A Natural Language Approach For Student's Speech Assessment

Mrs. R. Mridula
Sr. Lecturer,
Department of MCA,
Tamilnadu College of Engineering Coimbatore
mriduvenu@gmail.com

Dr. Kalugasalam P
Prof. & Head
Department of Science & Humanities
Tamilnadu College of Engineering Coimbatore
drponkalugalasam@gmail.com

ABSTRACT - The objective of reading is understanding. In order to understand a page, children must be able to decode the words on the page and to extract meaning. Many of the research focus on how children learn to decode text and how best to foster children's decoding skills. Oral Reading Assessment is an accessible computer program that uses speech recognition to provide an accurate measure of student's oral reading ability. It presents grade-level text passages to students, who read the passages out loud, and computes the number of words correct per minute (WCPM), a standard measure of oral reading fluency. Successful reconstruction of the text would transform these errorful utterances into fluent strings. This work is an extension of FLORA which currently measures reading speed and accuracy. Though the work is in progress, the components of the system architecture and the speech recognition subsystems are described, which is an essential computational linguistic tool needed for many natural language processing (NLP) applications

KEYWORDS: Natural Language Processing; Speech recognition; Oral Reading Fluency, Oral Reading Assessment;

I. INTRODUCTION

Oral reading fluency is used to assess an individual's reading level and proficiency. In recent years, natural language processing tasks such as machine translation, information extraction, and question answering have been steadily improving, but little importance was applied to the most natural form of language input: fluency in speech. Reading fluency has been defined as the ability to read a text quickly, accurately, and with proper expression .While assessment of reading is usually done with measures of reading comprehension, direct reading is often performed using oral reading. In this paradigm, students read texts aloud and their proficiency in terms of speed, fluency, pronunciation, intonation etc. can be monitored directly while reading is in progress. Oral reading has been one of the best diagnostic and predictive measures of foundational reading weaknesses and of overall reading. Reading assessments are used for identifying students who need immediate help, for making decisions about reading instruction, for monitoring the students' progress throughout their educational years, for comparing and evaluating reading programs and for reporting annual academic outcomes.

There are many significant benefits of automating or partially automating the assessment process. Data from automated assessments, including digitized recordings, can be entered into a database for all student assessments, enabling teachers to review progress reports for individual students as well as to listen to samples read aloud across successive assessments. The data could also be analyzed, summarized, and displayed to answer questions about changes in students' reading abilities for classrooms and colleges within and across college districts. This paper describes the system architecture, technology components, and performance of ORA, a fully functional Web-based system which estimates individual student's oral reading fluency which lasts for about five minutes. In Section 2 the background for developing ORA and the previous work on automatic fluency assessment are defined. Section 3 describes ORA architecture. Section 4 describes the data collection. Section 5 describes the results, and

Section 6 presents conclusions and discusses for future work that aims to improve performance.

II. PRIOR WORK

The main objective measures of speech, including rate of speech and phonation-time ratio, correlate well with fluency scores both for reading and spontaneous speech. There are many researches using speech recognition to assess and improve reading. Research conducted by Jack Mostow and his colleagues in Project Listen at Carnegie Mellon University has demonstrated the effectiveness of speech recognition for improving reading fluency and comprehension for both native and nonnative speakers of English [1]. Another approach in measuring oral reading fluency, used an ASR system to measure a student's interword latency, defined as the elapsed time between certainwords read aloud by the student that were scored as correctly read by the ASR system. They argue that latency "acts as a microscope to allow us to zoom in at the time the student takes to figure out how to pronounce a word, but does not include the time the student requires to actually say the word." Their model of interword latency produced a correlation of over 0.7 with independent WCPM measures of oral reading fluency using grade level passages.

In the context of Project T-ball (Technology Based Assessment of Language and Literacy) at UCLA and USC, investigated oral reading of 55 isolated words produced by kindergarten, 1st and 2nd grade children with the aim of detecting reading miscues automatically, such as sounding-out, hesitations, whispering, elongated onsets, and question intonations [2]reported preliminary results of a system for automatic scoring of oral reading fluency in text passages and word lists for middle school students. Pearson correlations between automated and human scores were 0.86 for passages and 0.80 for word lists. [3]describes the system architecture and initial performance of a reading coach that runs on both PCs and handheld devices. A series of studies by Bryan Pellom and Andreas Hagen and their collaborators [4][5] investigated ways to optimize the Sonic speech recognizer for children's speech. The research resulted in a reduction in the word error rate (WER) from 17.4% to 7.6%..[2][4]; developed a version of Sonic that uses subword modeling. The motivation for this work is the observation that a number of reading disfluencies in children's speech occur at the sub word level (e.g., "ba- ba- banana") so they are better modeled using subword lexical units, like syllables, as the basic unit for speech recognition. In the study several subword lexical units and approaches were evaluated for the detection of disfluencies and modest gains were reported [6] that additional detection gains can be achieved by using syllable graphs to represent hypotheses from the speech recognition system and to obtain confidence estimates. Investigated the potential benefits of using Support Vector Machines (SVMs), a powerful Machine Learning classifier approach, to verify children's speech while reading aloud [6]. SVMs were used during a second recognition pass to reclassify subword units previously recognized using Sonic[3]. His research led to further reductions in classification error rates during oral reading when tested on the same children's speech data used in Hagen's [2006] research[2]. During a training session college students were trained to score words as read correctly or incorrectly produced in recordings of children reading texts out loud from the test set of the CU Read and Summarized Stories [7][8]. Raters were trained to mark words on a printed copy of the text that were skipped over, mispronounced or substituted for other words. At least two judges scored each recorded passage independently. The judges scored the entire passage (not just one minute), and were able tostart and stop the recording while scoring. The WCPM score for each successive minute of speech was compared for the two judges. The WCPM score for Sonic was compared to the scores produced by each judge to produce the average agreement between Sonic and the two judges. The average agreement between the two judges was about 95%. The average agreement between Sonic and each of the judges was 92%.[9] Iinvestigated the potential benefits of FLORA which is a fully functional oral reading assessment system that can be accessed online from any major Web browser and used on any PC or Mac. It requires the user to read text passages aloud using a microphone. Here speech processing and data persistence are managed by the server machine. ORA currently measures reading speed and accuracy, but does not measure how expressively the passage was read[1].

III. THE ORA SYSTEM

(i) System Overview

ORA is an oral reading assessment system that can be accessed online from any major Web browser. It requires the user to read text passages aloud using a microphone. Speech processing and data persistence are managed by the server machine. ORA is planned to run in two different modes, which reflect the methods of assessing oral reading fluency in colleges today. In reading mode, the student is presented with a text passage at his or her grade level, and is instructed to read the passage out loud and to skip those words that he or she cannot read. In *word assists mode the* student is instructed to read the passage out loud. When ORA is used in assisting reading mode, the student can use a mouse to click on words they cannot read, and the words are spoken by the system.

During system use, the following features are supported for the student.

(a) *ORA enables* students to enroll in the system by providing information about the student's gender, age, and grade level. One of the screenshot showing gender is given in figure 1.

(b) *ORA* instructs the student that a text will be displayed for reading out loud, and displays the text. Figure 2 displays the text to be read by the student.

(c) *ORA stops* the recording after one minute, and thanks the student.

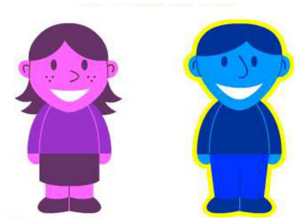

Fig. 1. ORA screenshot showing the gender

Faculty Interface. The Faculty interface was developed to enable individuals to test the system and to score passages. This enables the faculty to:

(a) View the WCPM score computed by ORA for the student and the percentile in which the student's score falls.

The Three Races

"In old times, fable retells the story of the young athletic boy hungry for success, for whom winning was everything and success was measured by such a result.

One day, the boy was preparing himself for a running competition in his small native village himself and two other young boys to complete. A large crowd had congregated to witness the sporting spectacle and a wise old man, upon hearing of the little boy, had travelled far to bear witness also.

The race commenced, looking like a level heat at the finishing line, but sure enough the boy dug deep and called on his determination, strength and power,, he took the winning line and was first. The crowd was ecstatic and cheered and waved at the boy. The wise man remained still and calm, expressing no sentiment. The little boy, however, felt proud and important.

A second race was called, and two new young, fit, challenges came forward, to run with the little boy. The race was started and sure enough the little boy came through and finished first once again. The crowd was ecstatic again and waved at the boy. The wise man remained still and calm, again expressing no sentiment. The little boy, however, felt proud and important.

"Another race, another race!" pleaded the little boy. The wise old man stepped forward and presented the little boy with two new challenges, an elderly frail lady and a blind man. "What is this?" quizzed the little boy. "This is no race "he exclaimed. "Race!", said the wiseman. The race was started and the boy was the only finisher, the other two challenges left standing at the starting line. The little boy was ecstatic; he raised his arms in delight. The crowd, however, was showing no sentiment toward the little boy.

"What has happened? Why not do the people join in my success?" he asked the wise old man."Race again" replied the wise man," this, finish together, all three of you, finish together' continued the wise man. The little boy thought a little, stood in the little boy walked slowly, ever so slowing, to the finishing line and crossed it. The crowd was ecstatic and cheered and waved at the boy. The wiseman smiled, gently nodding his head. The little boy felt proud and important.

"Old man, I understand not! Who is the crowd cheering for? Which one of us three?", asked the little boy. The wise old man looked into the little boy's eyes placing his hands on the boy's shoulders, and replied softly "Little boy, for this race you have won much more than in any race you have ever ran before, and for this race the crowd cheer not for any winner!"

Fig. 2. ORA screenshot showing the story to read

(b) Listen to a recording of the student reading a text passage while viewing the text, and click on those words that were skipped or read incorrectly. When the faculty clicks "Done," the

WCPM score is displayed, along with the associated percentile for each grade level. The percentile mapping is based on published grade level norms of oral reading of grade level passages by thousands of students collectedduring the fall, winter and spring of the college year.

(ii) ORA Architecture

Fig.3 shows the ORA architecture, with its modules, data flow and control. The figure also presents information about the communication protocols and the technology utilized.

Client-side. On the client side, a Web browser loads the ORA Web application. The application consists of:

(a) An embedded Flash application which supports the interactions with the user and providing the support for manual scoring.

(b) A Java Applet which records the audio and transmits it to the server by means of a socket. Both the Flash and the Java Applet are synchronized by means of a TCP connection on the client side.

Server-side. The server side consists of three modules which includes a Web server. , a Java application which receives the audio from the client, stores it in an audio repository and sends it to the speech processing module and the C++ based automatic speech recognizer tracker which receives the speech from the client and WCPM is calculated and percentile scores can be viewed immediately by a faculty.

IV. DATA

The speech data were parameterized using Discrete Wavelet Packets Transform and perceptual linear prediction coefficients (PLP) for noise robustness. Acoustic models were trained under Maximum Likelihood. After each re-estimation iteration, two Gaussian components were added to each mixture Speech recognizer uses a static decoding network organized as a prefix-tree and it was compressed using a forward-backward merging of nodes. An unsupervised adaptation of the means and variances of each of the Gaussian distributions was carried out using Maximum Likelihood Linear Regression (MLLR) before the second recognition pass. The regression tree is used to cluster the Gaussian distributions which comprise of base-classes and theminimum count to compute a transform was set to 300 feature frames. Expectation-Maximization clustering was used to clustertheGaussian means. The ORA system was configured to enroll each student, and then randomly select one passage from a set of 20 standardized passages of similar difficulty at the student's grade level. Depending upon the number of students that needed to be tested on a given day, each student was presented either two or three text passages to read aloud. During the testing procedure, the student was seated before the laptop, and asked to put on a set of headphones with anattached noise-cancelling microphone. Theexperimenter observes and helps the student to enroll

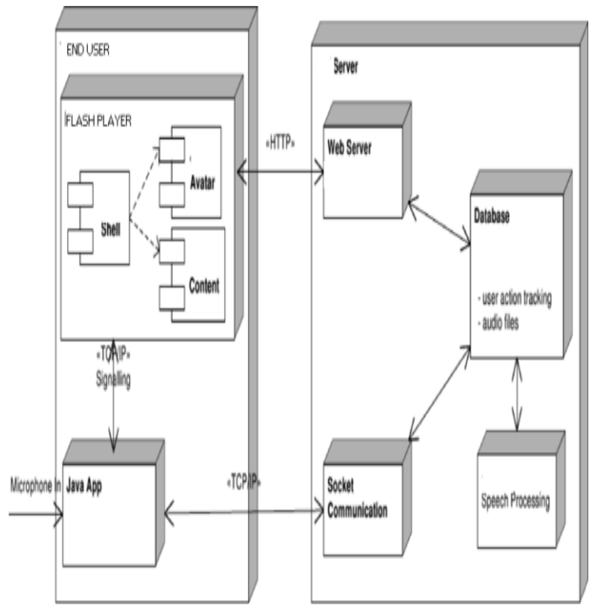

Fig. 2 ORA architecture

in the session that involves entering the student's gender, age, and grade level. ORA then presents a text passage, starts the one minute recording at the instant the passage was displayed, captured the student's speech and relayed the speech to the server. About 20% of the time, teacher directs that specific students are presented with text passages either one or two levels below or one or two levels above the student's grade level. Twenty text passages were available for reading at each grade level. The standardized text passages were downloaded from a net and are freely available for noncommercial use. The twenty passages were designed to be about the same level of difficulty at each grade level, and were designed specifically to assess oral reading fluency. Oral reading fluency norms have been collected for this text passage for tens of thousands of students at each grade level so that students can be assigned to percentiles based on national WCMP.

Grade	Ist	2nd	3rd	4th	all
# recordings	128	250	160	220	758
# colleges	2	3	2	2	4
# students	53	104	66	90	313
hours of audio	2:12	4:19	2:45	3:47	13:0

Table1. Summary of the Data Used in the evaluation

Grade	#words	#unique words	#sentences	# word / sentence
Ist	220	107	25	9
2nd	240	118	21	12
3rd	255	127	20	12.75
4th	381	184	28	13.61

Table 2. Statistics of the Stories for Each Grade

V. RESULTS

The partial results indicate that ORA gives an accurate estimate of mean WCPM scores for groups of students in colleges with different literacy achievement levels and different ethnographic characteristics. The relative inter-human agreement and the ORA to human agreement is very close which means that ORA performs very well at identifying students that might require additional reading assessments and instruction.

VI. CONCLUSION

The general pattern of results obtained with the initial ORA prototype with respect to WCPM scores were close to scores produced by human scorers, with mean differences of 3 to 4 words. The results suggest that ORA could be used as a tool for identifying students who may be at risk for learning to read and deserve further attention. In terms of progress monitoring—tracking changes in oral reading fluency through periodic assessment of students in response to an instruction—further research is needed to assess the potential of ORA as a valid and reliable assessment tool. There are many ways to improve the current system. The initial prototype, which was completed "just in time" to collect assessment data before the end of the academic year, has not been optimized in any way. Future work will focus on understanding the nature of the errors the system now produces (relative to human judgments), and pursuing established methods ofimproving the performance of the speech recognizer and reading tracker.

While additional research is needed to understand the relationship between expressive reading and comprehension, it is clear that developing valid automatic measures of expressiveness during oral reading will be desired and valued by teachers. Initial research in this area has produced promising results although much work remains to be done in this area, including defining and developing valid and reliable measures of expressive reading.

REFERENCES

[1] Mostow J.,Aist, G.,Burkhead P.,Corbett, A.,Cuneo, A.,Eitelman S.,Huang, C., Junker, B.,Sklar, M. B and Tobin B. 2003. Evaluation of an automated reading tutor that listens: Comparison to human tutoring and classroom instruction. J. Educ. Comput. Resea. 29*,* 1, 61–117.

[2] *Zechner,*K.,Sabatini, J., AND Chen, L. 2009. Automatic scoring of children's read-aloud text passages and word lists. In *Proceedings of the NAACL-HLT Workshop on Innovative Use of NLP for Building Educational Applications*

[3] Li, X., Y. C. Ju, L. Deng & A. Acero. (2007). Efficient and Robust Language Modelingin an Automatic Children's Reading Tutor System. Proc. IEEE International Conference on Acoustics, Speech and Signal Processing ICASSP 2007

[4] Hagen, A., Pellom, B., Van Vuuren, S., and Cole, R. 2004. Advances in children's speech recognition within an interactive literacy tutor. In Proceedings of the Conference of the North American Chapter of the Association for Computational Langustics—Human Language Technologies. 25–28..

[5] Hagen, A., B. Pellom, & R. Cole. (2007). "Highly accurate children's speech recognition for interactive reading tutors using subword units." Speech Communication 49(6): 861-873.

[6] Bolanos, D. 2008. Advances in the application of support vector machines for continuous automatic speech recognition. Ph.D. thesis, Computer Science Department, Universidad Autonoma de Madrid

[7] Cole, R.,Hosom, P., and Pellom, B. 2006. University of Colorado prompted and read children's speech corpus. Tech. rep. TR-CSLR-2006-02, Center for Spoken Language Research, University of Colorado, Boulder.

[8] Cole, R. and Pellom, B. 2006. University of colorado read and summarized stories corpus. Tech. rep. TRCSLR- 2006-03, Center for Spoken Language Research, University of Colorado, Boulder

[9] Daniel Bolanos, Ronald A. Cole, Wayne Ward, Eric Borts and Edward Svirsky, "FLORA: Fluent Oral Reading Assessment of Children's Speech", ACM Transactions on Speech and Language Processing, Vol. 7, No. 4, Article 16, Publication date: August 2011.