

[Colloquium3]

AI-Based Automatic Speaking Assessment: Challenges and Opportunities for L2 Learners Facing Limited English-Speaking Opportunities in Their Home Country

Heyoung Kim, Jin-Hwa Lee (Chung-Ang University),
YunDeok Choi (Chungnam National University)



Overview of This Colloquium



This colloquium aims to introduce an ongoing research project funded by the National Research Foundation of South Korea, **AI-based automatic system for diagnosing, practicing, and assessing Korean K-12 students' English speaking proficiency**

01

Introduction

English Education in Korea (10 min.)

02

System Design

Designing a task-based English speaking diagnosis and learning system using AI technology and automated assessment (30 min.)

03

Key Feature I

Fostering and Assessing L2 Speakers' Oral Interaction Ability with **AI Chatbot** (20 min.)

04

Key Feature II

Designing Accessible **Digital Score Reports**: Enhancing Stakeholder Understanding and Educational Opportunities for L2 Speaking (20min.)

05

Q&A

(20min.)



01

Introduction



English education in Korea



<https://www.hansannews.com/news/articleView.html?idxno=53997>

Public School English (3rd~12th grade)

- Teacher-centered (NNET)
- Large class size
- Limited opportunity for spoken interaction
- Mostly reading and grammar focused
 - College entrance exam: Academic Reading focused
 - Paper-based school exams



7 out of 10 middle and high school students are 'dissatisfied' with grammar-focused school English classes

A Reporter Jang Yoon-won | ©Approved 2015.01.14 14:29 | Comments 0



- According to the recent survey report(Jung 2015), **990 Korean middle and high school students** answered

From an new article <https://www.nbntv.kr/news/articleView.html?idxno=14126>

Q) Are you satisfied with your English class?

No  **67.5%**

Q) Which skill is the most emphasized in your English class?

Grammar  **58.4%**

Speaking  **9.5%**

Q) Which skill do you think the most important for learning English?

Speaking  **44.1%**

Grammar  **18.5%**

Q) Learning English in school is not enough to develop English proficiency

 **I agree 68.7%**

Q) Private lesson helps me study English in school

 **I agree 84.4%**

Q) Do you want to take private lessons for learning English

 **Yes 73.2%**

Speaking is a highly desired yet underexplored skill in public education.

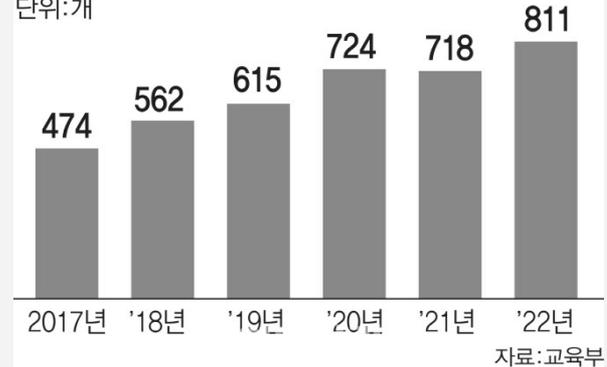
English education in Korea



Private English Kindergarten

The Number of English Kindergarten for Preschoolers

단위:개



Tuition Comparison English Kindergarten vs. University

Kindergarten

1351



University

679



자료:사교육걱정없는세상

<http://www.gnmaeil.com/news/articleView.html?idxno=196275>

“English Divide” = Social Issue in Korea



English education in Korea



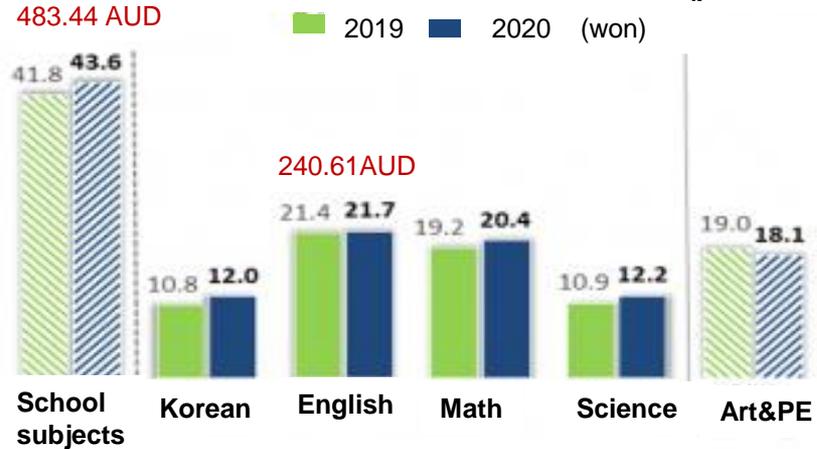
Private English lessons

2021

Q) What is the most necessary private lesson for your child if you choose only one?



The average monthly cost of private education (per student)



<https://www.edpl.co.kr/news/articleView.html?idxno=880>

<https://m.edaily.co.kr/News/Read?newsId=02840486628981680&mediaCodeNo=257>

Excessive Social Costs



English education in Korea

Statistics of English Education in Korea

N= 1000, of age 20s~50s of Koreans in 2022

I want to speak English well



나는 영어를
잘하고 싶다
(단위: 동의율 %)

2019 (N=1000)

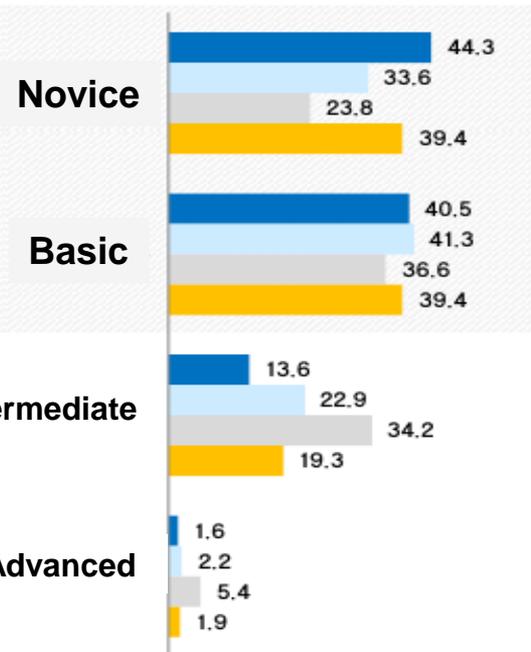
86.6

2022 (N=1000)

89.2

Self-Assessment of My Current English Proficiency

Speaking Listening Reading Writing



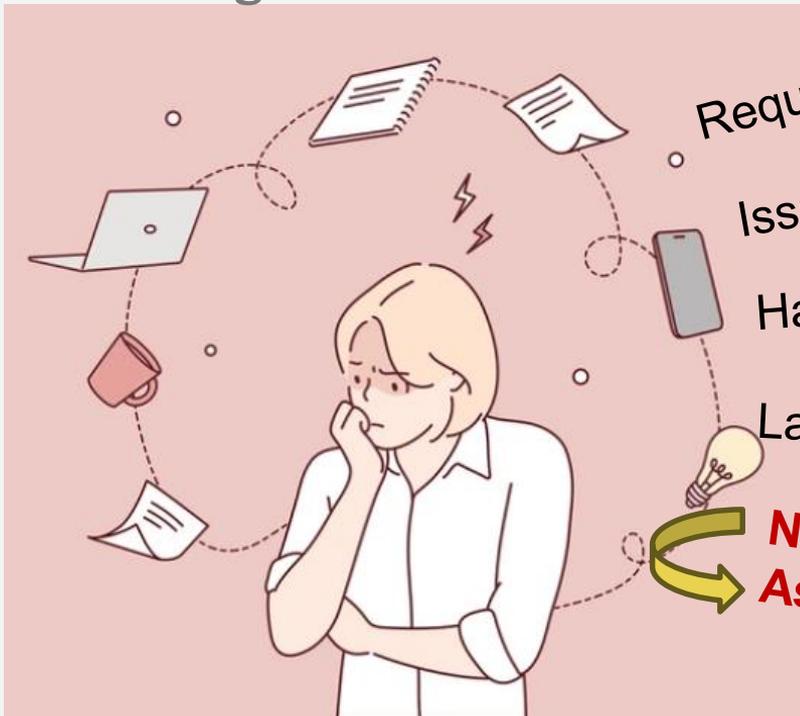
I have English
phobia
50.6%

There would be
more
opportunities if I
spoke English well

90.2%

English education in Korea

Korean English Teachers' Dilemma



Request for teaching speaking skills

Issue of objectivity of speaking evaluation

Hard to give feedback on too many students
(Goh & Burns 2012)

Lack of self-confidence with rating spoken data

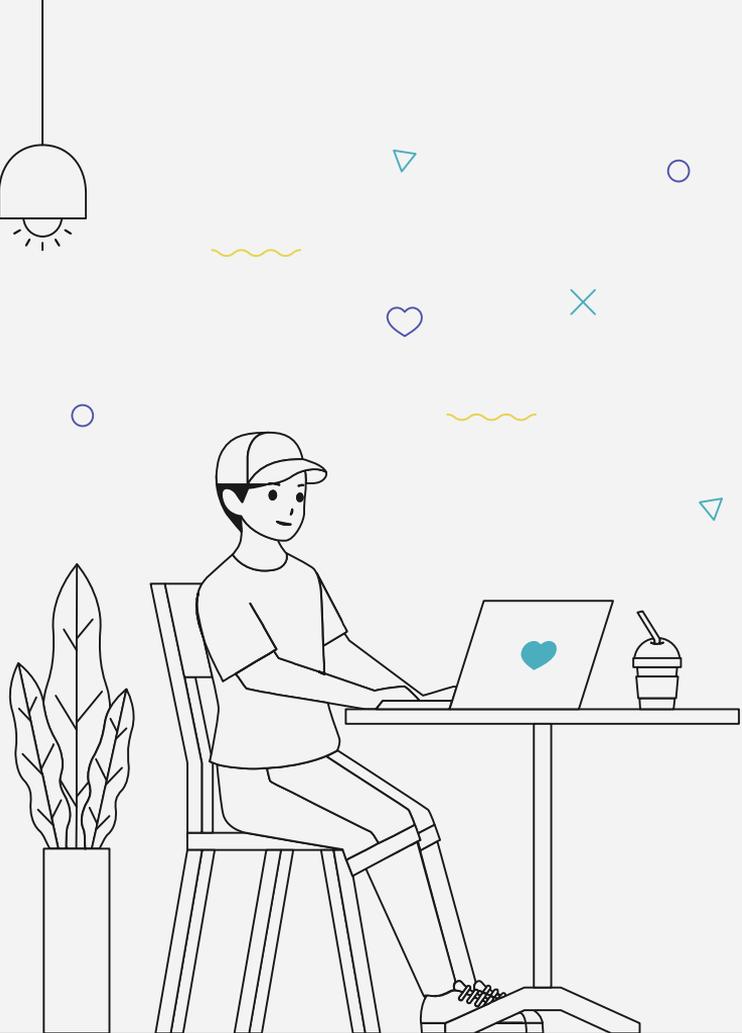
Need AI-based Automatic Speaking Assessment(ASA) system



02

Designing a task-based English speaking diagnosis and learning system using AI technology and automated assessment

Jin-Hwa Lee (Chung-Ang University)



Key features of an English speaking diagnosis and learning system



Assessment for learning

should diagnose the learner's speaking ability, and provide guidance for subsequent learning



Curriculum alignment

should align with the national curriculum



Task-based learning

should provide speaking tasks relevant to real world for EFL learners



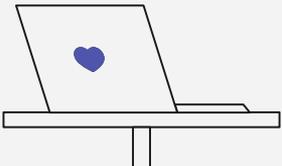
Automated scoring

should provide reliable and time-efficient scoring.



Applicable

should be easy for teachers and learners to access and use



The procedures of the project

01

Test development

Developing a speaking test and selecting scoring features

02

System development

Developing a diagnosis and learning system

03

System implementation

Implementing the system and collecting data

04

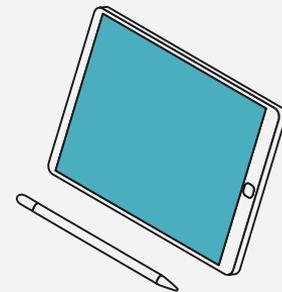
Validation

Validating the automated scoring and checking the system operation

05

System improvement

Improving the system performance

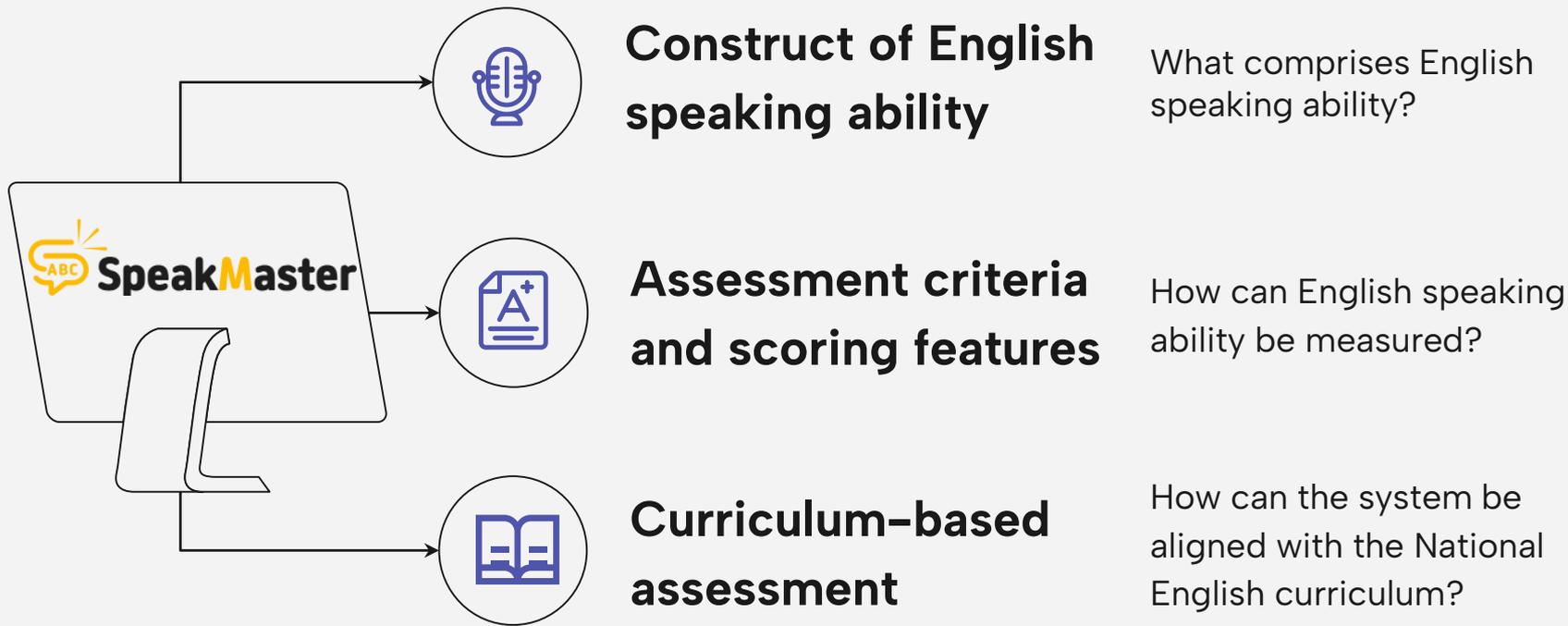


**Stage 1
Test
Development**





Issues in developing an English speaking test for Korean primary and secondary schools are ...



To solve the issues,

Literature review
of the papers on
L2 English
speaking and
testing

Review of CEFR
and NCSSFL-
ACTFL Can-Do
Statements

Review of
existing English
speaking
proficiency tests
(Lee et al., 2023)

Analysis of the
2022 Revised
English
curriculum of
Korea

TABLE 1

English Speaking Tests for Analysis

	Tests	Institutions	Country
ASS*	TOEFL® Practice Online	ETS	USA
	Duolingo English Test	Duolingo	USA
	EnglishScore Speaking Test	British Council	England
	Linguaskill General Speaking Test	Cambridge Assessment English	England
	PTE Academic	Pearson	England
	Versant English Test	Pearson	England
Non-ASS	TEPS	Seoul National University Language Institute	Korea
	NEAT (Levels 2 & 3)	Korea Institute of Curriculum and Evaluation (KICE)	Korea
	GEPT	Taiwan Language Training & Testing Center (LITC)	Taiwan
	TOEIC	ETS	USA
	TOEFL Junior	ETS	USA
	TOEFL iBT	ETS	USA

Note. * Automatized Speaking Scoring

Issue 1. The construct of English speaking ability



Imitative Speaking

Listen & repeat,
Read aloud



Interaction

Conversation



Presentation

Relatively long turns,

Table 3 – Macro-functional basis of CEFR categories for communication

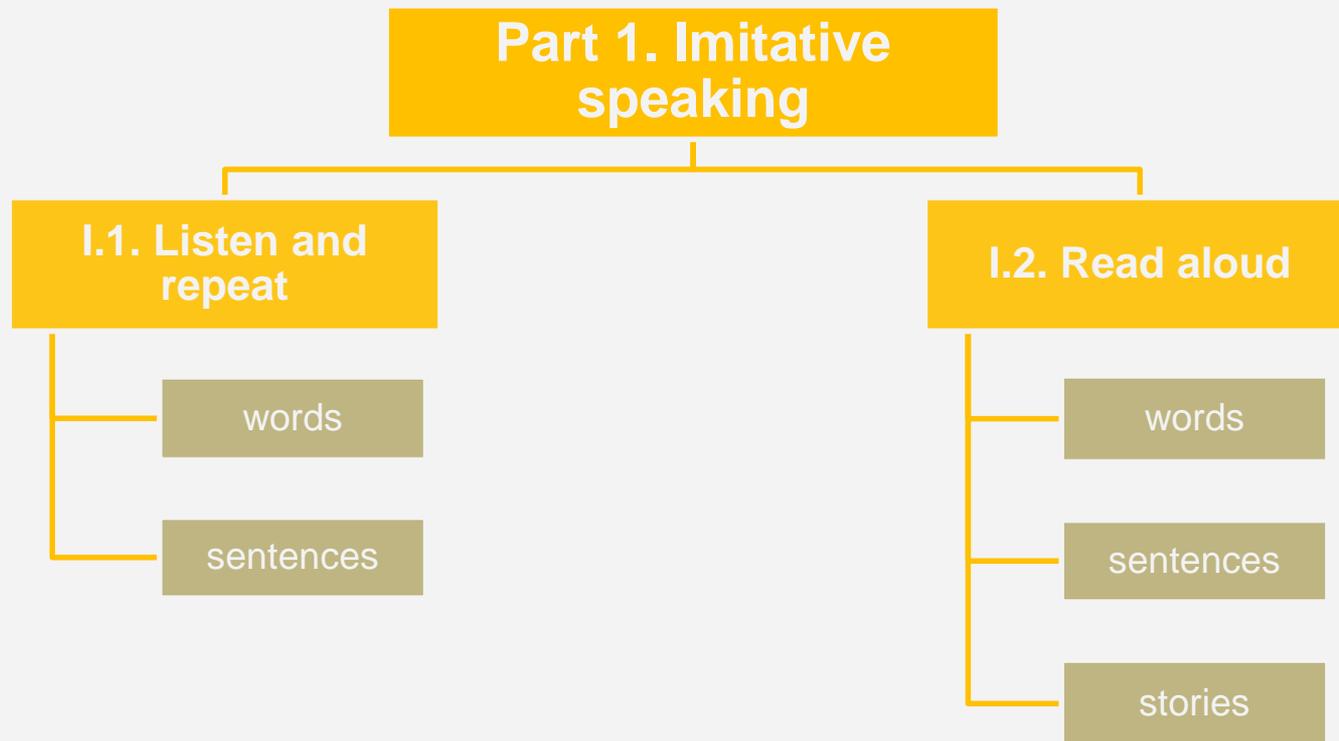
	Reception	Production
Creative, interpersonal language use	e.g. Reading as a leisure activity	e.g. Sustained monologue: describing experience
Transactional language use	e.g. Reading for information and argument	e.g. Sustained monologue: giving information
Evaluative, problem-solving language use	(merged with Reading for information and argument)	e.g. Sustained monologue: presenting a case (e.g. in a debate)

TABLE 3
Test Item Types in Automated vs. Non-Automated Speaking Scoring Tests

Type	Tests	Sounds & Forms			Conversation			Discourse		Integration	
		Listen & Repeat	Read Aloud	Order Phrases	Listen & Answer to Closed Questions	Listen & Answer to Open Questions	Answer Using Information Provided	Describe Pictures/Images	Tell a Story with Pictures	Extended Talk	Integrated Speaking
ASS*	TOEFL® Practice Online									V	V
	Duolingo English Test		V					V		V	
	EnglishScore Speaking	V	V				V				
	Linguaskill General		V			V	V			V	
	PIE Academic	V	V		V			V			V
	Versant English Test	V	V	V	V						V
Frequency		3	5	1	2		2	0		3	3
Non-ASS	IELTS					V		V	V	V	
	NEAT Level 2			V		V		V	V	V	
	NEAT Level 3					V	V	V	V	V	
	GEPT Elementary	V	V			V					
	GEPT Intermediate		V			V		V			
	GEPT High-Intermediate					V		V		V	
TOEIC			V			V	V	V		V	
TOEFL Junior			V						V		V
TOEFL iBT										V	V
Frequency		1	4(5)**	0	0	4(7)	2	4(5)	3(4)	5(6)	2

Note. *Automatized Speaking Scoring; **number of tests (number of tests counting sub-level tests of NEAT and GEPT)

Part 1. Imitative speaking



Part 2. Interaction

Part II. Interaction

II.1. Conversation

II.1.1. Everyday conversation

Greeting & talking about oneself

Talking about daily life
(e.g., daily schedule, hobby, weekend plans)

Requesting & responding
(e.g., invitation, suggestion, help, advice, approval)

II.1.2. Exchanging information/opinion

Exchanging information about surroundings
(e.g., people, things, places)

Exchanging opinions

II.2. Obtaining goods/services

II.2.1. Obtaining goods

Ordering food

Purchasing goods

Purchasing tickets

II.2.2. Obtaining services

Making reservations

Using transportations

Using an airport

Using hotels

Part 3. Presentation

Part III. Presentation

III.1. Describing/Narrating

III.1.1. Introducing

Introducing oneself

Introducing others

III.1.2. Describing

Describing people/things

Describing places/scenes

III.1.3. Narrating

Narrating experiences

Narrating plans

Narrating daily routines

Storytelling

III.2. Delivering information

III.2.1. Delivering information based on visual materials

Explaining tables and charts

Explaining signs and posts

III.2.2. Explaining relationship

Explaining methods

Explaining cause and effect

Explaining logical relationship

III.3. Expressing thoughts

III.3.1. Telling feelings/opinions

Expressing emotions

Expressing preferences

Expressing opinions

III.3.2. Making suggestions

Suggesting solutions

Issue 2. Assessment criteria & scoring features

Part	Criteria	Scoring features	scoring methods
I. Imitative speaking	Flow	Speed	WPM
		Chunking	Number and location of pauses
	Pronunciation accuracy	Phonemes, stress, intonation	AI-based assessment
II. Interaction	Interactional appropriateness	Speed	Time taken to complete a task
		Turn-taking	Number of turns for a task
	Task completion	Completion rate	Ratio of task completion
III. Presentation	Fluency	Speed	WPM
		Vocabulary amount	Repetition
	Grammar	Structural diversity	ASC (Argument structure construction) types, POS (part of speech) types
		Structural accuracy	Ration of error-free sentences
	Vocabulary	Vocabulary difficulty	In reference to VP Kids 250 levels and the National English curriculum vocabulary list
		Vocabulary diversity	TTR and Word types in reference to sample responses
Content	Keyword matching	Expected keywords	

2.1.2. Production

Category	Big Ideas	<ul style="list-style-type: none"> Fosters the ability to express one's emotions or opinions through spoken or written discourses using appropriate media according to communication purposes and situations. Expresses and exchanges meaning using appropriate thinking processes and strategies. Activities involving expressing through spoken or written discourses cultivate an attitude of co-operation and inclusivity in communication, based on an understanding of diverse cultures and perspectives, for expressing or exchanging meaning. 		
	Content Elements	Elementary School		Middle School
		Grades 3-4	Grades 5-6	Grades 1-3
Knowledge & Understanding	Language	<ul style="list-style-type: none"> Sounds, spellings, stress, rhythm, intonation of easy and simple words, phrases, and sentences [Appendix 2] Communicative Functions and Example Sentences [Appendix 3] Within 300 words based on the Guidelines for Basic Vocabulary [Appendix 4] Language forms recommended for use in elementary school 	<ul style="list-style-type: none"> Stress, rhythm, intonation of simple words, phrases, and sentences [Appendix 2] Communicative Functions and Example Sentences [Appendix 3] Within 300 words based on the Guidelines for Basic Vocabulary [Appendix 4] Language forms recommended for use in elementary school 	<ul style="list-style-type: none"> Stress, rhythm, intonation, linking or contraction in words and sentences [Appendix 2] Communicative Functions and Example Sentences [Appendix 3] Within 1,500 words based on the Guidelines for Basic Vocabulary [Appendix 4] Language forms recommended for use in middle school
		<ul style="list-style-type: none"> Stories, narratives, and prose (fairy tales, picture books, comics, songs, poems, etc.) Spoken and written discourses for interpersonal and social purposes (dialogues, letters, emails, etc.) Spoken and written discourses for informational delivery and exchange (signs, memos, descriptions, instructions, etc.) 	<ul style="list-style-type: none"> Stories, narratives, and prose (fairy tales, picture books, comics, songs, poems, diaries, etc.) Spoken and written discourses for interpersonal and social purposes (dialogues, letters, invitations, emails, etc.) Spoken and written discourses for informational delivery and exchange (announcements, instructions, descriptions, explanations, advertisements, etc.) Spoken and written discourses for conveying and exchanging opinions or making arguments (posters, reviews, etc.) 	<ul style="list-style-type: none"> Stories, narratives, and prose (diaries, essays, travel journals, novels, scripts, song lyrics, poems, etc.) Spoken and written discourses for interpersonal and social purposes (letters, emails, social media, etc.) Spoken and written discourses for informational delivery and exchange (announcements, instructions, explanations, reports, articles, advertisements, etc.) Spoken and written discourses for conveying and exchanging opinions or making arguments (editorials, articles, speeches, etc.)
	Context	<ul style="list-style-type: none"> [Appendix 1] Themes Topics around oneself Simple communication situations and purposes Non-verbal communication styles of people from diverse cultural backgrounds 	<ul style="list-style-type: none"> [Appendix 1] Themes Everyday topics Daily communication situations and purposes Verbal and non-verbal communication styles of people from diverse cultural backgrounds 	<ul style="list-style-type: none"> [Appendix 1] Themes Familiar topics Various communication situations and purposes Verbal and non-verbal communication styles of people from diverse cultural backgrounds

National English curriculum

Control

Contents

Curriculum
Vocabulary
List and
of words

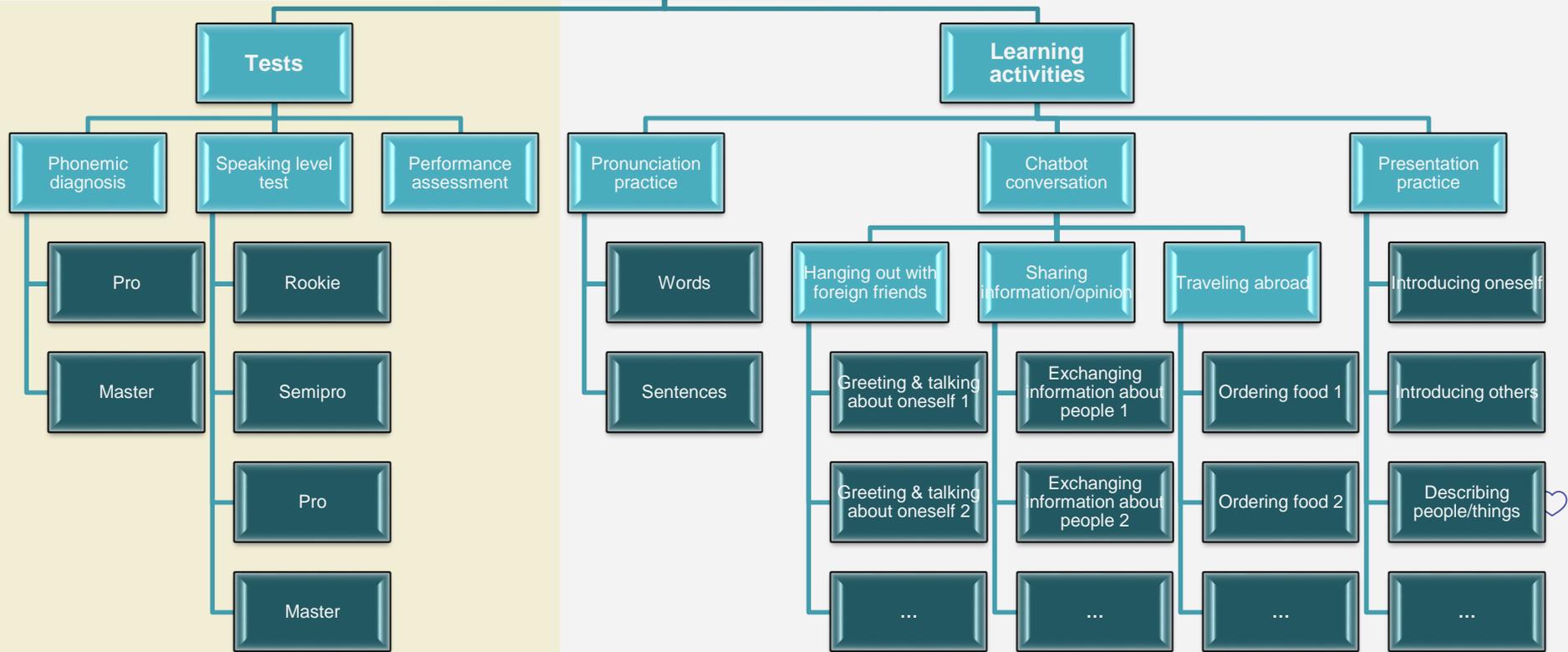
The National curriculum defines the content to be covered in English subject

Process & Skills	<ul style="list-style-type: none"> Listening and repeating in accordance with stress, rhythm, and intonation Writing in both uppercase and lowercase letters Writing words based on the relationship between sound and spelling Describing real objects, pictures, actions, etc., through speaking or writing Greeting others Introducing oneself Describing people or objects in the surroundings Giving instructions Expressing emotions Asking or answering important information Expressing creatively with facial expressions, gestures, and actions Expressing creatively using appropriate media Checking spelling and rewriting as needed 	<ul style="list-style-type: none"> Speaking with appropriate stress, rhythm, and intonation Speaking or writing based on real objects, pictures, actions, etc. Correctly using uppercase and lowercase letters and punctuation marks Introducing people in the surroundings Describing people or objects in the surroundings Explaining nearby places, locations, action sequences, or methods Expressing emotions, opinions, experiences, or plans Asking or answering specific information Writing in accordance with the purpose, referring to example sentences. Sustaining a conversation through repetition, verification, etc. Generating ideas through brainstorming Expressing creatively using various media Revising based on feedback 	<ul style="list-style-type: none"> Using assimilated and contracted sounds Describing people, objects, etc. Describing moods or emotions Explaining locations or places, etc. Describing pictures, photos, charts, etc. Explaining methods or procedures Describing experiences or plans Explaining the sequence and context of events Explaining causes and effects of events Asserting one's own opinions Listening or reading and summarizing Writing diaries, letters, emails, etc. Speaking or writing using appropriate media Speaking or writing using appropriate strategies Speaking or writing one's creative thoughts
	Values & Attitudes	<ul style="list-style-type: none"> Interest and confidence in speaking and writing Attitude of observing conversational etiquette, cooperating, and participating in communication activities 	<ul style="list-style-type: none"> Attitude of speaking or writing in consideration of others Attitude of speaking or writing in accordance with information ethics

A young boy with brown hair, wearing a dark blue long-sleeved shirt and tan pants, is climbing a green and red playground structure. He is holding onto a green ring with both hands and has one leg extended outwards. The background shows a large brick school building with many windows. A large white circle is overlaid on the left side of the image, containing the text.

Stage 2
Developing a
diagnosis and
learning
system

Architecture of the system



[Tests] 1. Pronunciation Diagnosis

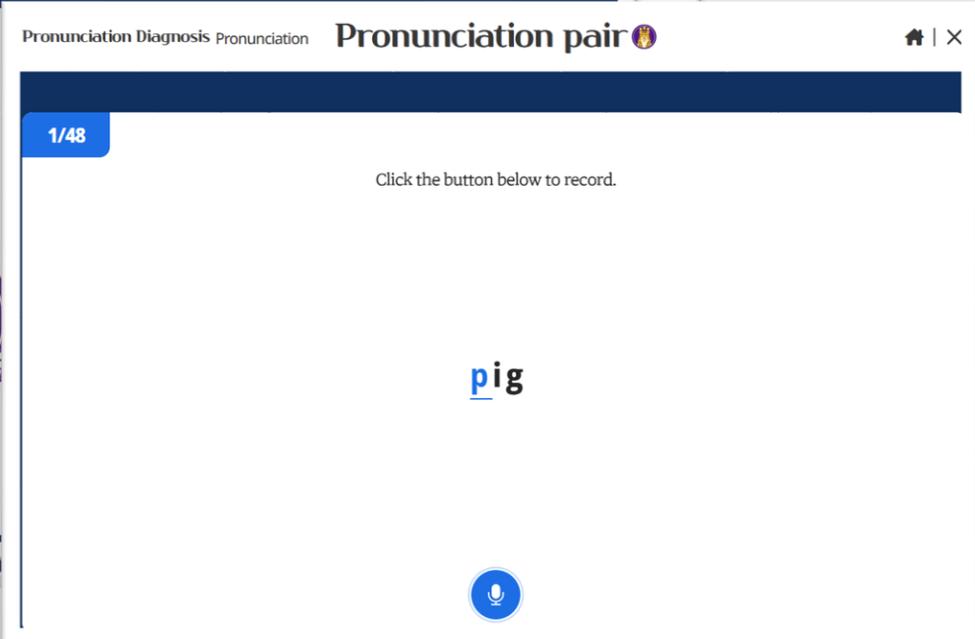
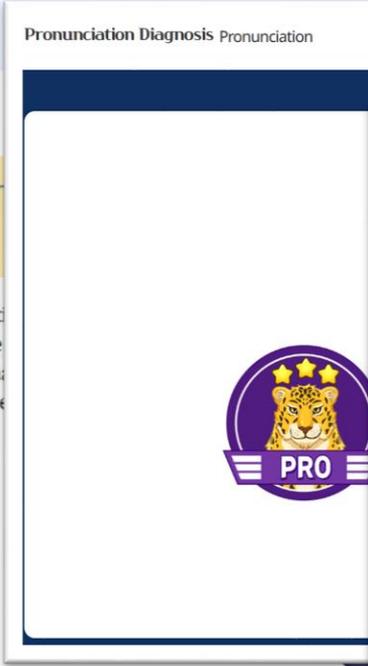


SpeakMaster test activity log out My record

test As I engage in realistic conversations and presentations, my confidence grows and I can speak English anytime, anywhere.

Pronunciation Diagnosis

Diagnose whether your pronunciation is understandable to others. Analyze your recordings and we will provide a pronunciation diagnosis chart that shows areas that need improvement.



Pronunciation pair

1/48

Click the button below to record.

pig



[Tests] 2. Speaking Level Tests

SpeakMaster test activity log out My record

Speaking Level Test Speaking

test As I engage in realistic conversat

Pronunciation Diagnosis

Diagnose whether your pronunciation is understandable to others. Analyze and provide a pronunciation diagnosis chart that shows areas that need improvement.

Assess your speaking skills.

Speaking Level Test Speaking **guide**

Structure of the questions

This test consists of three parts, totaling 21 questions (20 for rookies). In each part, you will perform different speaking tasks as follows.

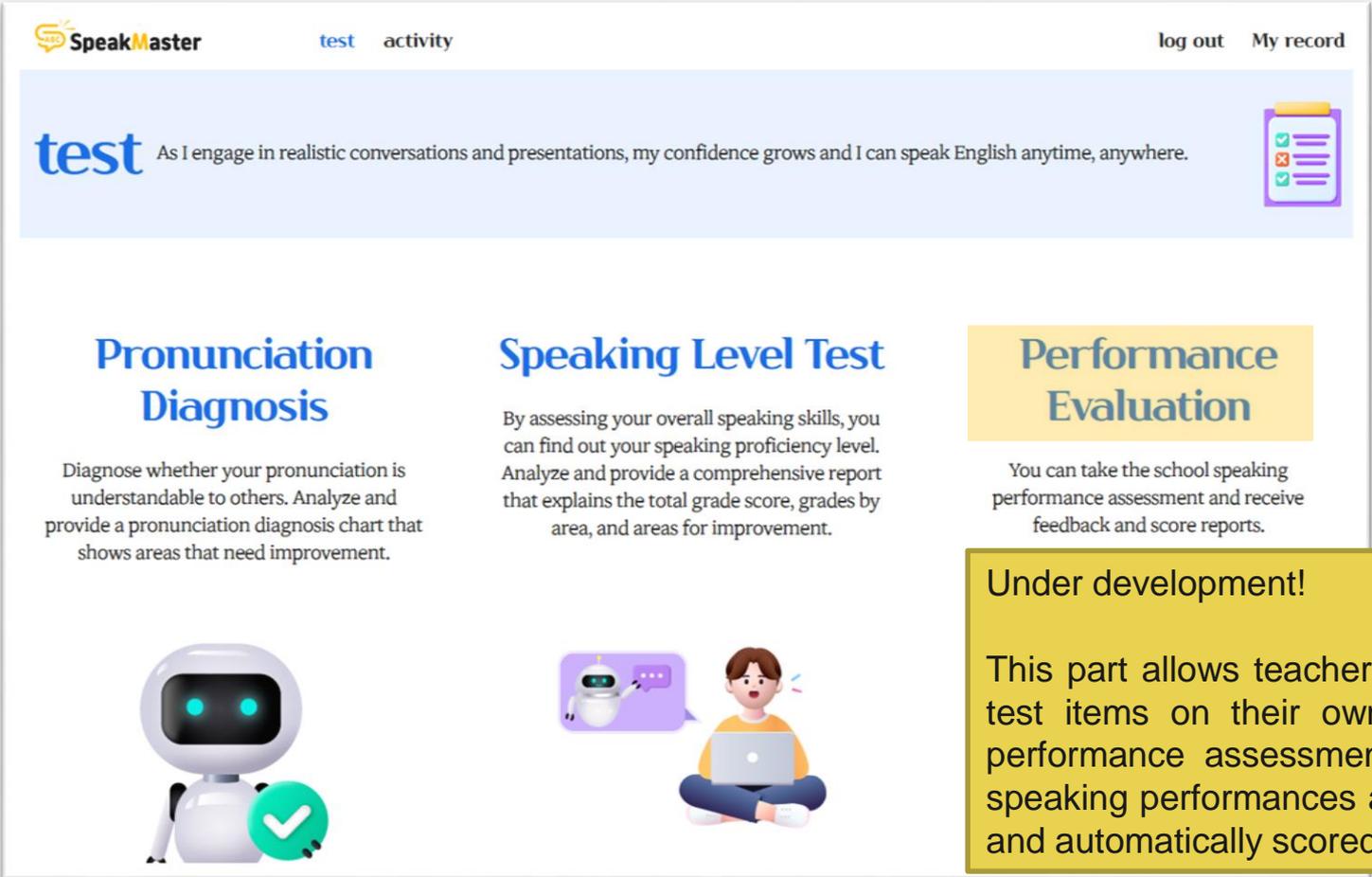
Part 1. Pronunciation
(Questions 14-15)

Part 2. Conversation
(4 questions)

Part 3. Presentation
(2 questions)

NEXT >

[Tests] 3. Performance Assessment



SpeakMaster test activity log out My record

test As I engage in realistic conversations and presentations, my confidence grows and I can speak English anytime, anywhere.

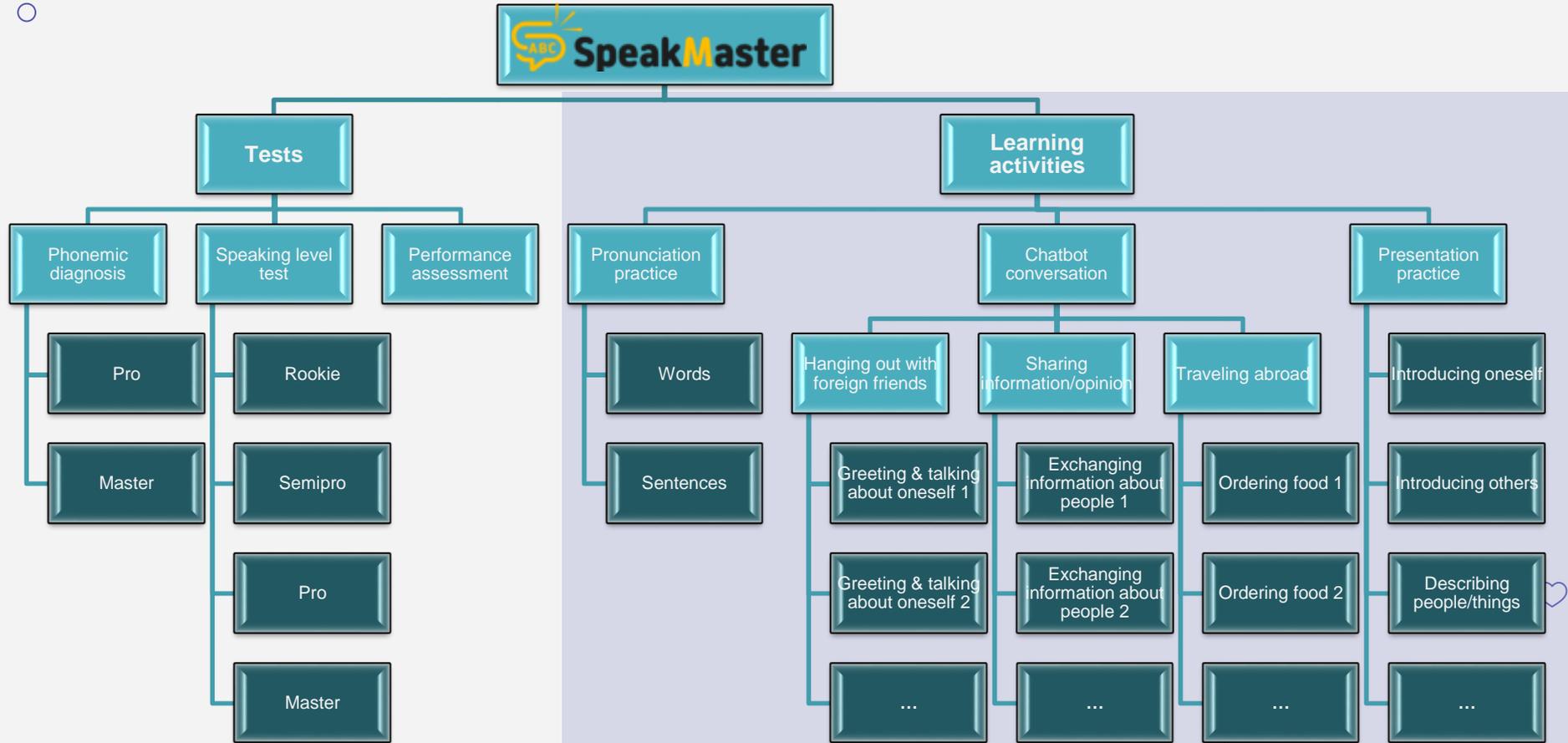
Pronunciation Diagnosis
Diagnose whether your pronunciation is understandable to others. Analyze and provide a pronunciation diagnosis chart that shows areas that need improvement.

Speaking Level Test
By assessing your overall speaking skills, you can find out your speaking proficiency level. Analyze and provide a comprehensive report that explains the total grade score, grades by area, and areas for improvement.

Performance Evaluation
You can take the school speaking performance assessment and receive feedback and score reports.

Under development!
This part allows teachers to develop test items on their own for official performance assessment. Students' speaking performances are recorded and automatically scored.

Architecture of the system



[Learning Activities] 1. Pronunciation Practice

pronunciation practice Pronunciation **Listen and follow along** X

log out My record

After choosing the area

pronunciation practice Pronunciation | Listen and follow **word** 🗑️

1. C

2. When you hear a "ding"

3. When the rec...

1/50

pronunciation practice Pronunciation | Listen and follow **word** 🗑️

1/50

Good

add

add

67% 85%

[Learning Activities] 2. Chatbot Conversation

SpeakMaster test activity log out My record

Chatbot conversation Chatbot Activity type X

rows and I can speak English anytime, anywhere.

Chatbot conversation Chatbot Talk X

Chatbot conversation Chatbot X

speakmaster.co.kr 내용:
미션이 종료되었습니다.
소요시간: 31초
유지점수: 3
획득가능: 0

확인

미션
On the first day of the semester, you will meet your native English teacher in the hallway. Say hello to your teacher and introduce yourself as the teacher asks you questions.

procedure
1. Say hello
2. Answer the teacher's questions

Very good. I remember you. What was your name?
I'm Jean.

Hi, Jean. What grade are you in? Grade 3? Grade 4?
I'm in grade four.

Great! I will see you in the English class. Bye.



When hanging out with foreign friends



[Learning Activities] 3. Presentation Practice

Presentation practice Presentation

Activity type

introducing people around

Presentation practice Presentation

introducing people

Presentation practice Presentation

미션

Introduce my best friend, Emily, who tran

Friend's name , #The Advantag

REC 00:00:2

Presentation practice Presentation

Presentation Practice Report

Teaching student

Test School 7 | Exam Date: November 16, 2024 | time taken: 02 minutes 33 seconds

My presentation

Hi Emily, this is my best friend Suji. Suji is very kind. She is very good at singing. You may like her voice very well. I hope you will make good friends each other.

Sample answer

Click the button to hear a sample answer.

You can check the text by clicking the button.

Speech Analysis

Fluency

Category	me	Sample Answer
속도	142wpm	126wpm
어휘량	5	5

Grammar

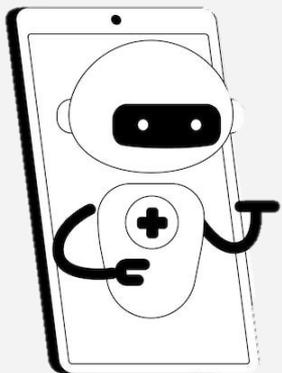
Category	me	Sample Answer
구문 다양성	1	1
문사 다양성	11	6
문법 오류	8	0

Vocabulary

Category	me	Sample Answer
난이도	10	8
다양성	13	12

Content

Category	me	Sample Answer
관련 단어 빈도	3	5
관련 단어 분포	10	12



03

Fostering and Assessing L2 Speakers' Oral Interaction Ability with AI Chatbot

Heyoung Kim (Chung-Ang University)



Interpersonal interaction

- Since speaking usually takes place in conversation, valid speaking test must measure conversation competency
- However, conversation is more transient, context-dependent, and dynamic, and unpredictable than unidirectional speech
- Assessing conversation ability even more difficult than monologue speech to be assessed due to tasks, raters, and interlocutors (de Jong, 2023, Fulcher, 2015, McNamara, 1996)
- Most standardized speaking tests do not include conversation (Lee et al. 2023)

Tacit class rule in Korea "No test No class activity"



Scope of ASS Technology Use in English Speaking Tests (Lee et al, 2023, p.230)

Tests	Test Items	Criteria
TOEFL® Practice Online	1. Independent speaking 2. Integrated speaking (campus-related, academic)	Some features (forms only) of pronunciation, fluency, vocabulary, and grammar
Duolingo English Test	1. Read aloud	Spoken feature (acoustic and fluency feature)
	2. Extended speaking (picture description, independent speaking with a written prompt, independent speaking with an aural prompt)	Grammatical accuracy, grammatical complexity, lexical sophistication, lexical diversity, task relevance, length, fluency & acoustic features
	3. Speaking sample (interview)	Ungraded
EnglishScore Speaking Test	1. Read aloud	Pronunciation (sound and stress), fluency
	2. Listen and repeat	
	3. Guided task (open-ended task with visual/textual prompts)	Pronunciation, fluency Communication skills
Linguaskill General Speaking Test	1. Interview (questions about oneself)	Pronunciation, fluency, language resource
	2. Reading aloud	Individual sound, stress, rhythm, and intonation
	3. Presentation (on a given topic)	Pronunciation, fluency, language resource, discourse management, task appropriateness
	4. Presentation with visual information	
	5. Communication activity (opinion-focused questions related to a scenario)	
PTE Academic	1. Personal introduction	Ungraded
	2. Read aloud	Content (form (word) correctness), oral fluency, pronunciation
	3. Repeat sentence	
	4. Describe image	
	5. Re-tell lecture	
	6. Answer short question (in a word)	
Versant English Test	1. Reading (read aloud)	Correct/incorrect
	2. Repeat (listen & repeat a sentence)	Pronunciation and reading fluency Mastery of sentence structure, pronunciation and fluency
	3. Short questions (listen a short question and answer with a single word or a short phrase)	Vocabulary
	4. Sentence builds (listen to three short phrases and rearrange them)	Mastery of sentence structure, pronunciation and fluency
	5. Story retelling (listen brief stories and retell them in their own words)	Mastery of sentence structure, pronunciation, vocabulary and fluency
	6. Open Question	Ungraded



Test item Types in Automated vs. Non-Automated Speaking Scoring Tests (Lee et al, 2023, p.233)

Type	Tests	Sounds & Forms			Conversation			Discourse			Integration
		Listen & Repeat	Read Aloud	Order Phrases	Listen & Answer to Closed Questions	Listen & Answer to Open Questions	Answer Using Information Provided	Describe Pictures/Images	Tell a Story with Pictures	Extended Talk	Integrated Speaking
ASS*	TOEFL® Practice Online									V	V
	Duolingo English Test		V					V		V	
	EnglishScore Speaking	V	V				V				
	Linguaskill General		V			V	V			V	
	PIE Academic	V	V		V			V			V
	Versant English Test	V	V	V	V						V
	Frequency	3	5	1	2	1	2	2	0	3	3
Non-ASS	TEPS		V			V		V	V	V	
	NEAT Level 2					V		V	V	V	
	NEAT Level 3					V	V		V	V	
	GEPT Elementary	V	V			V					
	GEPT Intermediate		V			V		V			
	GEPT High-Intermediate					V		V		V	
	TOEIC		V			V	V	V		V	
	TOEFL Junior		V						V		V
	TOEFL iBT									V	V
	Frequency	1	4(5)**	0	0	4(7)	2	4(5)	3(4)	5(6)	2

Note. *Automatized Speaking Scoring; **number of tests (number of tests counting sub-level tests of NEAT and GEPT)

Assessing interpersonal communication: Challenge

What to measure L2 speaker's

- cognitive skills to **produce** proper language rapidly in appropriate time length between turns of speakers (Bates et al. 2004; Pickering & Garrod 2013)
- cognitive skills to rapidly **recognize** a task, **predict and anticipate** the current incoming message in a conversation (Levinson 2016)
- Competency to successfully complete English conversation tasks, *target tasks* that very likely occur to (Korean) L2 speakers in their life and society(*for social purposes*)

Psycholinguistic
approach to speaking

Task-Based Language
Teaching approach to
speaking

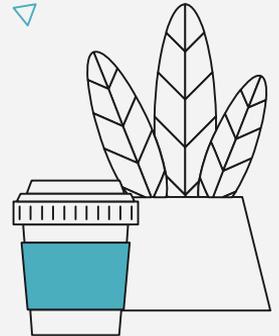
How to measure L2 speakers' conversation

- **conversation performance** The number of turns and time in conversation
- **task success rate** the participant's recognition, prediction and production in conversation



Voice Chatbot

- AI interlocutor
- Can assign context dependent roles.
- L2 learners' preference for chatbot over human partners (Fryer et al., 2020).
- L2 target task-based chats (Kim et al. 2021; Kim & Lee, 2022)
 - More than pattern drill or scripted role play
 - goal oriented tasks that require cognitive skills
 - Interlanguage, level-adjusted language
 - culturally sensitive, needs-identified topics
- Voice recognition (STT) & Voice synthesis (TTS) technology



L2 Chatbot

Language Learning needs

Intriguing topic, FAQs

Interlanguage

Age & cognition level

Easy language

Ice breaking strategies

Friendly human persona



ChatGPT? Not yet



Closed domain Chatbot, Task-based chatbot

- Call center, customer service bot, and other commercially developed human Assistant bot.
 - more goal-oriented and functional
 - predefined domain knowledge, topics, and contexts for conversation
 - Higher recognition rate than in open domain chatbots under the same conditions
- (Shum et al., 2018).



The previous chatbot studies

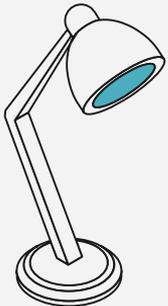


L1 Chats

- ALICE, and Kuki, Cleverbot, Amazon Alexa, Google Assistant, Apple Siri
- designed for L1 users' general use, but implemented in a language curriculum (e.g., Jia & Ruan, 2008; Lee et al., 2020, Kim, 2018; Shin, Kim, Lee, & Yang, 2021).

L2 chatbots

- CSIEC (Jia & Chen, 2008; Jia & Ruan, 2008), Dave ESLbot, or Lucy (Wang & Petrina, 2013), GenieTutor (Choi, Kwon, Lee, Roh, Huang & Kim, 2017), Mondly (see Fryer et al., 2020).
- enabled English learners to practice diverse expressions on given topics at a restaurant or a shopping mall, among others.
- close to form-focused dialogues.
- have centered more on pattern drills in certain situation
- limited opportunities to practice diverse meaningful exchanges in target languages
- not fully able to generate follow-up questions or relevant, natural responses to develop continuous conversations (Gallacher, Thompson, and Howarth, 2018)



Talk with Ellie

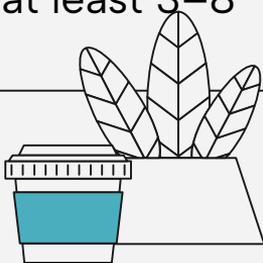
<https://chatbot.smartteaching.or.kr/>

(Yang, Kim, Lee & Shin, 2022 in *ReCALL*)

(Kim, Lee, Yang, & Shin, 2022, in *Language Learning and Technology*)



- **THE FIRST task-based L2 chatbot**
- predefined as a female native English speaker and EFL teacher living in South Korea.
- 71 tasks "Mission Challenge".
- Three levels sorted into task complexity levels,
- expected to exchange conversation turns at least 3-8 times for task



Guide (type 3-8)
Please keep Google Chrome up to date

Order burgers and drinks

Look at the menu and order a burger and a drink one by one. The three missions you need to do are:

- ✓ 1. Looking at the menu and choosing a hamburger
- ✓ 2. Understanding and answering a clerk's questions
- ✓ 3. Choosing a drink and a payment method

Burger Menu	Drink menu
1. Big burger	1. Coke
2. Cheese burger	2. Sprite
3. Shrimp burger	

English chatbot Ellie

Ellie
Look at the menu and order a burger and a drink one by one.

Ellie
The three missions you need to do are: 1) **Choosing a hamburger from the menu** 2) **Understanding and answering a clerk's questions** 3) **Choosing a drink and payment method**

Ellie
When you're ready, hit the [Click and Speak to Ellie!] button, say "Hello" and start the conversation.

Click and speak to Ellie

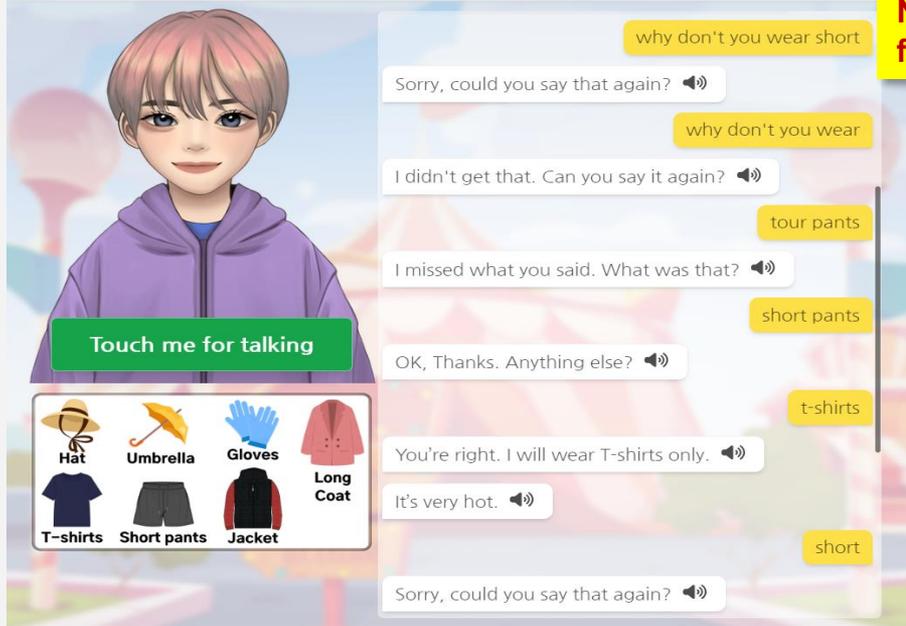
Yang et al (2022) “Implementation of an AI chatbot as an English conversation partner in EFL speaking classes (ReCALL)”

- 314 Korean EFL learners aged 10–15 years performed three speaking tasks with Ellie in their school classroom.
- The participants took **9.63 turns per session** on average using the first 1000-word band, indicating that the chatbot highly encouraged students to engage in conversation, which rarely occurs in general EFL classes in Korea.
- **The high task success rates (88.3%)** showed the design appropriateness of both L2 tasks and operational intents in terms of users’ successful understanding and completeness of the given chatbot tasks.

Kim and Lee (2023) “Task-based vs. scripted role play chatbot (MALL)”

- Six elementary third-graders participated in four different chatbot-based speaking sessions designed in two types.
- **More varied meaning-negotiations** were observed in the ‘task chatbot’ than the role-play ones.
- **The young learners recalled more words** in the delayed test and **showed more interest** and preference in the ‘task chatbot’.

Kim and Lee (2023) Kai Project “task-based smart mirror chatbot for young learners”



Negative feedback

Modified output

**Cognition process
Freedom to choose.**



<https://kai.smartteaching.or.kr/>

Some degree of freedom in the chatbot-based tasks, meaning-negotiation, which involves a divergence from scripted dialogue or situational drills

activity

As I engage in realistic conversations and presentations, my confidence grows and I can speak English anytime, anywhere.



pronunciation practice

Accurate speech is very important in communication. Practice is essential for accurate pronunciation! Improve your pronunciation by receiving feedback from AI.



Chatbot conversation

Depending on the situation, you can have conversations with chatbots and have a smooth conversation. Try challenging English usage missions that you will encounter in real life.



Presentation practice

True speaking skills lie in the ability to present! Build your confidence in speaking through presentation practice in a variety of situations.



Select the chatbot activity you would like to practice.

Click on the color that corresponds to your level.

When hanging out with foreign friends

Greetings, Introduction, and Greetings I	Greetings, Introductions, and Greetings II	answering questions about things around you I	answering questions about things around you II
Invite/Suggest and Respond I	Invite/Suggest and Respond II	Ask for help and answer I	Asking for Help and Responding II
Asking for and answering advice I	Asking and Responding to Advice II	Ask for permission and respond	
answering questions about people around you I	answering questions about people around you II	answering questions about things around you I	answering questions about things around you II
answering questions about people around you I	answering questions about people around you II	answering questions about things around you I	answering questions about things around you II

When sharing information or opinions in English

questions about
places around
you I



answering
questions about
places around
you II



answering
thoughts and
opinions I



answering
thoughts and
opinions II



While traveling abroad

Ordering Food I



Ordering Food II



Shopping for
stuff I



Shopping for
Things II



Buy Tickets I



Buy Tickets II



Make a
reservation



Using
transportation



Using the Airport



Use the hotel



Part 2. Interaction

Part II. Interaction

II.1. Conversation

II.2. Obtaining goods/services

II.1.1. Everyday conversation

Greeting & talking about oneself

Talking about daily life (e.g., daily schedule, hobby, weekend plans)

Requesting & responding (e.g., invitation, suggestion, help, advice, approval)

II.1.2. Exchanging information/opinion

Exchanging information about surroundings (e.g., people, things, places)

Exchanging opinions

II.2.1. Obtaining goods

Ordering food

Purchasing goods

Purchasing tickets

II.2.2. Obtaining services

Making reservations

Using transportations

Using an airport

Using hotels

Structure of the questions

This test consists of three parts, totaling 21 questions (20 for rookies). In each part, there are four different speaking tasks as follows.



Part 1. Pronunciation

(Questions 14-15)



Part 2. Conversation

(4 questions)



Part 3. Presentation

(2 questions)

Four
conversation
tasks

NEXT >

II.1. Conversation

Conversation Task goals

II.1.1. Everyday conversation

Greeting & talking about oneself

Talking about daily life (e.g., daily schedule, hobby, weekend plans)

Requesting & responding (e.g., invitation, suggestion, help, advice, approval)

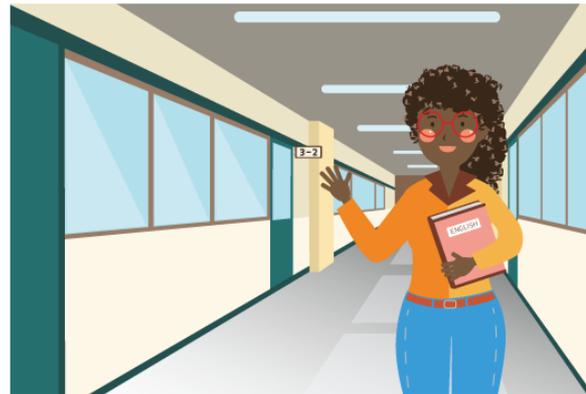
Level 2

Mission

On the first day of the semester, you will meet your native English teacher in the hallway. Say hello to your teacher and introduce yourself as the teacher asks you questions.

procedure

1. Say hello
2. Answer the teacher's questions



Procedure (two sub goals)

START >



II.1. Conversation

Conversation Task goals

Level 4

II.1.1. Everyday conversation

Greeting & talking about oneself

Talking about daily life (e.g., daily schedule, hobby, weekend plans)

Requesting & responding (e.g., invitation, suggestion, help, advice, approval)

Mission

You will be assigned to a group with international students who participated in the international camp. Exchange some basic information with the students sitting next to you and get to know them. Don't hesitate to start a conversation by talking to them first and asking questions naturally.

procedure

1. Say "Nice to meet you." first
2. Asking for name/age/hometown
3. Ask if someone likes sports



International Summer Camp



Procedure (three sub goals)

II.1. Conversation

II.1.2. Exchanging information/opinion

Exchanging information about surroundings (e.g., people, things, places)

Exchanging opinions

Conversation Task goals

Level 2

Mission

I lost my water bottle in the gym during English camp. Please go to the lost and found center and look for the water bottle.

procedure

1. Listen to the teacher's question and describe the lost location and the water bottle below.



Procedure (one sub goals)

START >

II.1. Conversation

II.1.2. Exchanging information/opinion

Exchanging information about surroundings (e.g., people, things, places)

Exchanging opinions

Conversation Task goals

Mission

I went to a Thai restaurant with my Thai friend yesterday. Ask your friend the name of the Thai food we ate yesterday.

procedure

1. Start a conversation with "Thank you for bringing me to a Thai restaurant yesterday."
2. Describe the food the arrow is pointing to your friend and find out its name



Procedure (two sub goals)

Level 4



START >

II.2. Obtaining goods/services

Conversation Task goals

Level 1

II.2.1. Obtaining goods

Ordering food

Purchasing goods

Purchasing tickets

Mission

I came to the Disney World merchandise shop in the US and found some socks on special sale. I asked the clerk if they had the design I wanted and my size.

procedure

1. Check the size and shape
2. Answer the clerk's questions



Procedure (two sub goals)

SPECIAL SALE
ONLY \$1.99

1. MICKEY MOUSE



2. SNOOPY



SIZE	2	3	4	5	6	7	8	9
mm	210	215	220	225	230	235	240	245

START >

II.2. Obtaining goods/services

Conversation Task goals

Level 4

II.2.1. Obtaining goods

Ordering food

Purchasing goods

Purchasing tickets

Mission

I'm coming to the Christmas market in Canada to buy some gifts. I'm going to pick out one decoration each for myself and a friend, check the prices, and buy them.

procedure

1. Ask the owner for the price of the items you want
2. Buying things within your budget (budget \$20)



Procedure (two sub goals)

START >



Demo

챗봇 대화 Chatbot

대화하기 

🏠 | ✕

미션

가족과 뉴욕에 왔습니다. 갑자기 오늘 저녁 6시 이후에 일정이 취소되어 브로드웨이 뮤지컬을 보려고 합니다. 박스오피스에 가서 원하는 뮤지컬 티켓을 구매하세요.

절차

1. 오늘 티켓이 있는지 물어보기
2. 6시 이후 티켓 3장 구매하기
3. 가격 물어보고 \$250 이내에서 구매하기



BOX OFFICE



START >

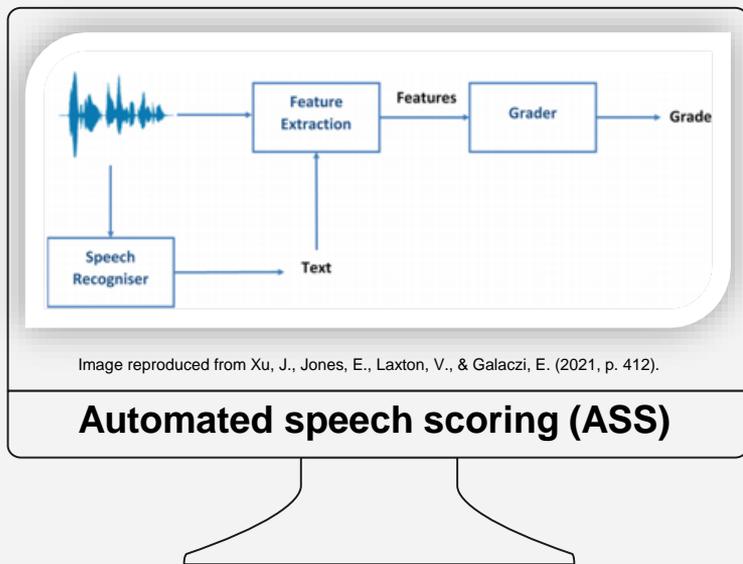
04

Designing Accessible Digital Score Reports: Enhancing Stakeholder Understanding and Educational Opportunities for L2 Speaking



YunDeok Choi (Chungnam National University)

Introduction

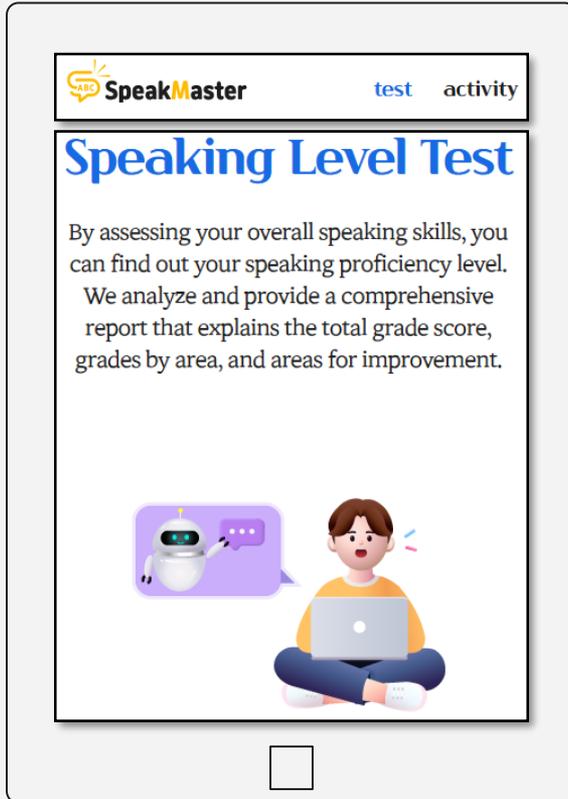


○ Automated speech scoring (ASS)

- Uses computer-assisted methods to evaluate spoken language proficiency (Singla et al., 2021)
- Assesses non-native speakers' speaking proficiency and provides immediate and individualized feedback (Evanini & Zechner, 2019)
- Is increasingly used in large-scale standardized exams (Evanini & Zechner, 2019; Gu et al., 2020; Xu et al., 2021)
 - E.g.) ETS's *SpeechRater*TM, Cambridge English Assessment's CASE v1.9
- Is rarely applied in classrooms, where it could support speaking improvement



Introduction

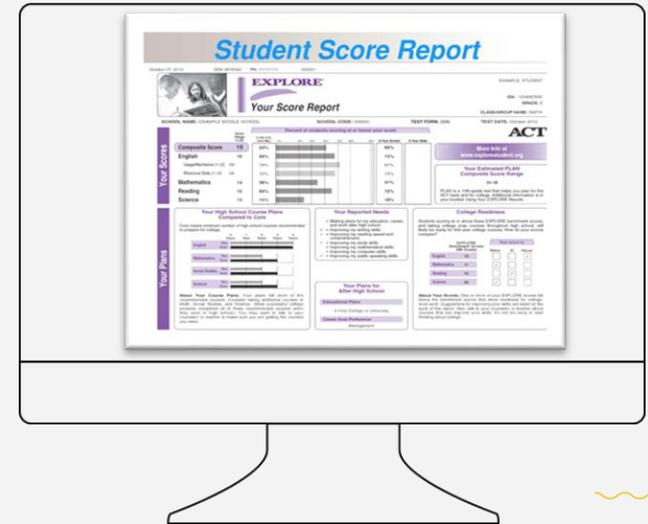


- **SPEAK MASTER's** *Speaking Level Test*

- Incorporates ASS technology tailored for classroom settings
- Diagnoses speaking abilities with detailed, immediate, and individualized feedback
- Features varied tasks beyond constrained tasks with interactive, & spontaneous tasks & provides task-specific feedback
 - E.g.) *Reading aloud*: pronunciation accuracy

Introduction

- Role of score reports & challenges
 - The effectiveness of automated evaluation systems for language learning relies on score reports to convey results accurately (Gu et al., 2020)
 - Stakeholders often struggle to interpret test scores correctly (e.g., Goodman & Hambleton, 2004; Kannan et al., 2018).
- Importance for validity
 - Validity: "the degree to which evidence and theory supports the interpretations of test scores for proposed uses of tests" (AERA, APA, & NCME, 2014)
 - Accurate, easy-to-understand score reports enhance validity by enabling accurate interpretations and minimizing misuse (AERA, APA, & NCME, 2014).



Introduction



- Gaps in current research on score report development
 - Growing research on score reports for
 - Low-stakes assessments (e.g., Brown et al., 2019; Zapata-Rivera et al., 2018; Vezzu et al., 2012)
 - High-stakes standardized proficiency tests (e.g., Ching-Ni Hsieh, 2023; Gu et al., 2020; Gu & Davis, 2019).
 - Limited focus on development processes for automated speaking assessment score reports in classroom contexts
- Research purposes
 - Presenting the ongoing development journey of a prototype score report system for SPEAK MASTER's Speaking Level Test
 - Outlining future steps to complete the development process



Design Process and Achievements:

A static online score report prototype for SPEAK MASTER's Speaking Level Test

25%



50%

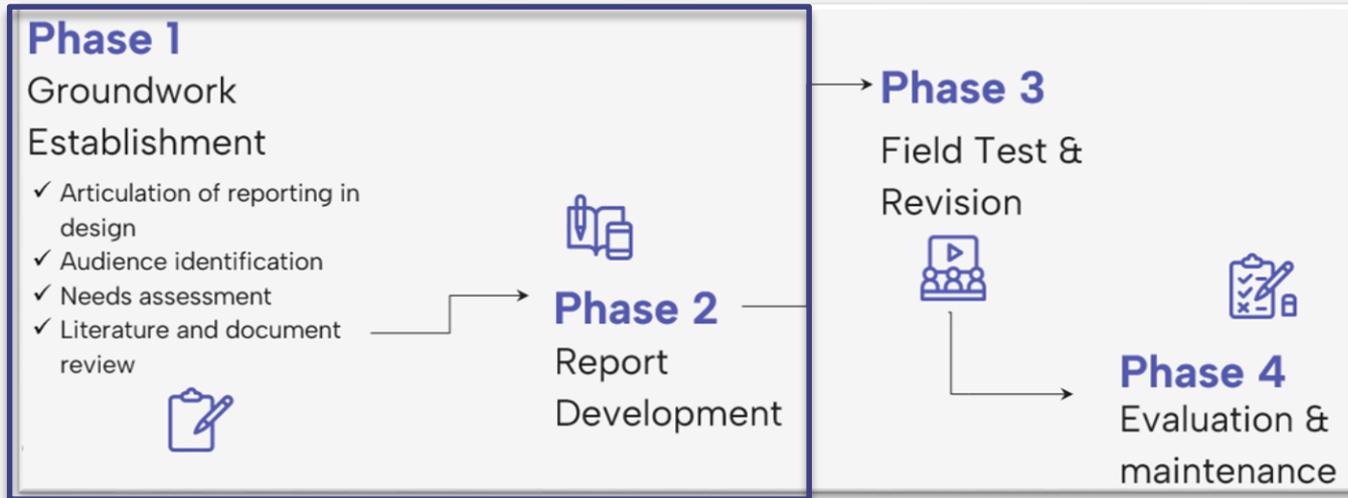


75%



Design Process and Achievements

- **Framework:** Hambleton & Zenisky's (2013) four-phase, seven-step model for score report development.
 - The model provides a systematic approach, integrating score report design into the overall test development process (Zenisky & Hambleton, 2016)



Design Process and Achievements

Phase 1:

Groundwork establishment

- ① **Articulation of reporting in design**
- ② Audience identification
- ③ Needs assessment
- ④ Literature and document review



Test purpose

Identifying strengths and weaknesses in speaking skills through **diagnostic assessment**

Test use

Guiding speaking ability improvement by **providing targeted feedback** for students and teachers

Main sections

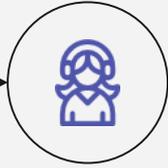
Overall results, domain-specific results, suggestions for improvement

Design Process and Achievements

Phase 1:

Groundwork establishment

- ① Articulation of reporting in design
- ② **Audience identification**
- ③ Needs assessment
- ④ Literature and document review



Students & Parents

EFL students in primary to secondary schools and their parents in South Korea



Educators

English teachers in primary to secondary schools in South Korea, guiding students' English-speaking improvement



Design Process and Achievements

Phase 1:

Groundwork establishment

- ① Articulation of reporting in design
- ② Audience identification
- ③ **Needs assessment**
- ④ Literature and document review



Internal Expert Judgment

Speaking domains (e.g., oral presentation) & targeted **constructs** (e.g., fluency, vocabulary, grammar)

Linguistic features to measure for *usefulness* and *construct coverage* (e.g., WPM) & task types ▾ (e.g., reading aloud)



Design Process and Achievements

Phase 1:

Groundwork establishment

- ① Articulation of reporting in design
- ② Audience identification
- ③ **Needs assessment**
- ④ Literature and document review



Internal Expert Judgment

Reporting methods aligned with the test purpose (e.g., numeric scores, written text feedback, visual graphs) & **score delivery mode** (digital) to ensure easy access & scalability for large-scale use.

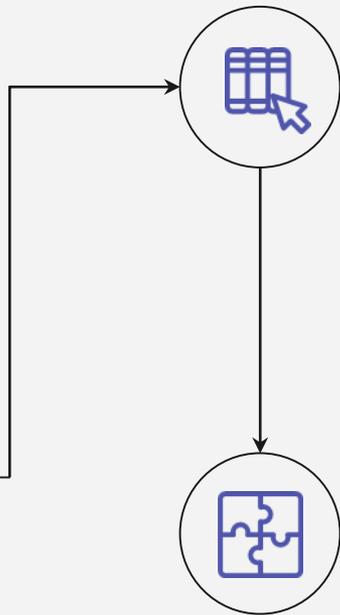
Separate **score reports for teachers** (reviewing individual and class-wide results) & **students** (personalized feedback on performance with actionable steps for skill improvement)

Design Process and Achievements

Phase 1:

Groundwork establishment

- ① Articulation of reporting in design
- ② Audience identification
- ③ Needs assessment
- ④ **Literature and document review**



Sources of review

Common features

Score reports from 12 major English-speaking tests both with/without automated scoring systems (Lee et al., 2023)

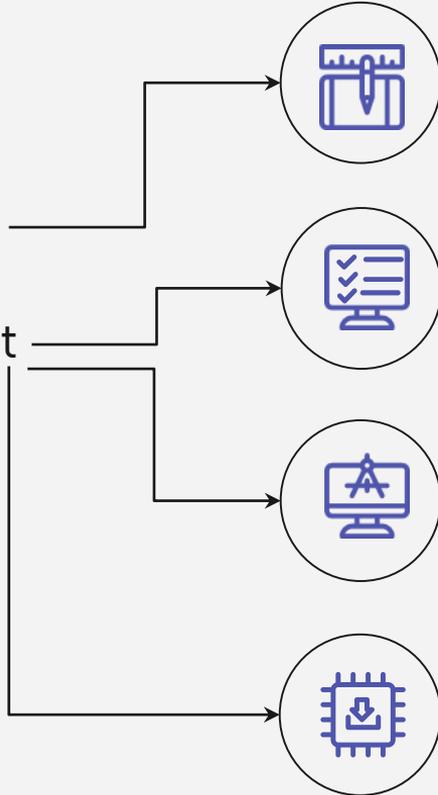
Literature on score report development and design guidelines, particularly focusing on "score reporting features" (e.g., Hambleton & Zenisky, 2013; Hattie, 2010; Zenisky & Hambleton, 2015)

Numeric scores, proficiency levels (e.g., CEFR), analytic scores, verbal description & visual graphs, & learning materials/activities

Design Process and Achievements

Phase 2:

Report development



Draft creation

Designing structure and content for student and teacher reports on PPT slides based on Phase 1 decisions

**Review,
Refinement, &
Finalization**

Reviewing and refining drafts through multiple meetings, followed by finalizing the prototype

**Visual design
by graphic artist**

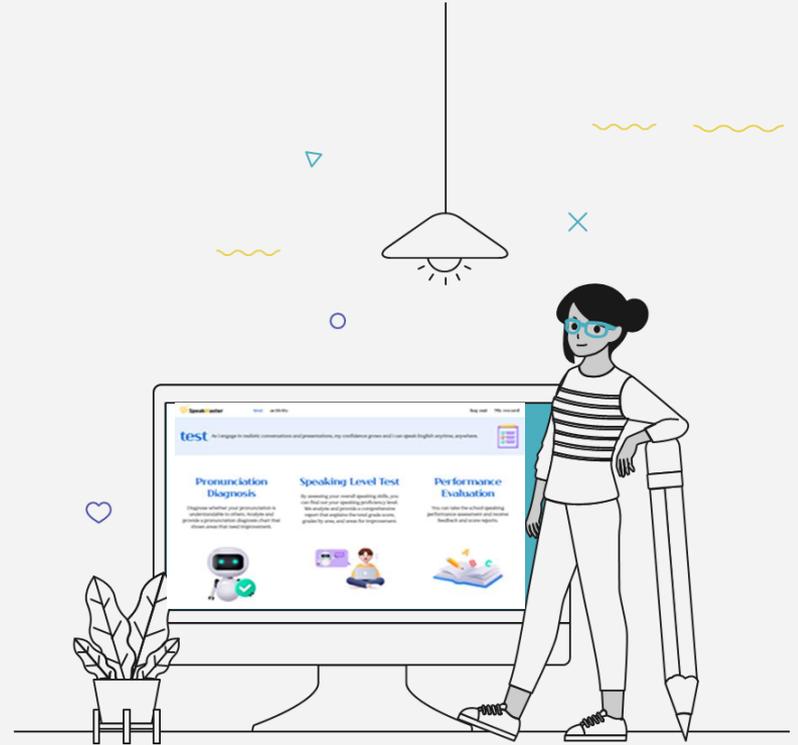
Translating the finalized prototype into simulated screens with the assistance of a graphic designer

**Online
implementation**

Integrating the prototype score reports into the SPEAK MASTER's Speaking Level Test web platform



SPEAKMASTER: Speaking Level Test's Prototype Individual Score Report



Prototype Individual Score Report Overview

Speaking Level Test Speaking

Descriptive Section

YunDeok Choi
Test School 7 | Exam Date: November 18, 2024 | Time taken: 2 minutes 18 seconds

Overall Evaluation Score and Feedback

Overall Score and Feedback Section

74

Detailed Diagnostic Results and Learning Suggestions

		Needs work	Good	Excellent
Oral skills	Fluency (7.0)	Needs work	Good	Excellent
	Accuracy (12.0)	Good	Excellent	Excellent
	Interaction (10.0)	Good	Excellent	Excellent
Comprehension	Listening (10.0)	Good	Excellent	Excellent
	Reading (10.0)	Good	Excellent	Excellent
Production	Writing (7.0)	Needs work	Good	Excellent
	Speaking (7.0)	Needs work	Good	Excellent

Recommended Learning & Practice Activities

I recommend the following practice activities to students!

pronunciation practice
 presentation practice

- **Main purposes of the online score report**
 - Highlighting strengths and areas for improvement in speaking
 - Providing personalized feedback to help students improve their speaking skills with actionable suggestions for targeted learning
- **Key design goals**
 - **User-friendliness:** simplified terms and explanations, complemented by visual graphs for ease of understanding
 - **Motivation:** a reward system featuring trophies to encourage and celebrate student achievements

Descriptive Section Overview



- **Purpose**

- Providing essential context for interpreting test results

- **Key components**

- Test administration details: test name, test date, & completion time
- Student information: student name, school name, grade
- Achievement recognition: Trophies awarded based on performance (1 to 3 trophies depending on results)

Overall Score and Feedback Section Overview

Overall Evaluation Score and Feedback



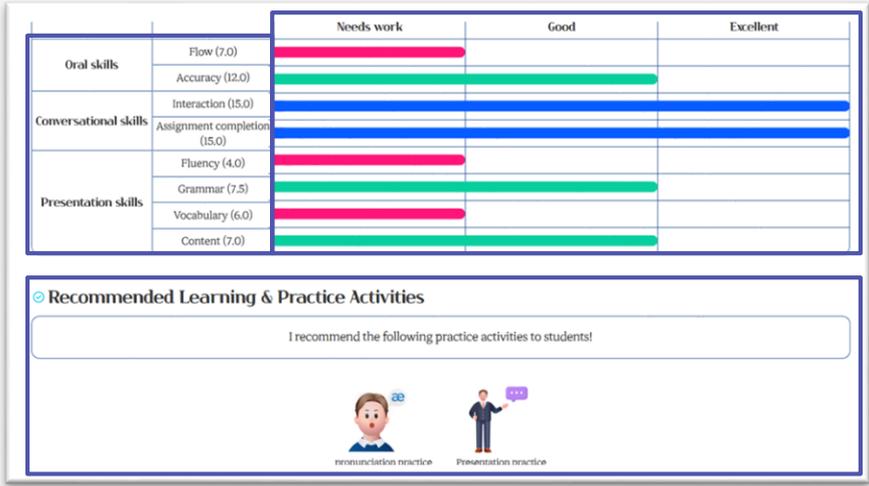
74

The student has almost reached the standard for speaking ability. He is doing well. [Oral ability](#) is very important as the foundation of speaking. Practice reading out loud so that you can read at a natural speed and pause. Compare words and sentences with native speaker speech so that you can pronounce each sound accurately. Try to imitate not only pronunciation but also stress and intonation. You will soon be satisfied with your fluent pronunciation skills. [Conversational ability](#) enables you to communicate with others. You are conversing at an appropriate speed without any problems in the communication area. You have perfectly performed the mission according to the situation. I praise you. [Presentation ability](#) can be useful in the future. You need to increase the amount and speed of speech so that you can sufficiently convey what you want to say in the given time. Listen to native speaker speech and practice repeatedly. Check if it is grammatically correct and find appropriate expressions (words and phrases) and read them repeatedly and practice the meaning. Above all, do not forget to carefully review the assignment instructions before recording to make sure that all the content is included.

- **Purpose**
 - Providing a comprehensive summary of performance across all domains
 - Delivering actionable feedback to guide improvement in weaker areas
- **Key components**
 - Test-taking level is aligned with the student's grade
 - Summative results: total raw score (scale: 0–100), calculated as the sum of three domain scores
 - Text-based feedback highlighting strengths and weaknesses in each domain and offering tailored learning advice to address identified weaknesses



Detailed Diagnostic Results and Learning Suggestions Section

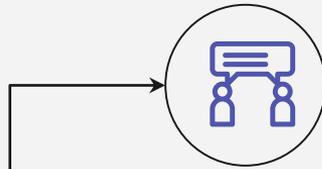


- **Purpose**
 - Providing domain-specific detailed performance results and actionable feedback to guide further improvement
- **Key components**
 - Subdomain performance: breakdowns of scores and level classification across constructs within each domain
 - Visual graph: Clear, color-coded performance bars indicating skill levels ("Needs work," "Good," "Excellent").
 - Learning suggestion: Links to recommended practice activities targeting areas marked as "Needs work."



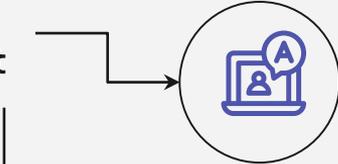
Future Directions and Next Steps

Phase 3: Field Test & Revision



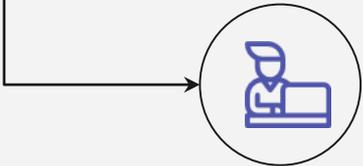
Internal reviews

(Completed) Conducting **iterative reviews and refining the structure, content, and usability** of the score reports



External review preparation

(Completed) **Designing surveys** for students and **interview questions** for teachers focusing on *comprehensibility* and *preferences* (e.g., Hambleton & Zenisky, 2013)



Piloting

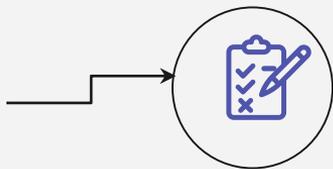
Pilot tests planned to start January 2025: **Over 1,000 students and 10+ teachers** in elementary and secondary schools **across multiple cities** will participate in SPEAK MASTER's Speaking Level test.



Future Directions and Next Steps



Phase 4:
Evaluation &
Maintenance



**Post-release
evaluation**

Gathering insights on score report design (**usability & clarity**) through surveys and interviews with stakeholders (O'Leary et al., 2016)

Monitoring how stakeholders interpret & utilize the score reports

Implementing major or minor revisions to the reports based on the feedback collected during post-release evaluations



Conclusions



Established groundwork for user-friendly and motivating score report development



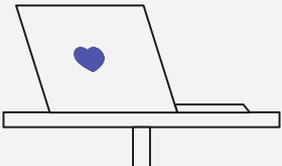
Informed by prior research on score report development



Systematically integrated the score report into the automated speaking assessment system (SPEAK MASTER)



Next steps conducting field tests and iterative revisions to improve usability and clarity for stakeholders



Thanks!

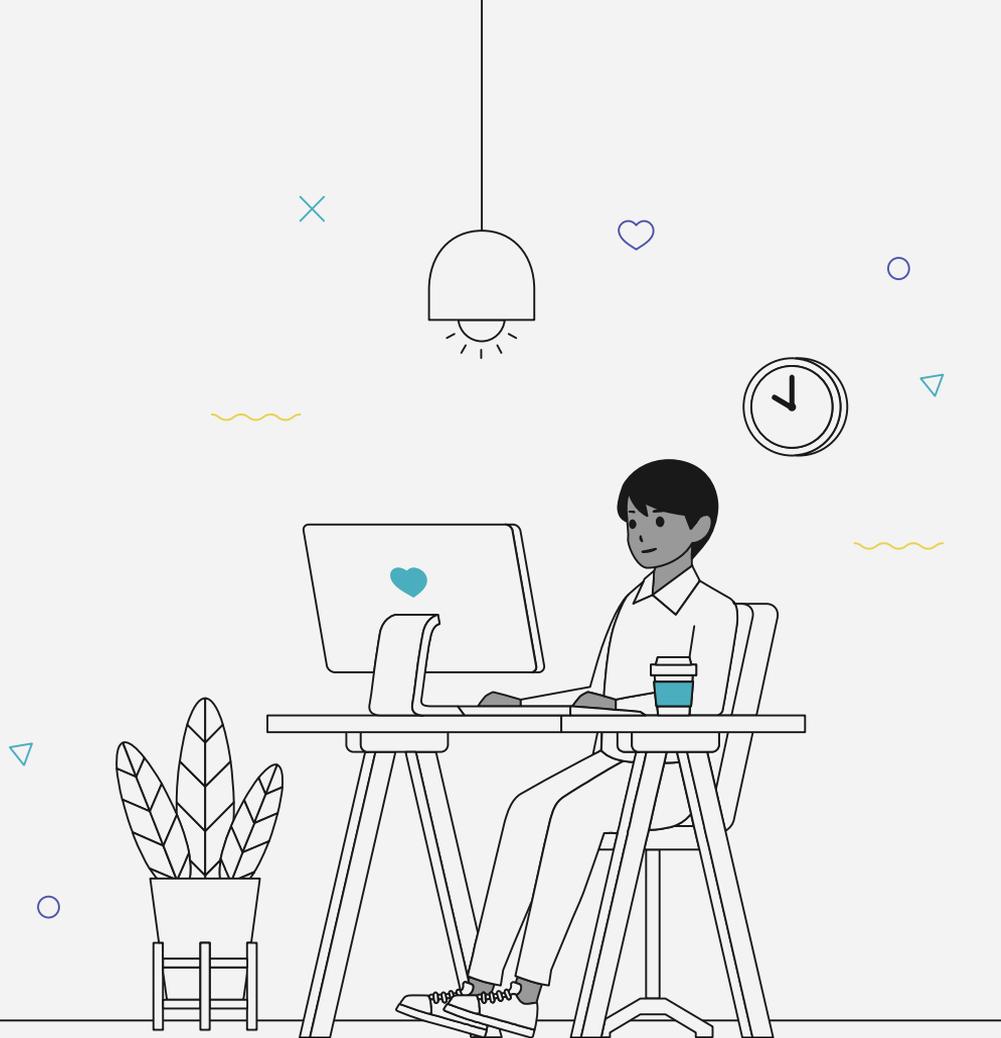
Do you have any questions?

englishnet@cau.ac.k

jinhlee@cau.ac.kr

yundeokchoi@cnu.ac.kr

CREDITS: This presentation template was created by [Slidesgo](#), and includes icons by [Flaticon](#), and infographics & images by [Freepik](#)





References

- American Council on the *Teaching of Foreign Languages* (ACTFL). (2017). *NCSSFL–ACTFL can-do statements: Performance indicators for language learners*.
- Brown, G. T., O’Leary, T. M., & Hattie, J. A. (2019). Effective Reporting for Formative Assessment: The asTTle case example. In D. Zapata–Rivera (Ed.), *Score reporting research and applications* (pp. 107–125). New York, NY: Routledge.
- Council of Europe. (2020). *Common European framework of reference for languages: Learning, teaching, assessment (Companion volume with new descriptors)*.
- De Jong, N. (2023). Assessing second language proficiency. *Annual Review of Linguistics*, 9, 541–60. <https://doi.org/10.1146/annurev-linguistics-030521-052114>
- Evanini, K., & Zechner, K. (2019). Overview of automated speech scoring. In K. Zechner & K. Evanini (Eds.), *Automated speaking assessment: Using language technologies to score spontaneous speech* (pp. 3–20). Routledge.
- Fryer, L. K., Ainley, M., Thompson, A., Gibson, A., & Sherlock, Z. (2017) Stimulating and sustaining interest in a language course: An experimental comparison of Chatbot and human task partners. *Computers in Human Behavior*, 75, 461–468. <http://10.1016/j.chb.2017.05.045>
- Fulcher, G (2015). Assessing second language speaking. *Lang. Teach*, 48(3), 198–216.
- Gu, L. & Davis, L. (2019). Overview of automated speech scoring. In K. Zechner & K. Evanini (Eds.), *Automated speaking assessment: Using language technologies to score spontaneous speech* (pp. 159 – 175). Routledge.
- Gu, L., Davis, L., Tao, J., & Zechner, K. (2021). Using spoken language technology for generating feedback to prepare for the TOEFL iBT® test: A user perception study. *Assessment in Education: Principles, Policy & Practice*, 28(1), 58–76.
- Hsieh, C. N. (2023). *Evaluating the use and interpretation of the TOEIC® listening and reading test score report: Perspectives of test takers in Japan*. ETS Research Report Series, 2023(1), 1–16.
- Hu, Y. (2019). *Do people want to message chatbots? Developing and comparing the usability of a conversational vs. menu-based chatbot in context of new hire onboarding*. Espoo: Aalto University, unpublished Master thesis.
- Jeon, J.-H., Lee, J.-Y., & Kim, J.-R. (2018) Development of survey to inquire learners’ awareness of language competence based on CEFR Basic User Level. *Asia-pacific Journal of Multimedia Services Convergent with Art, Humanities, and Sociology*, 8(3), 199–210.
- Jia, J., & Ruan, M. (2008). Use chatbots CSIEC to facilitate the individual learning in English instruction: A case study. In B. Woolf, E. Aimeur, R. Nkambou, & S. Lajoie. (Eds.), *ITS 2008: Proceedings of international conference on intelligent tutoring systems*. University of Quebec, 23–27 June, 706–708. https://doi.org/10.1007/978-3-540-69736-7_20
- Kannan, P., Zapata–Rivera, D., & Leibowitz, E. A. (2018). Interpretation of score reports by diverse subgroups of parents. *Educational Assessment*, 23(3), 173–194.
- Kim, H. (2018, June 27). *Designing L2 interactive tasks with an artificial intelligence robot*. Paper presented at Asia TE²¹ International Conference. University of Macau, China. Retrieved from <https://www.slideshare.net/heyoungkim/designing-l2-interactive-tasks-with-an-artificial-intelligence-robot>.
- Kim, H., & Lee, S. (2023). A study on task design of AI-based voice chatbot for elementary English speaking. *Multimedia-Assisted Language Learning*, 26(1), 31–58.





References

- Kim, H., Shin, D., Yang, H., & Lee, J. H. (2019). A study of AI chatbot as an assistant tool for school English curriculum. *Journal of Learner-Centered Curriculum and Instruction*, 19(1): 89–110.
<http://dx.doi.org/10.22251/jlcci.2019.19.1.89>
- Kim, H., Yang, H., Shin, D., & Lee, J. H. (2022). Design principles and architecture of a second language learning chatbot. *Language Learning & Technology*, 26(1), 1–18.
<http://hdl.handle.net/10125/73463>
- Lee, J.-H., Choi, Y., Sung, M.-C., & Kim, H. (2023). Analysis of English automated speaking scoring tests. *English Teaching*, 78(2), 223–244.
<https://doi.org/10.15858/engtea.78.2.202306.223>
- McNamara, T. F. (1996). *Measuring second language performance*. London: Longman.
- Ministry of Education. (2022). *English curriculum (No. 2022–33 [supplement 14])*. Sejong: Ministry of Education.
- Singla, Y. K., Gupta, A., Bagga, S., Chen, C., Krishnamurthy, B., & Shah, R. R. (2021, October). Speaker-conditioned hierarchical modeling for automated speech scoring. In *Proceedings of the 30th ACM international conference on information & knowledge management* (pp. 1681–1691).
- Shum, H.-Y., He, X., & Li, D. (2018) From Eliza to Xiaoice: Challenges and opportunities with social chatbots. *Frontiers of Information Technology & Electronic Engineering*, 19(1), 10–26.
<https://doi.org/10.1631/FITEE.1700826>
- Vezzu, M., VanWinkle, W., & Zapata-Rivera, D. (2012). *Designing and evaluating an interactive score report for students*. Princeton, NJ: ETS.
- Wang, Y. F., & Petrina, S. (2013) Using learning analytics to understand the design of an intelligent language tutor–chatbot Lucy. *Editorial Preface*, 4(11), 124–131.
<https://doi.org/10.14569/IJACSA.2013.041117>
- Xu, J., Jones, E., Laxton, V., & Galaczi, E. (2021). Assessing L2 English speaking using automated scoring technology: examining automarker reliability. *Assessment in Education: Principles, Policy & Practice*, 28(4), 411–436.
- Yang, H., Kim, H., Lee, J. H., & Shin, D. (2022). Implementation of an AI chatbot as an English conversation partner in EFL speaking classes. *ReCALL*, 34(3), 327–343.
doi:10.1017/S0958344022000039
- Zapata-Rivera, D., Kannan, P., & Zwick, R. (2018). Communicating measurement error information to teachers and parents. In *Score reporting research and applications* (pp. 63–73).
Routledge.

