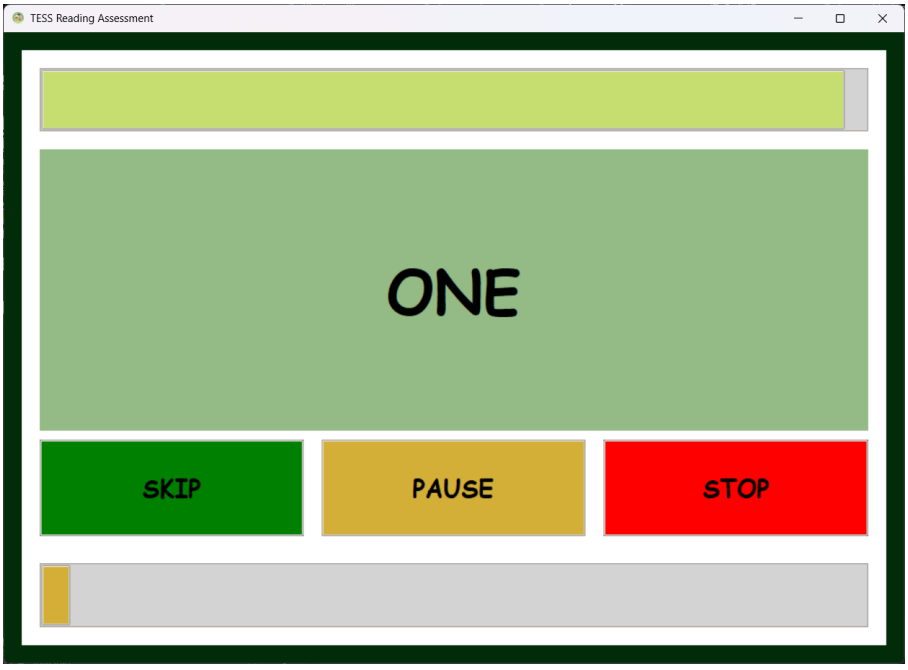

Figure 3.4: Test Screen.

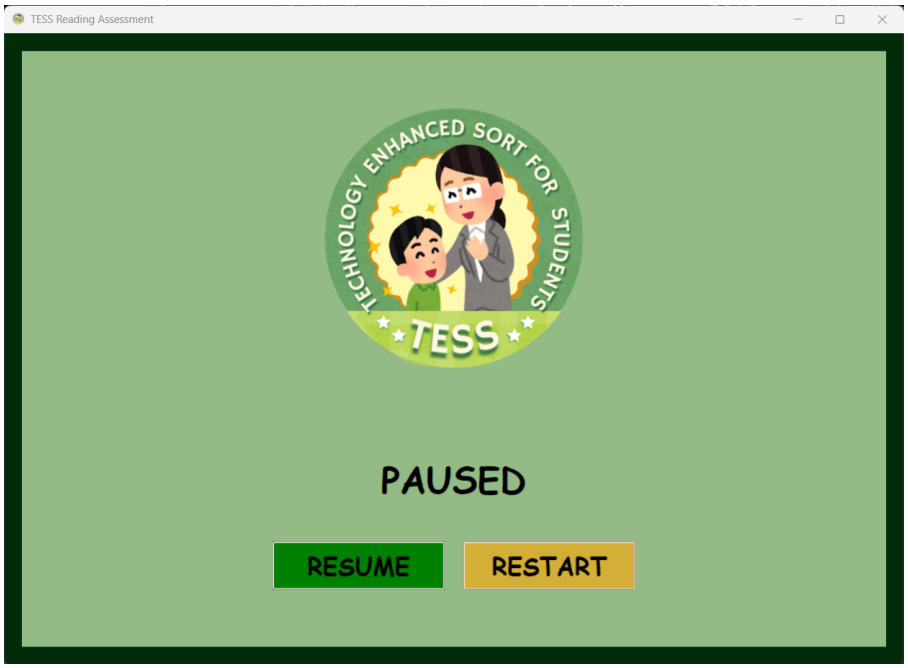

Figure 3.5: Pause Screen.

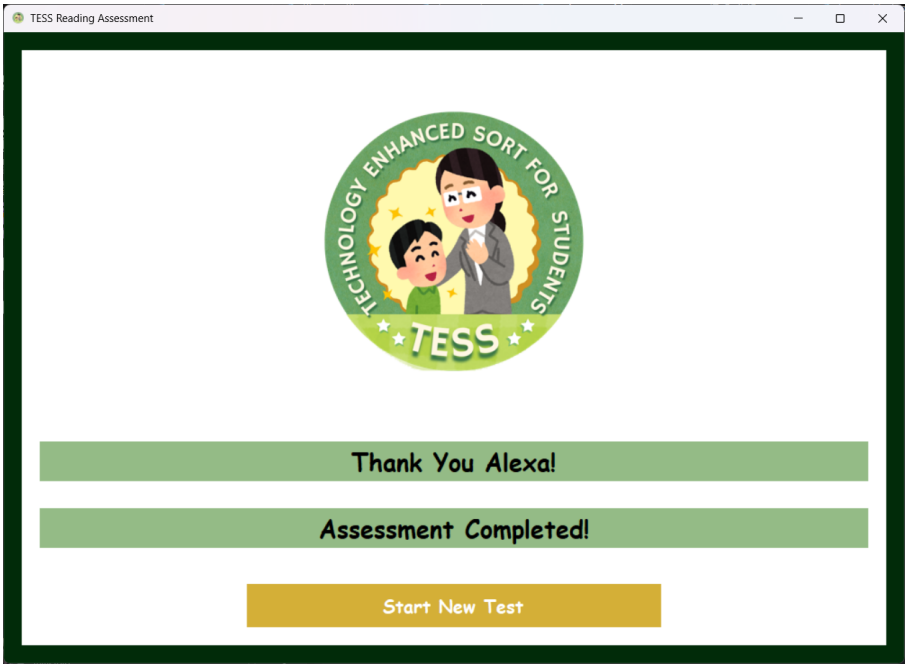

Figure 3.6: Post-test Screen.

3.5 Testing the Program

TESS was tested at the Iloilo City National High School with a group of 15 Grade 7 students. Five students of each reading profile (independent, instructional, frustration) were chosen to form the final group of 15.

The manual SORT and the computerized SORT through TESS was administered to the student simultaneously – which was done to avoid practice bias. The testing duration per student was timed and the resulting calculated reading level of the tests was used to determine the efficacy of TESS in relation to a manual SORT.

After the testing, a focus group discussion was held to determine what the user experience was like.

Consent Forms

Students participating in the TESS testing were required to fill out parental consent forms in order to participate.

Focus Group Discussion Outline

The focus group discussion was centered around the experience of the faculty and students while using TESS. The discussion lasted for around 30 minutes per batch. The following is the set of questions that was used as a guideline for the discussion.

1. What was the experience of using TESS to take the reading test like?
2. How did using TESS differ from taking / administering SORT manually?
3. What difficulties did you encounter while using TESS?
4. If you could add one feature or function to TESS what would it be?

3.5.1 Evaluating the Software

TESS was compared with Manual SORT along the following metrics:

- Speed - how fast does each student complete their assessment?
- Accuracy - how accurate are the results of each student?
- Ease of Use - how easy was the test to be administered / taken?

The Speed and Accuracy metrics were determined using the information gathered during the assessment. However, the Ease of Use metric was derived from a Focus Group Discussion composed of participating students and faculty.

3.5.2 Software Evaluation Form

To evaluate how effective TESS was along the specified metrics to the users, a software evaluation form was created for them to fill out.

While the main categories are Speed, Accuracy, and Ease of Use, they are further expanded into relevant questions which allows the researchers to better gauge TESS' performance in each of those metrics.

For each metric and its corresponding descriptive characteristic, TESS will be graded on a scale of 1 to 5; with 1 being "Poor" and 5 being "Very Good". The scale is as follows, from least to greatest.

1. Poor (P)
2. Fair (F)
3. Average (A)
4. Good (G)
5. Very Good (VG)

Questions	P (1)	F (2)	A (3)	G (4)	VG (5)	Comments (Notes)
SPEED						
(1) Program loads each word quickly and without delay.	☐	☐	☐	☐	☐	
(2) Program processes input without significant delay.	☐	☐	☐	☐	☐	
(3) Program finishes the SORT assessment quickly.	☐	☐	☐	☐	☐	
ACCURACY						
(1) Program accurately transcribes the spoken word.	☐	☐	☐	☐	☐	
(2) Program can identify successful and unsuccessful word readings.	☐	☐	☐	☐	☐	
EASE OF USE						
(1) The interface is easy to understand and use.	☐	☐	☐	☐	☐	
(2) The instructions provided are easy to understand and follow.	☐	☐	☐	☐	☐	
(3) The functionality of the program is easy to explain and teach..	☐	☐	☐	☐	☐	

Figure 3.7: Software Evaluation Form.

3.6 Statistical Treatment of Data

A mix of descriptive and inferential statistics were used to analyze the data and determine TESS's efficacy. Specifically, calculating for the mean was used to determine TESS' overall score in the software evaluation form and also to determine the average testing speed through TESS. And, to determine if there is a statistically significant difference between the test results of the Manual SORT and TESS, the Wilcoxon Signed Rank Test was used to analyze the difference between student's test scores. This test was chosen over others because the difference between the result pairs were not normally distributed and there was a need to evaluate the difference between the pair-set of test scores.

Chapter 4

Results

In this chapter, the results of the testing and development of TESS are presented. Included in this chapter are evaluations of TESS according to its speed, accuracy, and ease of use; as well as the summation of the results of the software evaluation form and focus group discussions with students and faculty.

The results in this chapter can help in determining if TESS fulfills the objective of the study which was to develop a computerized SORT that can competently administer SORT with minimal teacher involvement, be an easy-to-use and easy-to-teach application, be ran offline and with computer lab equipment, and can reliably transcribe non-native English speakers' reading of a word.

4.1 Speed

Below are the figures associated with the duration of each individual testing session:

Test duration for students with a lower reading score was shorter than those with higher reading levels – this was because those who scored higher made it further into the test than those who scored lower. Some factors to consider that affected that the testing duration is that, especially for the first batch of testers which were the Frustrated Level Readers that scored lower, the teachers and students alike were acclimating to the use of TESS and some difficulty arose in the instruction because of that lack of familiarity.

Later batches had more advantage, as the teachers had a better idea of how

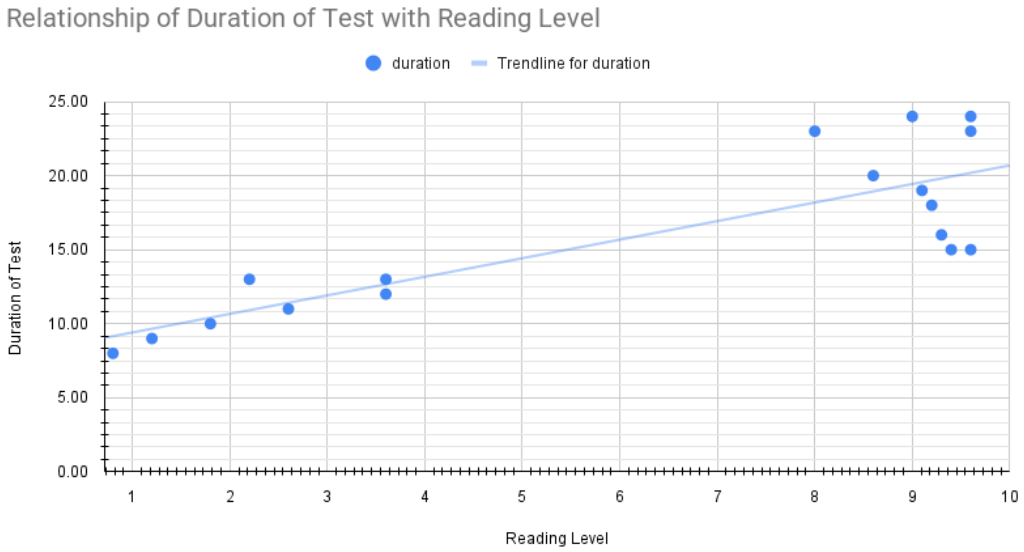

Figure 4.1: Scatter chart of the relationship between testing duration and the reading level

to proctor the test using TESS.

Notably, the duration for the entire testing process was just 2 hours and 54 minutes (inclusive of breaks and instruction times). Two students were taking the test simultaneously throughout the duration of the testing period. And at the end of those near-3 hours, 17 students were able to finish the SORT.

4.2 Accuracy

In the figure above, it is apparent that TESS calculates an overall lower reading level than that of the manual SORT. There are a couple of factors that affect the calculated reading level of a student: (1) in the manual SORT, a pass/fail for any given item can depend on the leniency of the teacher; (2) certain words can be read through the student's mother tongue and therefore is a pass in the manual SORT, and (3) TESS, as it is using Vosk, is too strict with its evaluation. However, it is already apparent that there is a significant difference between the reading level of the manual SORT and that of TESS.

Using the Wilcoxon Signed Rank Test we can determine how statistically significant the difference is between the manual SORT and the TESS results. Our

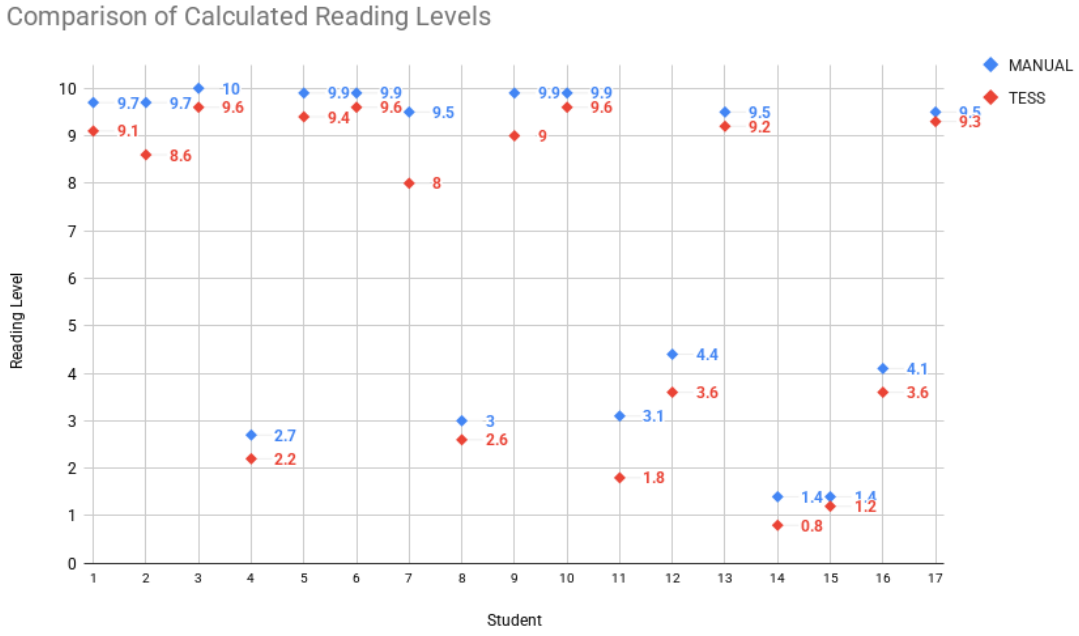

Figure 4.2: A chart comparing the reading levels derived from the Manual SORT and TESS

H0 is that the median difference between the results of manual SORT and TESS is zero, which means that there is no difference between the results. The HA is that the median difference between the results is not zero, which means that there is a significant difference between the results.

With a significance level of 0.05, a W statistic of 0, and a critical value of 34, we can reject the null hypothesis and conclude that there is a statistically significant difference in the results. Since the differences are mostly positive we can infer that TESS is too strict and scores students lower than what the teacher would.

However, the standard deviation of the difference is only 0.42 and so TESS is able to decently transcribe the reading of a non-native English speaker - with some adjustment to post-processing, TESS can become more reliable in this aspect.

4.3 Software Evaluation Form

A total of 17 anonymous student responses were received for the Software Evaluation Form. The average score for each item is as follows:

QUESTIONS	MEAN SCORE	QUALITATIVE SCORE	OVERALL MEAN SCORE	OVERALL QUALITATIVE SCORE
SPEED				
Program loads each word quickly without delay	4	GOOD	**4**	**GOOD**
Program processes input without significant delay	4	GOOD		
Program finishes the SORT assessment quickly	4	GOOD		
ACCURACY				
Program accurately transcribes the spoken word	3	AVERAGE	**3**	**AVERAGE**
Program can identify successful and unsuccessful word readings	-	-		
EASE OF USE				
The interface is easy to understand and use	4	GOOD		
The instructions provided are easy to understand and follow	4	GOOD	**4**	**GOOD**
The functionality of the program is easy to explain and teach	-	-		

Figure 4.3: A table of the average score per item and per category of TESS

This means that TESS performed well enough on the speed and ease-of-use metrics but was only average on accuracy.

4.4 Summary of Focus Group Discussion

For the focus group discussion with students, the researchers met with a total of 17 students. Seven students belonged to the frustrated readers group, five from the instructional readers group, and five from the independent readers group.

All three groups of students expressed excitement using the computerized SORT, and the majority preferred using the program over the manual method. TESS provided a game-like environment for students through instant feedback of the audio and visual cues and timer. However, some students in the frustrated reader group preferred the manual method because the manual method did not have the time pressure of a timer. The feeling of being pressured by the timer was also shared by other students, but it did not affect their preference for TESS and, to them, improved the game-like experience of TESS.

In particular, two students raised concerns about audio and visual feedback.

The first student expressed annoyance with the beeping, which affected their willingness to complete the SORT through TESS. The second student brought up that the flashing / changing of colors as a signal to reread the word caused them eye strain and suggested accommodation for that condition. Other concerns included some difficulty in getting used to the timing of the test – in regards, to reading the words and when to reread the word as the visual feedback was vague and there was a delay between the transcription process and the vocal input.

Lastly, the students requested graphical improvements to the application through including some logos for the command buttons (SKIP, STOP) and adding some slight animations were possible.

For the focus group discussion with the teachers, the researchers met with three teachers. These teachers provided essential assistance during the testing period and were able to observe and get hands-on experience with TESS.

The teachers were very positive about the experience as well. To them, it was apparent that TESS would be helpful to them and improve their process. Since the testing process was one-on-one, they were liable to be "bored" and using the computerized SORT could alleviate a proctor's stress and boredom throughout the testing period. They also believed that TESS would allow them to administer SORT with a group of at least five students, and that TESS would be able to do so quickly. The requested time period for TESS' testing period was one full day, but the test concluded after only around 3 hours – this really bolstered their faith in TESS' speed.

There were some slight technical issues that were noted, however, which included how to teach and coach students on how to use TESS and also how to properly use the headset.

The teachers had more significant suggestions for the improvement of TESS:

- Including celebratory animations or graphics at the conclusion of the test for positive reinforcement.
- Including an interactive tutorial with unique practice words, in order for the students to familiarize themselves with the program and also get used to the timing of the program's response time.
- Implement a system in which the program will 'remember' the current session of a professor to reduce the amount of time it takes to setup the program for a student.

Creating a manual for the teachers, so that they are able to study and understand the functions of TESS and, in turn, are able to assist the students..

Overall, the teachers were excited about the prospect of actually using TESS during enrollment.

Chapter 5

Conclusion

This Special Problem sought to develop a program that could provide a reliable Computerized SORT that not only met the requirements of the ICNHS Remedial Reading Committee, but was also:

- Easy to use with minimal teacher supervision,
- Able to run offline in a computer laboratory,
- Using voice recognition software that could reliably transcribe non-native English speakers.

The researchers developed Technology Enhanced SORT for Students (TESS) and conducted a test at ICNHS, with a total of 17 students from the frustrated, instructional, and independent reading level groups. Using two sets of computer units and headsets, and with the assistance of two teachers, all 17 students were able to be evaluated in around 4 hours.

In regards to speed, TESS is fairly efficient and is able to complete the testing of 17 students in around 3 hours while using two computer units simultaneously. It is easy to imagine that with the addition of extra units, the rate at which teachers can administer the SORT will increase. In the software evaluation form, students rated TESS' spred an average of 4 or "Good".

As for accuracy, TESS rated students at a lower reading level than what teacher's would through using the manual SORT. Through the Wilcoxon-Signed Rank Test it has been determined that the difference between both results are statistically significant enough that adjustments need to be made to TESS' processing and post-processing in order for TESS to be more reliable and effective. A

potential improvement or follow up to this study would be to create a data-model on Ilonggo English speakers, particularly of the junior high demographic. Doing so would ensure that programs such TESS is able to work off of more relevant data. In the software evaluation form, student's rated TESS' accuracy an average of 3 or "Average".

When it comes to ease of use, TESS rated (4) "Good" on the software evaluation form. Both students and teachers have also expressed excitement and positivity towards the use of TESS. Although some improvements to the UI as well as an implementation of a more robust tutorial for students and documentation for teachers was suggested. Overall, there were minor technical issues involving the headset and the timing for reading the words, but these can be solved through the suggestions provided. It is recommended that focus on accessibility features in the UI and UX be prioritized; especially in regards to flashing and eye-strain inducing graphics.

The teacher's greatly appreciated and were very excited for TESS and its potential for alleviating the stress of the screening for remedial readers.

It can be concluded that TESS as a computerized SORT could improve ICNHS' Remedial Reading Assessment at the beginning of the year, however improvements to TESS' accuracy and additional feature implementations are required before it can be fully launched.

References

Abergos, L. I., Cruz, J. R., Lasala, J. C., Prado, S. N., Tapar, P. K., Cañeza, D. C., & Ocampo, D. M. (2024). Effectiveness of remedial reading to struggling readers of grade 7 students. *SHS Web of Conferences*, *182*, 01004. doi: 10.1051/shsconf/202418201004

Auphan, P., Ecalle, J., & Magnan, A. (2020, Sep). High potential of computer-based reading assessment. *Canadian Journal of Learning and Technology*, *46*(1). doi: 10.21432/cjlt27847

Estuya, J. A. (2024). *Home and school based-factors affecting the reading performance of frustration-level readers* (Unpublished doctoral dissertation). Columban College, Inc.

Lee, G.-L., Gable, R., & Klassen, V. K. (2012). Effective reading remediation instructional strategies for struggling early readers. *Procedia - Social and Behavioral Sciences*, *46*, 822–827. doi: 10.1016/j.sbspro.2012.05.206

Lovett, M. W., Frijters, J. C., Steinbach, K. A., Sevcik, R. A., & Morris, R. D. (2021, May). Effective intervention for adolescents with reading disabilities: Combining reading and motivational remediation to improve outcomes. *Journal of Educational Psychology*, *113*(4), 656–689. doi: 10.1037/edu0000639

Rafanan, J. A., & Raymundo, J. L. (2024, Jul). Impact of customized reading remediation and intervention in reading fluency of high school students. *International Journal of Linguistics, Literature and Translation*, *7*(7), 66–74. doi: 10.32996/ijllt.2024.7.7.8

Sancada, R. T. (2022). Remedial reading program in the schools division of iloilo: Challenges and effective practices. *Psychology and Education: A Multidisciplinary Journal*, 1–19. doi: 10.5281/zenodo.6957238

Second Congressional Commission on Education. (2024). *Miseducation: The failed system of philippine education, EDCOMM II year one report* (Tech. Rep.). Second Congressional Commission on Education.

Slosson Educational Publications, Inc. (n.d.). *Slosson Oral Reading Test - Revised.* Retrieved from https://www.slossonplus.com/sort-r3

United States Agency for International Development. (2022, Nov). *Computer-based reading assessment pilot report* (Tech. Rep.). United States Agency for International Development. Retrieved from https://shared.rti.org/content/computer-based-reading-assessment-pilot-report