

## ENGLISH SPEAKING SKILLS ASSESSMENT IN THE "ENGLISH ON TUESDAY" PROGRAM: A CIPP-BASED EVALUATION APPROACH

Witri Luthfiyanti<sup>1</sup>, Nisa Aliyah<sup>2</sup>, Dhika Arya<sup>3</sup>

<sup>1</sup> Al-Irhaam Elementary School Bandung, Indonesia

<sup>2</sup> UIN Sunan Gunung Djati Bandung, Indonesia

<sup>3</sup> UIN Sunan Gunung Djati Bandung, Indonesia

*Corresponding E-mail: [witriluthfiyanti27@gmail.com](mailto:witriluthfiyanti27@gmail.com)*

### ABSTRACT

This study evaluates the speaking assessment component of the "English on Tuesday" program at Al-Irhaam Elementary School, Bandung, a vocabulary enrichment initiative for primary school students in Grades 1-6. Using the Context, Input, Process, and Product (CIPP) evaluation framework, the research examines how speaking tasks embedded in the program affect young learners' oral proficiency. The context evaluation emphasizes the program's relevance to students' communicative needs and motivation to use English in daily school activities. The input evaluation demonstrates that multimedia materials, vocabulary lists, and simple phrase patterns serve as the foundation for speaking practice, supported by educational videos and interactive digital games. The process evaluation indicates that speaking activities, including sentence construction, role-playing, and oral reporting, promote dynamic and engaging language use among students aged 6-12 years old. The product evaluation reveals significant improvements in students' fluency, confidence, and ability to apply vocabulary in meaningful contexts, as documented through teacher reflection journals, classroom observations, and student speaking samples collected over one academic semester. The findings indicate that integrating structured speaking tasks with vocabulary learning, accompanied by age-appropriate resources and interactive approaches, can considerably improve speaking skills in early English language instruction. This study provides practical insights into developing effective formative speaking assessments for vocabulary-focused programs targeting young learners in Indonesian elementary school contexts.

**Keywords:** CIPP model, English Tuesday Program, Speaking Assessment.

### INTRODUCTION

In today's globally interconnected world, English language proficiency has become increasingly valued, especially in international educational environments. Speaking, as one of the four fundamental language skills, is often considered the most immediate and apparent form of communication because it requires learners to produce language in real-time situations. For primary school students, particularly those in global learning environments, speaking promotes not only linguistic competence but also social integration, cognitive growth, and academic advancement (Putri Refani, 2006). However, assessing speaking skills in young learners presents unique challenges due to developmental diversity, short attention spans, and the need for engaging, age-appropriate tasks (Nur Alwiyah, 2019).

English is widely recognized as a global language and serves as an international medium of communication in various countries across the world. In the current era of globalization, English functions as a universal and multinational language that facilitates

communication at international levels (Sigalingging et al., n.d.). As stated by Rinny S Rorimpandey et al (2013), English plays a crucial role as a national language in numerous countries, making it essential for students to develop proficiency from an early age. Consequently, English teachers need to develop their students' speaking ability as a foundation for future communicative competence.

Speaking is used for various purposes, including expressing opinions, describing objects, making complaints, persuading others, and making polite requests (Tirtanawati & Salsabila, 2021). From this perspective, speaking is considered one of the most important skills in learning English as a Foreign Language (EFL), particularly in contexts where English is not the primary language of instruction (Titik Komariyah et al., 2022). Students will use this skill when they want to express everything that is in their minds. Therefore, the correct way of speaking will guarantee the success of interaction between students in the learning process (Puji et al., n.d.).

However, speaking skills are usually one of the most challenging skills to teach for teachers and to learn for students. Most students think that they cannot speak fluently, which leads them to feel that they cannot develop their speaking skills adequately. This is in line with what Sun et al (2017) observed, noting that students often feel unable to develop their fluency, resulting in a lack of communication ability. Teachers must use appropriate resources and assessment methods in teaching and helping students improve their speaking skills. As stated by Syifa (2022), EFL students' speaking skills can be improved by using appropriate learning resources combined with effective assessment practices.

Evaluating young students' speaking ability provides important information about their language development, fluency, and communication skills. It enables teachers to identify students who require additional assistance, customize their lessons, and promote self-expression among learners. Additionally, successful speaking evaluation increases students' motivation, encourages goal-setting, and boosts their confidence, all of which are critical qualities in early education, especially in global curricula that prioritize intercultural competency and communication (Ian Clark, 2012).

The use of innovative learning media and assessment tools is one method of applying 21st-century learning principles. As Sharma (2019) pointed out, social media applications and digital tools are useful resources to improve EFL students' English language skills and motivate learning. Language learning has been transformed by technology, with applications offering fun approaches to learning through interactive language games, quizzes, and audio-visual information. In this context, the integration of multimedia resources in teaching and assessment becomes increasingly relevant.

The "English on Tuesday" program is a vocabulary enrichment initiative implemented at Al-Irhaam Elementary School in Bandung, aimed at helping students in Grades 1 to 6 acquire 1,000 English words over six years of primary education. While the program primarily focuses on vocabulary acquisition through various multimedia resources, including YouTube videos and educational games, speaking tasks are embedded throughout the curriculum to reinforce contextual usage and oral production. This integration of speaking activities within a vocabulary-focused program presents a unique opportunity to examine how assessment practices can support both vocabulary learning and oral proficiency development.

Despite the growing body of research on speaking assessment in EFL contexts, there remains a gap in understanding how speaking assessments can be effectively

integrated into vocabulary-based programs for young learners in Indonesian primary schools. This study addresses this gap by investigating how speaking assessments are incorporated into the "English on Tuesday" program, examining their design, implementation process, and learning outcomes through the comprehensive lens of the CIPP evaluation model. The research aims to provide practical insights for educators seeking to develop effective formative speaking assessment practices within vocabulary enrichment programs.

Given these considerations, this research seeks to answer the following questions: How are speaking assessments designed and implemented within the "English on Tuesday" program? To what extent do embed speaking tasks contribute to students' oral proficiency development? What are the strengths and areas for improvement in the current speaking assessment practices? By addressing these questions through systematic evaluation, this study contributes to the broader discourse on effective assessment practices for young English language learners.

### **Theoretical Background**

Understanding the theoretical foundations of vocabulary-based speaking tasks is essential for designing effective assessments for young learners. This section reviews key linguistic and pedagogical theories related to vocabulary acquisition, communicative competence, and task-based learning, highlighting how these frameworks inform the development and evaluation of speaking activities in the "English on Tuesday" program.

#### **2.1 Vocabulary-Based Speaking Tasks**

One of the basic parts of speaking ability is vocabulary. Nation (2001) highlighted that in order to talk successfully, one must possess both scope (the number of words one learns) and level (the level of one's word knowledge). Fluency and coherence in oral performance are strongly impacted by lexical access fluency. Vocabulary is a core component of fluency and communication skills; without mastery of vocabulary, speaking ability is limited.

Research shows that in addition to vocabulary measurement (how many words one knows), vocabulary depth (understanding of meaning in context, collocations, registers, etc.) greatly influences speaking performance. The importance of vocabulary mastery in speaking particularly at the elementary level, teachers need to use some approaches to give student's task i.e TBLT (Teaching Based Language Task). Through interactive tasks that encourage real vocabulary use, TBLT improves learning a language. Ellis (2003) states that tasks tend to be gap-based, meaning-focused, and outcome-oriented. This method encourages conversation and contextual vocabulary use. There were several theories which were conducted in vocabulary tasks to encourage students' knowledge vocabulary, namely:

- a. Rooted in Communicative Language Teaching (CLT) and Information Processing Theory by (Anderson, 1983), teacher can ask the student to do some tasks by giving them picture descriptions Students describe a scene using target vocabulary (animals, actions, locations),
- b. Informed by Tasks act as "artifacts" in Vygotskian terms, Sociocultural Theory (Vygotsky) and Task-Based Language Teaching (TBL) (Feryok, 2017), role play, simulated real-life situations like shopping, ordering food through providing an authentic communication scenario.
- c. Draws from Narrative Theory and Schema Theory from British Council, use a story using some clues to encouraging use of learned vocabulary.

- d. Dual Coding Theory (Paivio, 1971), this theory states In the human brain process information through a verbal system for language and a non-verbal system for images. According to Paivio, when new information is presented both verbally (spoken or written words) and visually (pictures, symbols, or videos), learners create richer, interconnected mental representations. In the context of vocabulary-based speaking in Al-Irhaam elementary school, these theories confirm the use of picture, story illustrations and multimedia resources during English lessons. So that teachers help students form stronger associations between words and their meanings. Making young learners remember and use vocabulary during speaking activities.
- e. Multimedia Learning Theory (Mayer,2001), this theory states through proposing the meaningful learning occurs when words and pictures are presented together in well designed format. In Al-Irhaam elementary school's "English on Tuesday" program, Multimedia Learning theory supports the integration of videos, interactive games as a part of speaking tasks. For example, students might watch a short video or animated clip introducing vocabulary about animals, then practise it using words by describing scenes from the video.
- f. CIPP Evaluation Framework Integration (Daniel, Stufflebeam, 1960), Context (Analyze student needs for vocabulary acquisition to support speaking). Input (Examine teaching materials and task designs that target key vocabulary), Process (Monitor how vocabulary tasks are implemented and student engagement), Product: Assess speaking performance and vocabulary retention gains

## 2.2 CIPP Evaluation Model

The four components of the Stufflebeam-created CIPP evaluation model i.e Context, Input, Process, and Product provide a thorough framework for evaluating educational initiatives. The evaluation of initiatives like "English on Tuesday" in primary schools like Al-Irhaam is a perfect fit for this approach since it guarantees a comprehensive analysis of both execution and results.

1. Context evaluation focuses on the program's past, requirements, and purpose. When it comes to speaking evaluation in "English on Tuesday," this stage aids in determining if the curriculum is in line with students' communicative needs and learning goals (Widayanthi,2023)
2. Input Evaluation assesses the strategy, resources, and action plans used in the program. For example, the availability of English-speaking materials, qualified instructors, and time allocation would be evaluated to determine feasibility and readiness (Agustina Noor, 2019)
3. Process Evaluation focuses on the actual implementation, including classroom interactions, speaking activities, and formative assessments. It identifies strengths and bottlenecks during the program execution
4. Product Evaluation measures outcomes such as improvements in students' speaking proficiency, confidence, and participation. It ensures the program meets its intended objectives (Akhadi, 2024)

## 2.3 Second Language Acquisition Theories

The Context, Input, Process, Product (CIPP) evaluation approach will be used in this study to assess the speaking component of Al-Irham Elementary School's English on Tuesday program. By focusing the evaluation on significant Second Language

Acquisition (SLA) theories, the study seeks to determine if the program effectively supports the development of young learners' speaking abilities and whether its design is consistent with best practices in language acquisition research.

### **Applying Krashen's Input Hypothesis and the Affective Filter**

According to Krashen's data Hypothesis (1977), learners develop in their second language when they receive understandable information that is partially in addition to their current competence level ( $i+1$ ). This hypothesis also presents the Affective Filter Hypothesis, which argues that high anxiety might hinder language acquisition while low-anxiety conditions stimulate it (Krashen, 1982). In the Context evaluation, the researcher should look into students' levels of anxiety or motivation, as these affective aspects have a direct impact on how well they process and produce spoken English. Input evaluation should look at whether class materials provide comprehensible but seeking language exposure through teacher speech, audio recordings, or planned dialogues, allowing students to gradually improve their speaking skills without overcapacity.

#### **Analyzing Long's Interaction Hypothesis in Classroom Practices**

Long's Interaction Hypothesis (1981) states that language acquisition is facilitated when learners engage in conversational interaction that includes meaning discussion, such as clarification requests, confirmation checks, or repetition requests, which makes input more understandable and prompts learners to adjust their output (Long, 1996). In the Process component of the evaluation, the researcher should observe whether the English on Tuesday sessions include interactive speaking tasks such as pair or group work, role plays, or games that require students to converse for real. The analysis should focus on how frequently meaning is negotiated in these activities, as more opportunities for negotiation are associated with higher speaking fluency and accuracy.

### **Integrating VanPatten's Input Processing Theory for Comprehension and Production**

According to VanPatten's Input Processing Theory (1993), language learners place a higher value on extracting meaning from input than on noticing grammar structures. This means that students must first comprehend what they hear before being able to correctly develop new language elements (Van Patten, 1996). When studying the Input and Process stages, the researcher should consider whether speaking assignments help students understand before expecting them to speak correctly. For example, employing pictures, gestures, or simpler explanations helps learners understand meaning, which is critical for developing speaking confidence. Materials and classroom interactions can be evaluated to determine whether the program meets students' understanding needs or prematurely expects sophisticated output.

### **Applying Pienemann's Processability Theory to Task Progression**

In line with Pienemann's Processability Theory (1998), language structures are acquired in a predictable order based on learners' cognitive processing capacities; learners cannot produce structures that are not developmentally appropriate (Pienemann, 1998). Speaking activities should be evaluated in terms of input and process to see if they follow this developmental pattern by beginning with simpler utterances such as basic noun-verb phrases and proceeding to more complicated structures, such as subordinate clauses or conditional sentences. If the program asks kids to produce language beyond their developmental stage, it risks frustrating them and fossilizing their blunders.



### Incorporating Communicative and Natural Approaches to Speaking

Communicative Language Teaching (CLT) and the Natural Approach emphasize the value of meaningful conversation, authentic situations, and prioritizing fluency over form in the early stages of learning (Krashen & Terrell, 1983). Evaluating the Process component should include determining whether the English on Tuesday program provides authentic speaking opportunities, such as describing personal experiences, expressing opinions, or responding to open-ended questions, rather than limiting students to memorized dialogues or scripted exercises. Assessments for the Product component should focus on students' ability to convey meaning effectively and maintain conversational flow rather than just on grammatical accuracy.

### CIPP Framework and SLA Alignment

The Context evaluation must look into the program's match with students' needs, motivation, and readiness for speaking practice. Input evaluation should look at whether materials provide intelligible, developmentally appropriate verbal input and whether they incorporate engaging visuals or real-world situations. Process evaluation should look at educational approaches, specifically the level of engagement and meaning negotiation, whereas product evaluation should look at outcomes such as speaking fluency, correctness, and complexity. Each component of the CIPP model should be examined in light of SLA ideas to discover gaps in theory and practice.

In the context of pedagogical theories, it can be described based on the instructional design adopted in this study is grounded in the TPACK framework developed by Koehler and Mishra (2011), which emphasizes the integration of Technological Knowledge (TK), Pedagogical Knowledge (PK), and Content Knowledge (CK) to enhance effective teaching in digitally enriched environments.

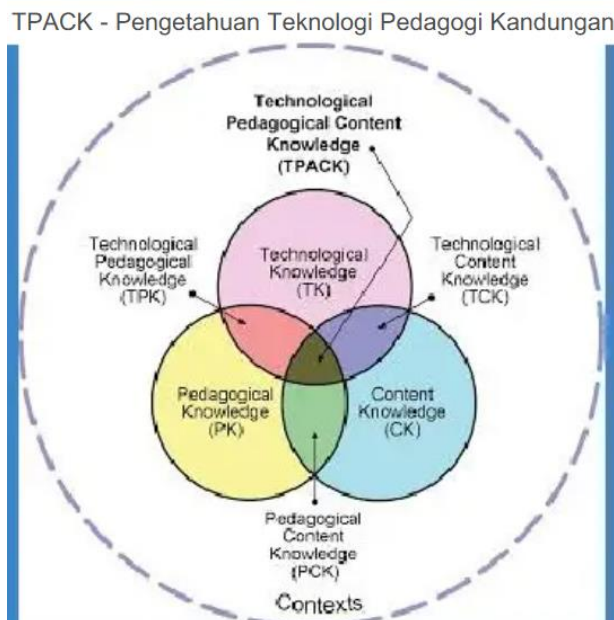


Figure 1, Koehler, M. J. (2011). *TPACK – Technological Pedagogical Content Knowledge*. Retrieved Nov 30, 2014, from <http://www.matt-koehler.com/tpack/tpack-explained>

The diagram below illustrates the theoretical foundation that guides the design, implementation, and evaluation of speaking assessments within the "English on Tuesday" program, emphasizing how cognitive and sociocultural theories influence the

development of visual-based speaking activities and the expected outcomes for young English language learners. This framework uses elements from Dual Coding Theory, Multimedia Learning Theory, Sociocultural Theory, and Communicative Language Teaching to help students learn vocabulary and communicate fluently. Using these theories, the program at Al-Irhaam Elementary School incorporates interactive and age-appropriate speaking tasks such as picture descriptions, story sequencing, and multimedia-supported games to meet the developmental requirements of young learners. This technique seeks to establish a low-anxiety, engaging environment where students can study vocabulary, develop oral fluency, and grow confidence in speaking English.

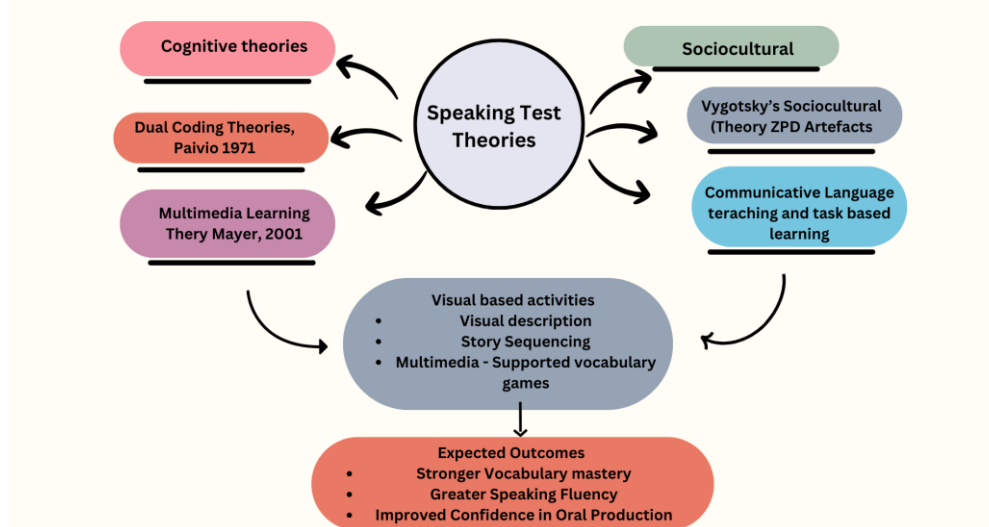


Figure 2, Pedagogical Framework for Speaking Tasks

## Literature Review

A journal entitled *Assessing Speaking Skill: A Study in the Third Level Education* (Nurkhofifah, Juliani, Manik, & Juniardi, 2023) examines in depth the concept and process of speaking skills at the higher education level using the library research method in the form of collecting and identifying factual data from several previous writings. The assessment techniques used are dynamic assessment, public speaking assessment, peer assessment, self-assessment, direct, semi-direct, and indirect techniques. In addition, technological developments have encouraged the emergence of new assessment techniques in the form of digital formative assessment and mobile-based formative assessment. The findings in this study show that the use of technology (technology-based assessment) provides more effective results compared to traditional methods.

Then the second journal entitled *Developing Assessment for Speaking* written by Rahmawati and Ertin (2014). They revealed research on contextual speaking skills in college students by adopting the criteria from Brown and Abeywickrama (2010) to design effective speaking assessment research with specific criteria, appropriate tasks, maximum output, and practical and reliable scoring procedures. There are also five types of speaking skills categorized as imitative, intensive, responsive, interactive, and extensive. In the practice of their research, students were asked to make an oral presentation for 10 minutes, with assessment criteria covering aspects of fluency, grammar, vocabulary, pronunciation, task accomplishment, and presentation skills. This assessment was designed to meet the principles of practicality, reliability, validity, authenticity, and positive washback. The assessment was carried out directly, did not require large tools or

costs, and provided specific feedback to students. The conclusion of this article is that speaking assessments must be carefully designed to be fair and valid. The authors recommend that teachers be given training and support from institutions or the government so that they can develop speaking assessments professionally and contextually.

## METHOD

This study employs a descriptive qualitative method to explore how speaking assessments were implemented in the "English on Tuesday" program. The research was conducted over one academic semester in a private Islamic elementary school in Indonesia, focusing on students in Grades 1 until 6.

The methodology includes the following components:

**3.1 Participants:** Participants included 12 homeroom teachers, 3 English Teachers, and approximately 144 students from Grade 1 until 6. The homeroom teachers were directly involved in delivering the English on Tuesday sessions and documenting student performance.

### 3.2 Data Collection Techniques:

- **Classroom Observations:** Weekly non-intrusive observations were conducted to record how speaking tasks were introduced and how students responded. Observation sheets focused on student engagement, accuracy, fluency, and classroom interaction.
- **Teacher Reflection Journals:** Teachers maintained weekly logs to reflect on the teaching-learning process, including notes on student difficulties, successful strategies, and spontaneous speech production.
- **Student Speaking Samples:** Student oral performances were collected during vocabulary reinforcement activities. Tasks included sentence-making and question-and-answer sessions using target vocabulary.
- **Artifacts and Visual Media:** Photos of student worksheets, screenshots from educational videos (e.g., YouTube), and results from digital games (TinyTap) were included as supporting evidence.

**3.3 Instruments:** A checklist and speaking performance rubric were developed based on the adapted CEFR (Common European Framework of Reference for Languages) speaking descriptors for young learners. The rubric assessed vocabulary accuracy, sentence formation, pronunciation, and confidence on a 4-point scale.

**3.4 Data Analysis:** All qualitative data were coded thematically, guided by the CIPP model. Codes were grouped under four categories: (1) context relevance, (2) input support and material quality, (3) process fidelity and classroom practice, and (4) observable outcomes. Visual data were triangulated with teacher reflections and observation logs to ensure reliability.

## FINDING AND DISCUSSION

### 4.1 Context Evaluation

The program responds to the need for improved English vocabulary among primary students. Speaking assessments were introduced to reinforce contextual understanding of vocabulary through oral use.



## 4.2 Input Evaluation

Speaking tasks are supported by multimedia inputs (e.g., YouTube videos), vocabulary worksheets, and games. Homeroom teachers are equipped with general guidelines, including vocabulary lists and simple sentence patterns using target words (e.g., "There is one pencil"). In this input evaluation process, the teacher prepares open materials as a guide to train students' speaking skills in class. Examples of teaching materials created by teachers can be seen in the following image.



### 2<sup>nd</sup> Grade

#### February Week 2

Num	Vocabularies
Video	<a href="https://www.youtube.com/watch?v=hjFaqDNUVFo">https://www.youtube.com/watch?v=hjFaqDNUVFo</a>
1	Pencil / pen
2	Rubber/eraser
3	School Bag
4	Book
5	Crayon
Game	<a href="https://www.tinytap.com/activities/g43mw/play/school-supplies">https://www.tinytap.com/activities/g43mw/play/school-supplies</a>

#### February Week 3


Num	Vocabularies	Example Sentences
Video	<a href="https://www.youtube.com/watch?v=eUPOGiAEwqg">https://www.youtube.com/watch?v=eUPOGiAEwqg</a>	
1	There is	<b>There is</b> one pencil
2	There are	<b>There are</b> two books
	Students mengucapkan 5 kalimat menggunakan ( alat-alat tulis dengan <i>there is</i> dan <i>there are</i> seperti di kolom <b>Example Sentences</b> )	

In the example image above, there is an input evaluation process carried out by the teacher on grade 2 students. The material taught is in the form of objects found in the classroom. As the first input, the teacher presents an animated video containing a song about various objects in the classroom along with how to write and pronounce them. Students then repeat what is in the video or what is said by the teacher to get used to and introduce new vocabulary. After that, students practice through the available games. Then in the following week (week three), students are introduced to how to use "there is" and "there are" which will then be made into five simple sentences using these keywords.

## 4.3 Process Evaluation

Speaking is assessed during classroom activities such as sentence-making, oral reporting, and role-playing. Teachers guide students to say five sentences using weekly vocabulary in a structured format. Creativity and engagement are fostered through digital

games and peer interaction. One of the methods used in the evaluation process is as shown in the image below.



**SD AL-IRHAAM GLOBAL ISLAMIC SCHOOL**

JL. A.H. NASUTION NO.133 BANDUNG  
 TELP/FAX : (022)-6376173  
 WEBSITE : WWW.ALIRHAAMBANDUNG.SCH.ID  
 E-MAIL : AGISBANDUNG@GMAIL.COM

---

**February Week 3**

No	Vocabularies	Your Notes	Teacher's Signature


**February Week 4**

No	Vocabularies	Your Notes	Teacher's Signature

When students have known and understood the material presented in class, the teacher makes an assessment sheet as a formative assessment while the lesson is in progress. This aims to see how far students understand the material that has been given previously. In practice, students are asked to write down any new vocabulary they get during that week's learning and write it on a table. After students can write down what was previously taught and pronounce it correctly, the teacher will then provide a signature on the table and give notes as feedback to students.

#### 4.4 Product Evaluation

End-of-semester evaluations include oral tests where students must use vocabulary in meaningful contexts. Observations show increased fluency and confidence, especially among Grades 1–6. Teachers document student progress through vocabulary logs and anecdotal records.



**SD AL-IRHAAM GLOBAL ISLAMIC SCHOOL**  
 JL. A.H. NASUTION NO.133 BANDUNG  
 TELP/FAX : (022)-6376173  
 WEBSITE : WWW.ALIRHAAMBANDUNG.SCH.ID  
 E-MAIL : AGISBANDUNG@GMAIL.COM

---

**REPORT VOCABULARIES ASSESSMENT**  
2024 - 2025

**Name** : Andhra Rasya Ramadhan  
**Class** : 2 Ali bin Abi Thalib

NO	Criteria of Assessment	Assessment Categories			
		Excellent	Good	Fair	Poor
1.	Vocabulary Mastery	✓			
2.	Fluency	✓			
3.	Pronunciation	✓			
4.	Complete Sentence	✓			

**Explanation of Assessment:**

- **Fluency:** Speaking smoothly without pauses or errors that hinder communication.
- **Vocabulary Mastery:** Ability to master and use a diverse and relevant vocabulary.
- **Complete Sentences:** Ability to construct sentences with complete and correct structures
- **Pronunciation:** Clear pronunciation that adheres to English language standards.

**Assessment Categories:**

- **Excellent:** Meets the criteria very satisfactorily.
- **Good:** Meets the criteria well.
- **Fair:** Some deficiencies but still adequate.
- **Poor:** Requires significant improvement and development

**Description:**

Alhamdulillah, overall, the vocabulary assessment is excellent. Keep up the great work and continue to maintain your English skills.

The principal

Dian Andrayani, S.Pd  
 NUKS : 19023L0750260222158019

Bandung, December 20<sup>th</sup>, 2024

English Coordinator

Witri Luthfiyanti, S.Pd

As the final step of learning, when the material has been taught, students receive an assessment report as the final product of the CIPP method, namely Evaluation. The teacher presents a table and scoring rubric for each student along with a description of the student's development during learning. This report will then be given to the student's parents as a learning development report.

## CONCLUSION

This CIPP-based evaluation of speaking assessment in the "English on Tuesday" program reveals that integrating structured speaking tasks within a vocabulary enrichment initiative can significantly enhance young learners' oral proficiency while simultaneously supporting vocabulary acquisition. The evaluation demonstrates that when speaking assessment is thoughtfully designed with age-appropriate tasks, supported by multimedia

resources, implemented through engaging classroom practices, and evaluated formatively, substantial benefits emerge for elementary students' English language development.

Context evaluation confirmed that the program responds appropriately to identified needs for improved communicative competence beyond vocabulary memorization. The integration of speaking assessment aligns with institutional goals, stakeholder expectations, and students' developmental needs for active, participatory learning experiences.

Input evaluation revealed that multimedia resources, particularly educational videos and interactive games, provide crucial support for speaking practice. The combination of visual, auditory, and kinesthetic inputs creates rich learning environments that facilitate vocabulary retention and oral production. However, variability in teacher preparation and confidence regarding speaking assessment indicates need for enhanced professional development and ongoing instructional support.

Process evaluation demonstrated that diverse speaking activity types including sentence construction, picture description, role-play, and oral reporting effectively engage students and promote meaningful language use. Teacher facilitation strategies such as modeling, scaffolding, and positive reinforcement create supportive environments where students feel comfortable attempting oral communication. The embedded nature of formative assessment within regular classroom activities reduces anxiety while providing continuous feedback for learning.

Product evaluation documented substantial improvements in students' vocabulary application, fluency, pronunciation, and confidence across all grade levels. End-of-semester assessments indicated that 87% of students met or exceeded age-appropriate expectations for speaking tasks, representing significant achievement for a program primarily focused on vocabulary acquisition. Observable outcomes validate the effectiveness of integrated speaking assessment while also identifying students requiring additional support and highlighting needs for greater assessment consistency across teachers.

These findings carry important implications for elementary English language education in Indonesian contexts and similar EFL settings. The study demonstrates that vocabulary programs need not treat speaking as a separate skill but can effectively embed oral production tasks throughout vocabulary instruction. This integration enhances both vocabulary retention and speaking development while making efficient use of limited instructional time.

The theoretical frameworks guiding this evaluation including Dual Coding Theory, Multimedia Learning Theory, principles of Task-Based Language Teaching, and the CIPP model itself proved valuable for understanding how and why particular practices succeeded. These frameworks provide guidance for educators seeking to design effective speaking assessment practices grounded in research evidence.

## REFERENCES

Anne Feryok (2017). Sociocultural Theory and Task-Based Language Teaching: The Role of Praxis. <https://www.jstor.org/stable/i40211825>. Vol. 51, No. 3, SPECIAL TOPIC ISSUE: COMPLEMENTARY PERSPECTIVES ON TASK-BASED CLASSROOM REALITIES (SEPTEMBER 2017), pp. 716-727 (12

- pages). Published By: Teachers of English to Speakers of Other Languages, Inc. (TESOL)
- Cameron, L. (2001). *Teaching Languages to Young Learners*. Cambridge University Press.
- Daniel Stuffelebeam, 1960, Curriculum Development
- Dannie Yusfi Akhadi and Imam Shofwan Evaluation of the CIPP Model of The Sewing Training Program, <https://orcid.org/0000-0003-2459-5988>, Published 2024-02-15
- Ellis, R. (2003). *Task-based Language Learning and Teaching*. Oxford University Press.
- John Benjamins Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, 1(1), 1–47.
- Krashen, S. D. (1982). *Principles and Practice in Second Language Acquisition*. Pergamon <https://www.researchgate.net/journal/Biotechnology-Letters-1573-6776>
- Long, M. H. (1996). The role of the linguistic environment in second language acquisition. *Handbook of Second Language Acquisition*.
- Luh Made Dwi Wedayanthi, [et.al](#) (2023) The Implementation of CIPP as Evaluation Model on Teaching Assistance in SD Bali Bilingual School as The Practice of Kurikulum Merdeka
- Mayer, R. E. (2001). *Multimedia learning*. Cambridge: Cambridge University Press.
- McKay, P. (2006). *Assessing Young Language Learners*. Cambridge University Press.
- Nandang, A., Behesyti, S. H., & Nugraha, D. (2024). The Use of Augmented Reality (AR) Technology to Improve Students' Arabic Vocabulary Comprehension and Pronunciation at Madrasah Aliyah. *Tadris Al-'Arabiyyah: Jurnal Pendidikan Bahasa Arab Dan Kebahasaaraban*, 3(2), 274–291. <https://doi.org/10.15575/ta.v3i2.39432>
- Nugraha, D., Husni, F. A. N., Ruhendi, A., & Suhartini, A. (2025). *Evaluation The Development Of Diversity Students Elementary School*. 1(1).
- Nugraha, D., Husni, F. A. N., Ruhendi, A., & Suhartini, A. (2025). Evaluation The Development Of Diversity Students Elementary School. *Japdimas: Jurnal Pengabdian Masyarakat*, 1(1), 1–8.
- Nation, I.S.P. (2001). *Learning Vocabulary in Another Language*. Cambridge University Press.
- Noor Qomaria Agustina and Fanny Mukhtaruddin , published 2019-11-07. The Cipp Model-Based Evaluation on Integrated English Learning (IEL) Program at Language Center DOI: <https://doi.org/10.12928/eltej.v2i1.1043>
- Nurkhofifah, M. H., Juliani, A., Manik, U., & Juniardi, Y. (2023). Assessing the Speaking Skill: A Study in the Third Level Education. *JELTS (Journal of English Language Teaching and Studies)*, 6 (2).
- Paivio, A. (1971). *Imagery and verbal processes*. New York: Holt, Rinehart and Winston. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413–468). San Diego: Academic Press.
- Pienemann, M. (1998). *Language Processing and Second Language Development: Processability Theory*.
- Pinter, A. (2006). *Teaching Young Language Learners*. Oxford University Press.



- Rahmawati, Y., & Ertin. (2014). Developing Assessment for Speaking. *IJEE (Indonesian Journal of English Education)*, 1 (2), 199-210.
- Stufflebeam (Eds.), *International handbook of educational evaluation* (pp. 31–62). Dordrecht: Kluwer Academic.
- Stufflebeam, D. L. (2003). The CIPP Model for Evaluation. In the *International Handbook of Educational Evaluation*.
- VanPatten, B. (1996). *Input Processing and Grammar Instruction in Second Language Acquisition*. Ablex
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.