

---

# Interactive English E-Learning Based on Cloud Speech-to-Text API

Andi Nurkholis <sup>1</sup>, Renda Bimantara <sup>2</sup>, Neneng <sup>3</sup>

<sup>1,2,3</sup> Department of Informatics, Teknokrat Indonesia University, Lampung, Indonesia  
email:<sup>1</sup> andinh@teknokrat.ac.id, <sup>2</sup> rendabm.mhs@teknokrat.ac.id, <sup>3</sup> neneng@teknokrat.ac.id

DOI: <https://doi.org/10.21107/edutic.v9i1.12486>

Diterima: 17 November 2021 | Direvisi: 3 Oktober 2022 | Diterbitkan : 2 Nopember 2022

---

## Abstract

*Good English skills will be a competitive asset in the world of work and education. Mastery of English is essential to be prepared early on. Schools allocate study time and learn how to use textbooks, which sometimes makes students bored and bored. Students are expected to be more motivated in learning and practicing English language skills. So we need a suggestion for other learning media that is more applicable. This study aims to develop mobile interactive e-learning media based on Cloud Speech-To-Text API at PGRI Pasir Sakti vocational high school. The method used is adopting the waterfall type system development life cycle. The material feasibility test was carried out based on expert validation by two English teachers, which resulted in an agreed category. Meanwhile, the ISO 25010 standard testing covers four aspects: functional suitability, performance efficiency, portability, and usability, which resulted in strongly agreed categories.*

**Keywords:** *e-learning, english skills, iso 25010, speech-to-text api, system development life cycle*

---



© Author (s)

## **INTRODUCTION**

English as an international language is vital to be mastered in the current era of globalization (Bayu, 2020). Good English language skills will undoubtedly be one of the assets to face global competition (Apsari et al., 2020). Realizing the importance of mastering English, learning English should be given and applied as early as possible to informal and non-formal education institutions (Bayu, 2020). English skills can be mastered with more interaction by listening, reading, speaking, and writing (Megawati, 2016). English subjects at the SMK level allocate 2 hours of study time in 1 week (Permendikbud, 2018). This time is often not effective for students, so students must study independently at home. However, students usually become bored and even bored if they have to learn at home, relying on textbooks.

The utilization of technology in education can be applied to the development of science for students (Saputra et al., 2017). One example of using technology in education is to create interactive learning media, which aims to generate motivation and stimulate learning activities (Saek et al., 2017). Interactive learning media can be developed by utilizing the Cloud Speech API (Application Programming Interface) developed by Google as an interactive medium for learning a language, especially in speech recognition into text, and vice versa (Hu et al., 2019; Iancu, 2019; Shakhovska et al., 2019). The voice recognition feature is essential to improve English speaking skills (Saputra et al., 2017), which will determine the suitability of the user's pronunciation with what it should be (Anggraini et al., 2018). Voice recognition implementations can take advantage of the Cloud Speech API service, a combination of the Google Translate API and Cloud Vision API.

In a previous study, a speak English chatbot application was developed using the Google Text-to-Speech API to translate sentences into sound, which obtained 63.96% of users satisfied and found it helpful in learning English (Afrianto et al., 2019). Another study developed a media application to correct Koran memorization by implementing the Cloud Speech API to obtain 100% accuracy in converting voice into text (Iancu, 2019). Finally, a study that also studied the Cloud Speech API succeeded in designing and implementing it into an Android application (Saputra et al., 2017). However, the application is limited to the pronunciation of messages that can only pronounce a word in English, so it cannot accommodate articulation in sentence form. Based on this, it is necessary to develop English interactive media that can accommodate sentence pronunciation by utilizing the Cloud Speech-to-Text API.

This study aims to develop interactive English learning media using the Cloud Speech-to-Text API. The research method applied is an adaptation of the Waterfall System Development Life Cycle (SDLC). The android application was developed using the Java programming language assisted by the Integrated Development Environment (IDE) Android Studio. As an implication, the application is expected to help increase students' learning motivation and improve the quality of English learning for schools by providing opportunities for students to be more independent and broaden their horizons.

## **METHOD**

The study area of this research was conducted at the Vocational High School (SMK) of the Indonesian Teachers Association (PGRI) Pasir Sakti, having its address at Merdeka No. 1 Street, Pasir Sakti District, East Lampung Regency. At the SMK, English subjects are guided by the 2013 Curriculum, which has basic competency standards to support the mastery and development of four language skills: listening, speaking, reading, and writing. Among these skills are linguistic elements such as structure or grammar, pronunciation, and vocabulary and the four skills above to gain comprehensive English language skills (Muttaqien, 2017). The material included only covers three skills in this study, namely listening, speaking, and reading. This research is a development of previous research that has not accommodated speech recognition of more than one word in speaking skills.

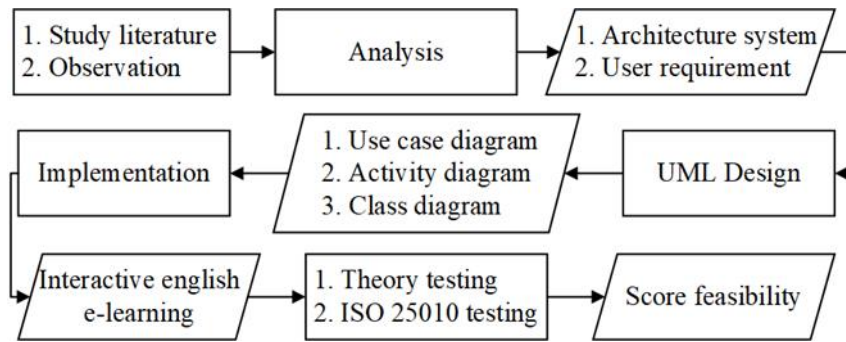


Figure 1. Research stage

The stages carried out in this study are adaptations of the SDLC waterfall. The main reason for adopting the waterfall method is a structure that can organize and control software development projects with accurate identification of user needs (Kramer, 2018). In general, this research includes six stages: planning, analysis, design, implementation, and testing, shown in Figure 1.

1. Planning

At this stage, the data needed for research is collected through literature study and observation (Rozi & Khomsatun, 2019). Literature study is done by reading literacy from scientific reading sources and books related to the research topic. To achieve the usefulness of the developed application, data collection was also carried out through direct observation to the principal and English teacher of SMK PGRI Pasir Sakti to obtain the required data. Based on the words, the English material used in making learning media in this study was limited to class X by referring to the English module of SMK PGRI Pasir Sakti.

2. Analysis

The observations obtained that the application needs include an attractive and fun display, can optimize the learning process and a single user. In addition, the learning media developed must be interactive so that students' English language skills can be further improved (Megawati, 2016). Based on this, a system architecture design was made, especially in applying Cloud Speech-to-Text API technology. The results of the system architecture design can be seen in Figure 2.

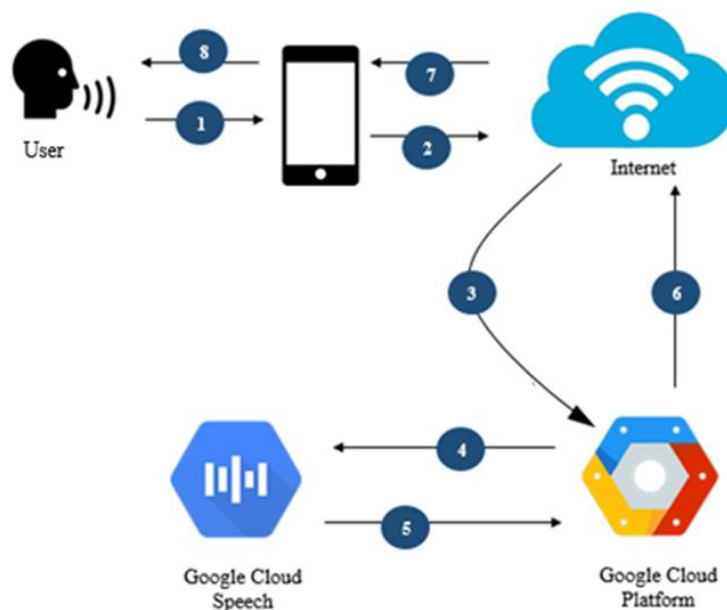


Figure 2. System architecture design

Figure 2 shows the required system process flow in detail as follows:

- a. The user pronounces the English sentence that he wants to assess via his smartphone
- b. The recorded user's speech will be transferred to Google Cloud Platform using the internet network
- c. Google Cloud Platform chooses the type of cloud speech service to process voice format into text format
- d. The sound design that has been processed is then returned to the Google Cloud Platform to send the results via the internet to the user's smartphone. The result is that the user gets a true or false rating based on the sentence that has been said before

### 3. Design

The design stage serves to describe and specify a product to facilitate the development of application products. The design stages are described in a Unified Modeling Language (UML) diagram, including use case, activity, and class diagram to illustrate the steps of the developed system process. The use case diagram describes the scenario of the relationship and interaction between the user and the system, shown in Figure 3 (a). Activity diagrams illustrate the flow of activities in the design, how each flow begins, the alternatives that may occur, and how each flow ends, shown in Figure 3 (b). At the same time, the class diagram describes the interaction between objects in and around the system (including users, displays, and so on) in the form of messages depicted against time, shown in Figure 4. Based on Figure 3, the system developed has only one user role, namely SMK students. Students can choose theory, exercises, or menus. In the theory menu, students can choose several materials based on adopting the English learning module at SMK PGRI Pasir Sakti. In the exercises menu, students can select the form of listening, reading, and multiple choice practice questions. In the About menu, students can view information about the system developer's biodata.

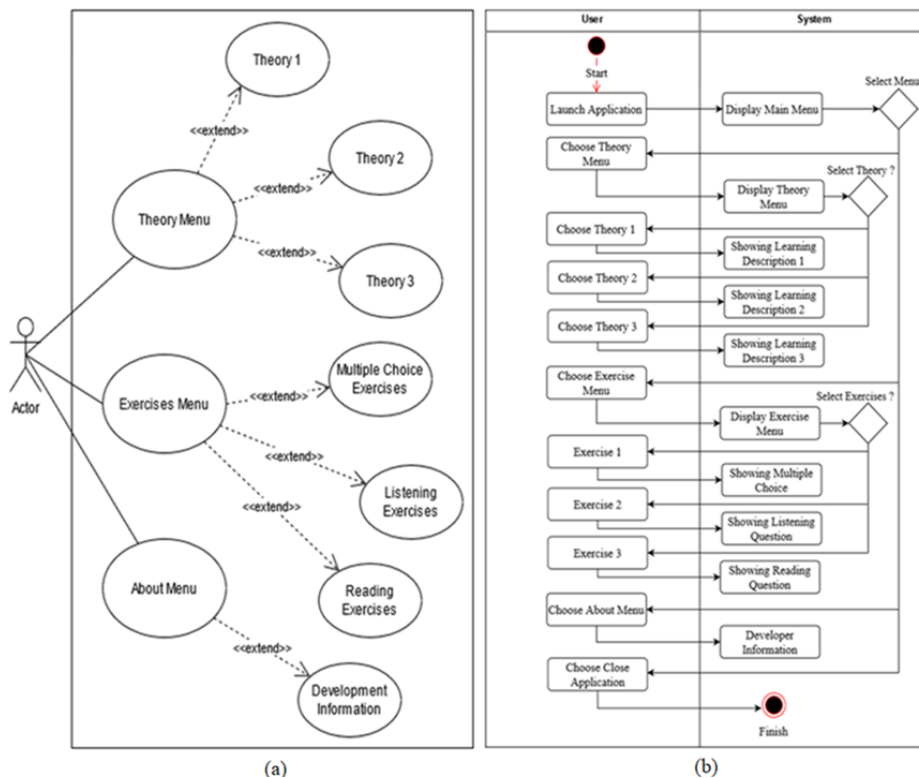


Figure 3. UML design of (a) use case and (b) activity diagram

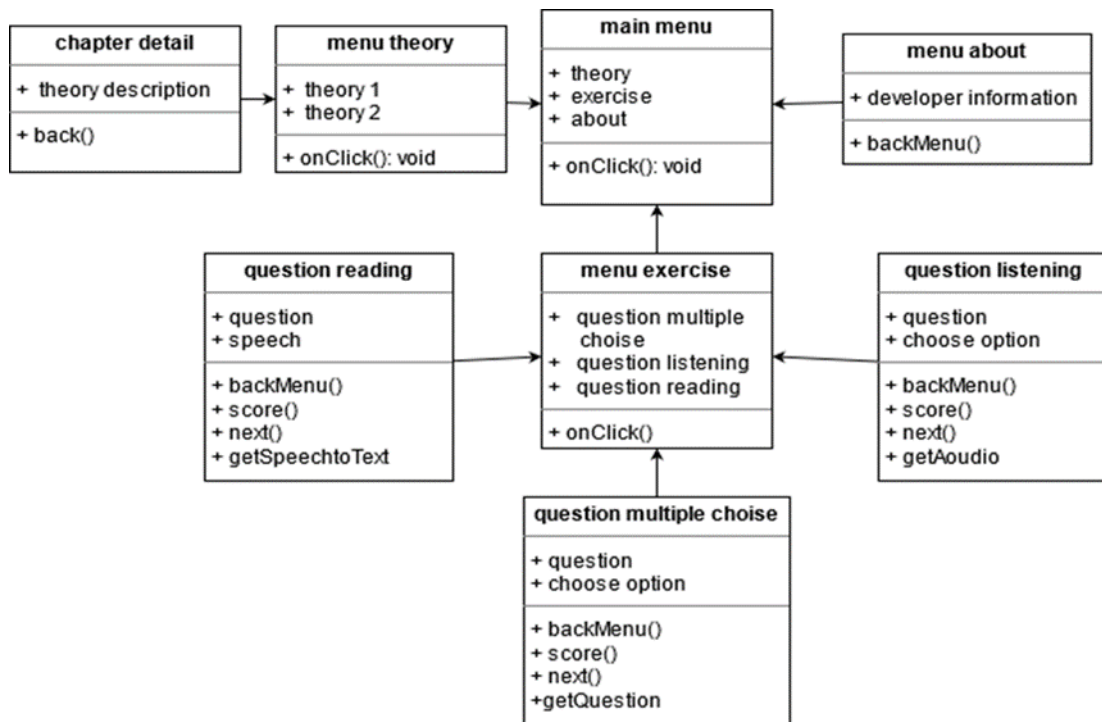


Figure 4. Class diagram

It can be seen in Figure 4, class diagram contains seven classes that are all connected to the main menu class. That is, the entire class can be accessed/displayed by running the main menu first.

4. Implementation

At this stage, the construction of the system based on the design is carried out by implementing it into the program code. The developed system uses several tools as follows.

- a. Java programming language version 8.0 for android application development
- b. Android Studio version 4.1 as programming language coding IDE
- c. Cloud speech-to-text API version v1p1beta1 as a library for sending audio and receiving transcription back in text form
- d. XML (Extensible Markup Language) version 1.0 as a markup language for interface design

5. Testing

At this stage, the material and application systems are tested to obtain optimal results based on the Likert scale. Material testing aims to determine the advantages and disadvantages of the material in the learning media developed. A material expert/validator carried out the test by giving a questionnaire to two Class X English teachers at SMK PGRI Pasir Sakti, namely Made Hengky, S.Pd., and Wahyu Eka Sumaryati, S.Pd. At the same time, system testing aims to determine the quality and feasibility of the developed application. System testing is carried out based on the ISO 25010 standard (ISO, 2013), covering four aspects, namely functional suitability, performance efficiency, portability, and usability. The selection of four of the eight aspects is an adjustment to the needs of the desired application (Rozi & Khomsatun, 2019). The following is the formula for calculating the testing percentage in Equation 1 (Nurkholis et al., 2021).

$$Testing (\%) = \frac{Actual\ score}{Ideal\ score} \times 100 \tag{1}$$

Table 1. Likert scale

No	Criteria Range (%)	Feasibility Score
1.	0-10	Very Not Agree
2.	21-40	Not Agree
3.	41-60	Neutral
4.	61-80	Agree
5.	81-100	Very Agree

The actual score is the result of the answers of all respondents from the questionnaire given. Meanwhile, the ideal score is the highest value from the questionnaire provided. The test results obtained are then calculated using Equation 1, followed by an interpretation of the system's feasibility based on the Likert scale to get conclusions and suggestions for future development. The following represents the Likert scale to assess the feasibility of the design listed in Table 1 (Awang et al., 2016).

**RESULTS AND DISCUSSION**

**1. Interactive English e-Learning**

The following are the results of the English learning media application interface that has been developed:

**a. Main Menu Interface, Theory, and Chapter Menu**

The main menu interface is an application display that contains a slide banner and theory, exercises, and about menu buttons. The main menu interface can be seen in Figure 5 (a). The theory menu interface contains five chapters containing explicit material, essential competencies, and practice buttons. The material interface can be seen in Figure 5 (b). The chapter interface includes a description of the learning, essential competencies contained in the syllabus. There is also a practice button that displays multiple-choice questions. The chapter interface can be seen in Figure 5 (c).

**b. Practice Menu Interface, Multiple Choice, and Listening**

The practice interface contains three different types of questions that can be selected by the user, multiple-choice types, listening questions types, and reading questions types. The Exercise interface can be seen in Figure 6 (a). The multiple-choice practice interface contains questions in text form; this question uses a random type. The user must choose an answer to continue with the following question; the question ends when the questions provided have run out, and the timer exceeds the set duration. The multiple-choice exercise interface can be seen in Figure 6 (b). The listening practice interface contains questions in audio form; this question uses a random type. Users must choose an answer and press the submit button first to continue to the next question. The listening interface can be seen in Figure 6 (c).

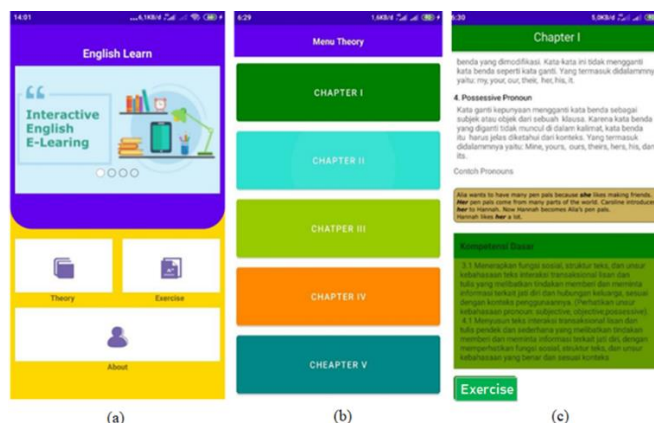


Figure 5. Interface of (a) main, (b) theory, and (c) chapter menu

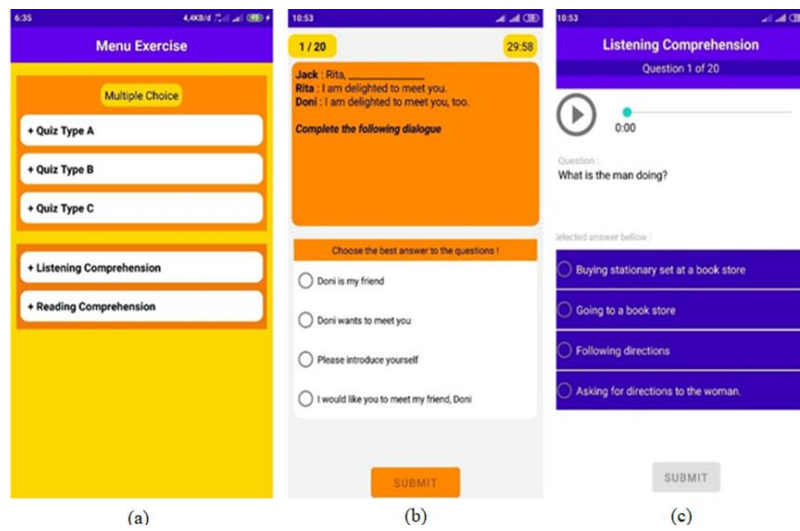


Figure 6. Interface of (a) exercise, (b) multiple-choice, and (c) listening menu

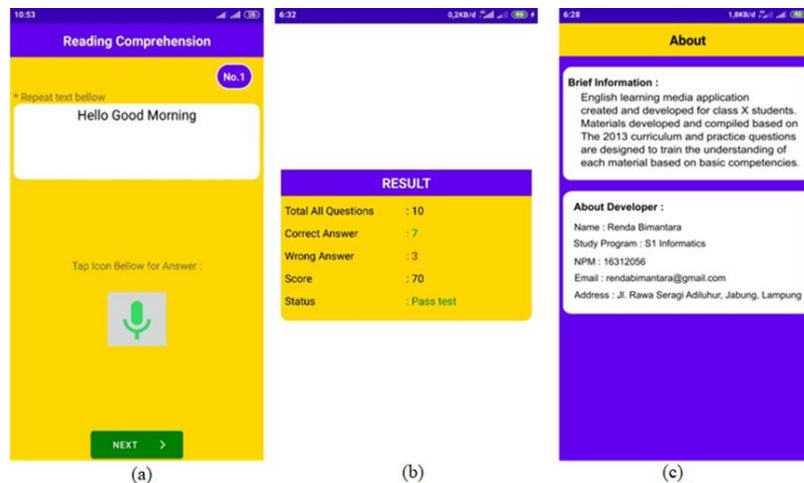


Figure 7. Interface of (a) reading, (b) score, and (c) about menu

c. Training Menu Interface Reading, Score, and About

The reading practice interface contains text-form questions; the user answers by pressing the microphone image and repeating the question sentence. The user is given three opportunities to answer questions. The Chapter interface can be seen in Figure 9 (a). The score result interface appears when the user finishes working on the questions. The score results page will display the number of questions, correct answers, incorrect answers, scores, and descriptions. The interface of the score results can be seen in Figure 9 (b). The developer interface contains application information and the identity of the creator or developer of learning media. The developer interface can be seen in Figure 9 (c). Figure 7 (a) is an example of implementing the Cloud Speech-to-Text API. When a question in a sentence appears on the smartphone, the user can answer by saying the correct English pronunciation. Then, the application will respond based on using voice recognition technology by the Google Cloud API. If it is accurate or appropriate, a message will appear that the pronunciation is correct. However, an error warning will appear if the accent is wrong, along with the pronunciation similarity score.



## 2. Testing Result

The following is a breakdown of the results of two tests that include material quality validation and the ISO 25010 standard.

### a. Theory

Material testing is done by giving a questionnaire containing 12 questions. The questionnaire obtained shows the percentage result of 79% using the calculation of Equation 1. Based on the Likert scale representation of the above analysis, it is found that the material in the learning media is in the agreed category. So, it can be said that the quality of the material is feasible to be given to students.

### b. ISO 25010

The application testing phase uses four aspects of ISO 25010: functional suitability, performance efficiency, portability, and usability. Here are the results of four elements of ISO 25010 testing:

- 1) Functional suitability testing is done by giving 22 function questions based on user needs analysis. The application test obtained a percentage of 100% based on calculations using Equation 1. Based on the results of these calculations, it can be concluded that the rate obtained by this system has met the aspects of functional suitability. So, it can be said that the operating system is very feasible to be used by users.
- 2) Tests were carried out with the Android Profiler tool found in Android Studio version 4.1. The test results show that the CPU usage used by the application is the highest at 7%, and the memory usage used by the application is the highest at 79.6 MB. With the use of CPU and memory resources, the application does not experience crashes and memory leaks.
- 3) Portability testing is done by testing the application on four versions of the Android OS, namely Lollipop, Marshmallow, Nougat, Oreo. The test is run on various OS with a minimum standard Android OS version of Lollipop. Based on the results of tests carried out on four different OS, it can be concluded that the application has a portability standard, which is 100%.
- 4) Usability testing is done by giving a questionnaire to 20 students of class X. The questionnaire showed consists of 15 questions. The test results obtained from student questionnaires obtained a percentage of 87.7%.

Based on the Likert scale of the four aspects of ISO 25010 testing, it was found that the overall application quality was in the category of strongly agree. So, it can be said that interactive English learning media is very feasible to use, especially for class X students, SMK PGRI Pasir Sakti.

## CONCLUSIONS

This study developed an interactive English language learning media based on cloud speech-to-text API for class X students, especially at SMK PGRI Pasir Sakti. The system was developed by adopting a waterfall type system development life cycle, which includes planning, analysis, design, implementation, testing stages. Testing the quality of the material aspects carried out by two English teachers of class X showed that the material presented obtained a percentage of 79% with a quality scale agreed. The ISO 25010 quality test results covering aspects of functionality stability, performance efficiency, portability, and usability obtained test results with a ranking of strongly agree or, in other words, very good.

## REFERENCES

- Afrianto, I., Irfan, M. F., & Atin, S. (2019). Aplikasi Chatbot Speak English Media Pembelajaran Bahasa Inggris Berbasis Android. *Komputika: Jurnal Sistem Komputer*, 8(2), 99–109. <https://doi.org/10.34010/komputika.v8i2.2273>
- Anggraini, N., Kurniawan, A., Wardhani, L. K., & Hakiem, N. (2018). Speech recognition application
-



- for the speech impaired using the android-based google cloud speech API. *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, 16(6), 2733–2739.
- Apsari, Y., Lisdawati, I., & Mulyani, E. R. (2020). Alat Permainan Edukatif Sebagai Media Pembelajaran Bahasa Inggris. *Abdimas Siliwangi*, 03(01), 38–47.
- Awang, Z., Afthanorhan, A., & Mamat, M. (2016). The Likert scale analysis using parametric based Structural Equation Modeling (SEM). *Computational Methods in Social Sciences*, 4(1), 13.
- Bayu, K. (2020). Penggunaan Google Translate Sebagai Media Pembelajaran Bahasa Inggris Paket B Di Pkbn Suryani. *Comm-Edu (Community Education Journal)*, 3(1), 62. <https://doi.org/10.22460/comm-edu.v3i1.3764>
- Hu, S., Liu, S., Chang, H. F., Geng, M., Chen, J., Chung, L. W., Hei, T. K., Yu, J., Wong, K. H., & Liu, X. (2019). The CUHK Dysarthric Speech Recognition Systems for English and Cantonese. *INTERSPEECH*, 3669–3670.
- Iancu, B. (2019). Evaluating Google Speech-to-Text API's Performance for Romanian e-Learning Resources. *Informatica Economica*, 23(1).
- ISO. (2013). ISO / IEC 25010 : 2011 Systems and software engineering - Systems and software Quality Requirements and Evaluation ( SQuaRE ) - System and software quality models. *ISO / IEC 25010 : 2011 Systems and Software Engineering*.
- Kramer, M. (2018). Best practices in systems development lifecycle: An analyses based on the waterfall model. *Review of Business & Finance Studies*, 9(1), 77–84.
- Megawati, F. (2016). Kesulitan Mahasiswa Dalam Mencapai Pembelajaran Bahasa Inggris Secara Efektif. *Child's Nervous System*, 32(9), 1709–1714. <https://doi.org/10.1007/s00381-016-3174-3>
- Muttaqien, F. (2017). Penggunaan Media Audio-Visual dan Aktivitas Belajar Dalam Meningkatkan Hasil Belajar Vocabulary Siswa pada Mata Pelajaran Bahasa Inggris Kelas X (Quasy Experiment: SMAN 8 Garut). *Jurnal Wawasan Ilmiah*, 8(1).
- Nurkholis, A., Susanto, E. R., & Wijaya, S. (2021). Penerapan Extreme Programming dalam Pengembangan Sistem Informasi Manajemen Pelayanan Publik. *J-SAKTI (Jurnal Sains Komputer Dan Informatika)*, 5(1), 124–134. <https://doi.org/10.30645/j-sakti.v5i1.304>
- Permendikbud. (2018). *Peraturan Menteri Pendidikan Dan Kebudayaan Republik Indonesia Nomor 36 Tahun 2018*.
- Rozi, F., & Khomsatun, K. (2019). Rancang bangun game edukasi pengenalan warna untuk pendidikan anak usia dini menggunakan adobe flash berbasis android. *Jurnal Ilmiah Penelitian Dan Pembelajaran Informatika*, 04, 12–18.
- Saek, F. A., Purba, K. R., & Purbowo, A. N. (2017). Pembuatan Media Pembelajaran Bahasa Inggris Menengah Atas Kelas X dengan Fitur Speaking dan Listening Practice Berbasis Android. *Jurnal Infra*, 5(2).
- Saputra, D. I. S., Handani, S. W., & Diniary, G. A. (2017). Pemanfaatan Cloud Speech Api Untuk Pengembangan Media Pembelajaran Bahasa Inggris Menggunakan Teknologi Speech Recognition. *Jurnal Telematika*, 10(2), 92–105.
- Shakhovska, N., Basytiuk, O., & Shakhovska, K. (2019). Development of the Speech-to-Text Chatbot Interface Based on Google API. *MoMLLeT*, 212–221.