Unified Spoken Language Proficiency Assessment System

Sunil Kumar Kopparapu, Ashish Panda

TCS Research, Tata Consultancy Services Limited, India

27th Oriental COCOSDA Oct 17-19, 2024

Introduction

- Language proficiency assessment is a common requirement for L2 speaker of English.
- There exist separate tools for
 - **spoken language assessment** (SLA) to assess articulation in terms of pronunciation or oral fluency and
 - written language assessment (WLA) to assess language grammar and vocabulary.
- Unified spoken language proficiency assessment (USLPA) system
- Desired
- easy to administer,
- not repetitive,
- * automatic,
- short assessment duration, and
- * should be reliable and robust in its assessment.
- Advantages
- * removal of inherent human bias in the assessment process
- ★ greater flexibility of assessment
- dynamic assessment; **not** same assessment tests for all candidates

Language Assessment (Written | Spoken)

- Spoken Language Assessment
 - quality of articulation of speech (in terms of pronunciation)
 - > speech delivery in terms of (oral fluency which includes speaking rate, recognition of pauses, filler words, and analysis of intonation)
- Written Language Assessment
 - language grammar, vocabulary

The Problem

- Speech quality and language grammar assessment are administered separately (together they assess the complete language proficiency)
- The use of SLA and WLA separately
 - makes the language proficiency test lengthier
 - ▶ allows the learner to be slack in terms of language grammar, for example during a SLA test.
- Tests are **repetitive**; WLA tools choose from a large bank of text paragraphs

Why Grammar not assessed through Speech?

- Language assessment are fairly automated when administered on text
 Why? Availability of robust natural language text processing (analysis) tools
- SLA systems do not cater to language grammar and vocabulary assessment
 Speech Analysis tools not as robust, yet
 - * Why? Wide differences between spoken and written language text
 - \star /when uhh where do does one use CNN/ (spoken) \leftrightarrow (written) "Where does one use CNN"
- Grammar assessment, requires the precise sequence of words spoken as spontaneous speech
 - Most automatic speech recognition (ASR) engines transcribe *spontaneous speech* erroneously compared to *read speech*
 - Leading to erroneous language assessment scores

USLPA (Unified spoken language proficiency assessment)

Non-repetitive Speech Friendly Assessment

- Use GenAl to generate a **new** paragraph on the fly
- Elicit read speech instead of spontaneous speech
- Use custom language model in ASR

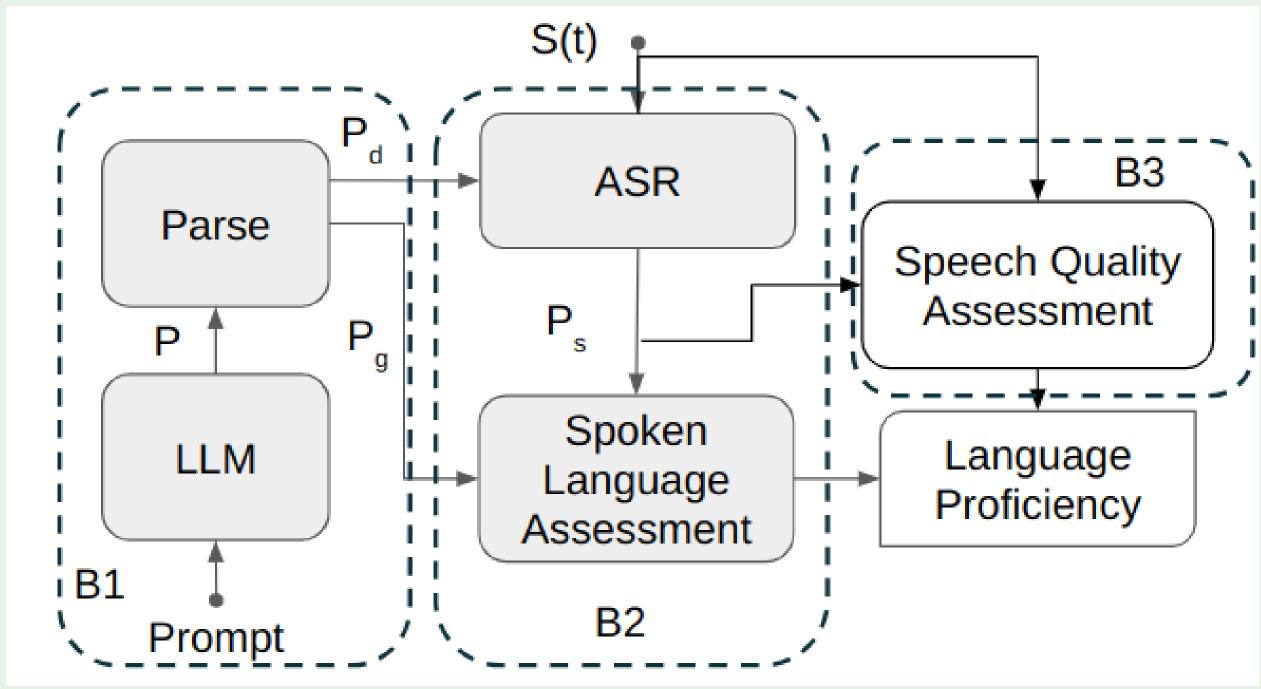


Figure: USLPA System.

Our system consists of three parts,

- ullet (B1) dynamic text generation using a large language model (LLM) which generates a wide variety of non-repetitive paragraphs P which makes it possible to administer unique assessment test
- (B2) a spoken **language assessment** module that assess language grammar and vocabulary from spoken speech S(t) and
- (B3) **speech quality** assessment module.

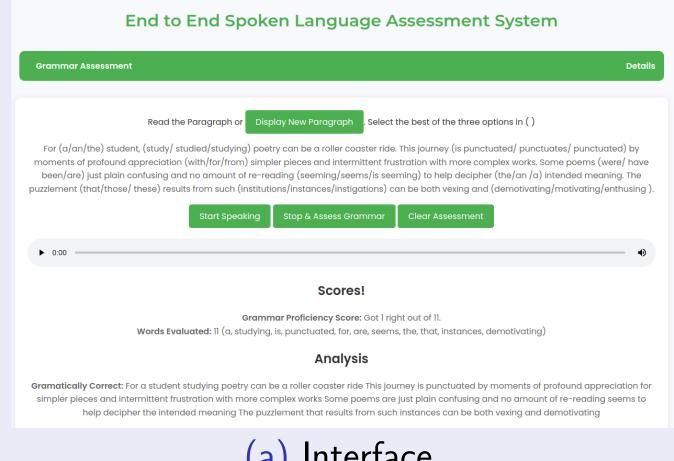
(B1) Paragraph Generation (Using GenAl | LLM)

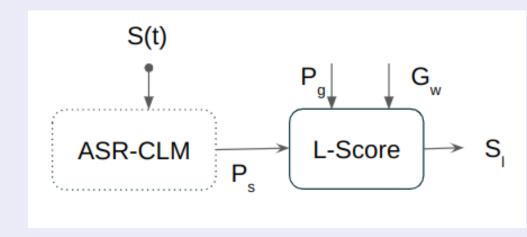
- 1-shot learning prompting style for generating new paragraph (P)
- provide a sample paragraph (P) which has both correct and incorrect grammar options to ChatGPT.

Example (Using ChatGPT)

- **#1 User:** """P"". Generate paragraph like P. One <correct></correct> tag within <grammar> </grammar> tags. Each <grammar> tag to have three options separated by "/".
- **#1 ChatGPT:** Thank you for providing the specific format and instructions. The grammar choices are marked within <grammar>, with the correct option indicated using <correct>.
- #2 User: Generate a paragraph similar to the example shown.
- **#2 ChatGPT:** In <grammar><correct>an</correct>/a/the</grammar> bustling city, <grammar>exploring/ explored/ <correct>exploration</correct> </grammar> can be an exciting adventure. <**snip**> The challenge <grammar><correct>that</correct>/ those/these</grammar> comes from such <grammar>adventures/<correct>explorations</correct>/explorers</grammar> can be
 - <grammar>adventures/<correct>explorations</correct>/explorers </grammar> can be
 both thrilling and <grammar>eye-opening/<correct>exhausting</correct>/
 insightful</grammar>.
- #3 User: Generate a para. With subject "learning physics".

(B2) Spoken Language Assessment





(a) Interface.

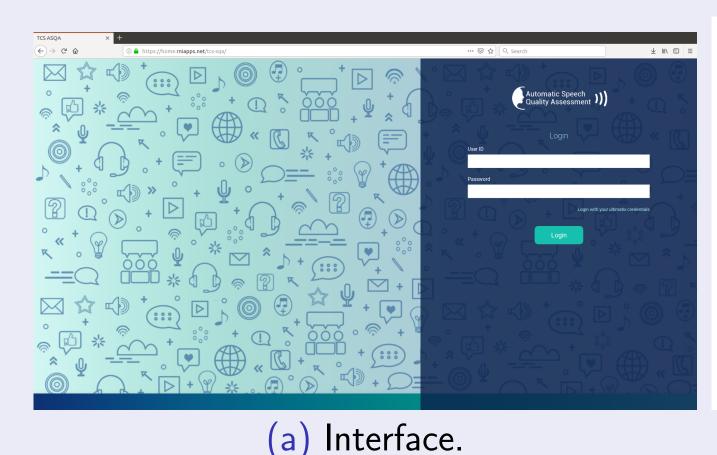
(b) Custom language model ASR.

Figure: Spoken Language Assessment. G_w is the list of words in P that are being evaluated for language (words between the "<comment>" tag).

- Use P to build a customized language model (CLM) to be used with ASR
 Why? Performance of ASR-CLM for assessment of language grammar is far superior than the state-of-the-art end-to-end ASR system like whisper
- Grammar score $S_l = \text{L-SCORE}(P_s, P_d, G_w)$ is computed by
 - extract the set of words $\in P_d$ but $\notin P_s$ (transcript); $p_1 = \{w \in P_d \mid w \notin P_s\}$
 - extract $p_2 = \{ w \in G_w \mid w \notin p_1 \};$
 - the language score, $S_1 = |p_2|$; cardinality of the set p_2

(B3) Speech Quality Assessment

- The read speech S(t) that was used for language assessment is used for speech quality assessment as well.
- Assess speech quality
 - pronunciation (Pr) speaking rate (Sr), stress (Ln), oral fluency (Of), emotion (Ta) which have been developed in-house based on **our** published literature



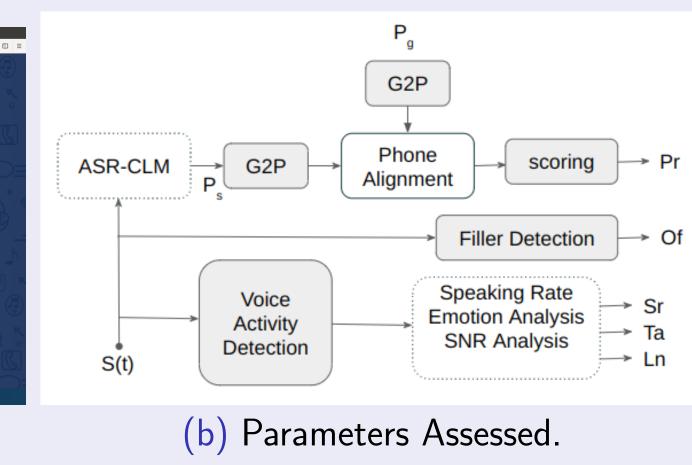


Figure: Speech Quality Assessment.

Conclusion

- Language proficiency assessment is a common requirement for L2 speakers
- SLA assessment tools to analyze speech (pronunciation, oral fluency) exist;
 none for assessing grammar
 - we designed and implemented
- ▶ a practical, scalable and a robust USLPA system that can assess
 - speech quality (pronunciation, oral fluency) andspoken language (grammar, vocabulary) assessment.
- End-to-end language assessment
 - scalable: non-repetitive tests using LLM's,
 - practical: hard to be memorized by students
 robust: read speech instead of spontaneous speech;
 - automatic: No human bias

O-COCOSDA (Oct 17-19, 2024)