



GOOGLE TRANSLATE'S SPEECH-TO-TEXT AS A PRONUNCIATION COACH? STUDENTS SHARE THEIR VIEWS

Wahyu Andi Pratama¹, Ashari², Erna Wardani³

Jenderal Soedirman University

¹wahyuandypratama@gmail.com, ²asharimintarno@gmail.com

³erna.wardani@unsoed.ac.id

Received: 2024/10/10

Accepted: 2024/09/01

Available Online: 2024/09/30

Abstract

This study explored students' perceptions of using Google Translate's Speech-to-Text feature for learning English pronunciation. Using a descriptive qualitative method, the research was conducted with seven purposively selected X TKJ 5 (Computer and Network Engineering) students at SMK HKTI 2 Purwareja Klampok during the 2023/2024 academic year. Data were collected through interviews, questionnaires, and observations, and analyzed using Miles and Huberman's qualitative approach. The findings revealed that students had predominantly positive perceptions. They found that Google Translate's Speech-to-Text was enjoyable, engaging, and more effective than traditional methods. Observational data confirmed students' enthusiasm, while 71.4% appreciated its low data consumption and all students (100%) agreed that Google Translate's Speech-to-Text effectively identified pronunciation errors and corrections. Additionally, students highlighted that Google Translate's Speech-to-Text was ease of use, flexibility, and user-friendly interface, affirming its value as an effective medium for learning pronunciation.

Keywords: *perception, pronunciation, Google Translate, Speech-to-Text.*

Introduction

Pronunciation is one of English basic skills that should be well mastered by the students. This statement is relevant with statement from Hidayatullah (2018), stated that English should be well-trained to be clearly and understandably communicated by speakers and listeners. because wrong pronunciation may lead to different meanings and confuse people. Some students in Indonesia have problems with pronunciation, because English still considered as a foreign language.



As we know, there are so many regional accents in Indonesia. Therefore, for some students, they might find difficulty in pronouncing English words. As stated by Suprayogi (2020), students still struggle to pronounce English words. It is because the feature of English phonology is different from Indonesian phonology. As a result, it influences the students in learning English pronunciation skill. Furthermore, most of the students were lazy to check the correct pronunciation of the words in the dictionary and the students cannot develop their speaking skill (Pratiwi & Ayu 2020). It might happen because they were not interested in using it.

As Numonjonov (2020) stated, traditional pronunciation learning methods are dominated by memorization and lack direct pronunciation practice that make tend to get bored. The teacher should offer students opportunities to argue to assess their comfort with pronunciation learning methods. Traditional methods rely on memorization, lacking direct pronunciation practice, leading to boredom. Alternative methods should be introduced.

Since Covid-19 in 2019-2021, hybrid learning became popular, encouraging students to use gadgets for learning and become more tech-savvy (Sakkir, Dollah, & Ahmad, 2021). Technology plays a crucial role in the learning process. Smartphones offer convenience for self-directed learning and accessing study materials. One of the advantages of technology that can be used as learning media is Google Translate.

Google Translate is a platform providing free multilingual translation founded by Google. Maulida (2017) stated that Google Translate can help us translate speech, text and images into various languages. In this case, Google Translate has a unique feature. It is speech to text detector that students can use to learn pronunciation.

Based on preliminary observations AT SMK HKTI 2 Purwareja Klampok, the students had difficulty in producing accurate pronunciation, lack of confidence in speaking English, and less interested in the pronunciation learning method that the teacher provided. The teacher said that once some students complained that they had difficulty in memorizing and pronouncing words because they only had English lessons once a month, so they have lack of chance. The researcher argued that students' perceptions on the use of Google Translate Speech-to-Text need to be analyzed to find out students' perception dealing with the learning media used

Theoretical Framework

Pronunciation

Pronunciation is the way a word in a language is pronounced to convey meaningful information. It is the production of sounds that aim to convey information or meaning (Yates, 2002 as cited in Gilakjani, 2016). It is a vital component of speaking, because failure in pronunciation can result in misunderstanding. This happens because the sound of speech in a language is a distinctive unit when pronounced differently can cause different meanings (Yule, 2022). They will be



able to communicate easily using English if they pronounce words well. Therefore, it is important to know their perceptions of learning pronunciation for the sake of better quality afterwards. Problems and learning of pronunciation should be discussed further.

Problems in pronunciation can be caused by several factor. They can be the difference in phonological characteristics, the length of learning period (Kemendikbud, 2022), and the teacher with the conventional learning method (Numonjonov, 2020). Many regional accents in Indonesia cause difficulties in producing accurate pronunciation. The length of the English subject taught is short. Conventional learning methods of pronunciation bore students.

Students should hear a word correctly before pronouncing it, since learning of pronunciation and listening is related (Damiati, 2007; Wilga, 2018). The listener is an important factor in communication because the relation between pronunciation production and perception of speech (Pennington & Bishop, 2009). Pronunciation includes two stages: receptive stage and productive stage (Frankel, 1984 as cited in Roberts and Cooke, 2009). Receptive stage involves listening to words, while productive stage involves the attempt to speak what students have learned before. This far, the students learnt pronunciation by pressing the microphone button in the application, record, and detect the pronunciation for accuracy.

Learning Media

Google Translate with its text-to-speech feature is one of the learning media available. It is a type of digital learning media of the three types of learning media. Hikmah (2019) classified learning media into three kinds: visual, audio, and audio visual. Visual media express messages through verbal communication symbols which need to be understood as a whole. It serves to attract students' attention, clarify the material presentation, and illustrate visualization. It is attractive but requiring more effort to prepare. Pictures, augmented reality, boards, and so on are examples of this type. In Audio media, messages are conveyed in the forms of auditive symbols, both verbal and nonverbal (Sadiman, et al., 2002). The benefits of it are its simplicity and portability, but children will find it boring due to no display. Tape recorders, podcasts, and conversations are examples of this type. Finally, audio visual media seems to be the combination of both previously mentioned media. It is a series of electronic images accompanied by audio sound elements included in the video recording (Anderson, 1994). It offers more complex and realistic messages, making students interested, but it needs more effort to prepare. The examples may be television, Youtube, movies, etc.

Technology development causes digital learning media to become prevalent with the existence of Google Translate Speech-to-Text for pronunciation learning. It is expected to ease the process of learning pronunciation. It is available on all gadgets owned by students. Overall, learning media plays a significant role in modern education by providing learners with diverse and interactive resources that foster effective leaning experiences.



Google Translate

Google Translate is a free multilingual translation service by Google. By it, students can translate text, voice, and images directly between languages (Maulida, 2017). It has four main features: an online dictionary, a thesaurus, a spell checker, and a pronunciation tool called speech-to-text and text-to-speech which can help students learn pronunciation. Google Translate provide several benefits such as its ability to help students learn pronunciation, portability, and fun learning experience (Duong, Tran & Nguyen, 2021).

Perception

Perception is an individual's response to something on their environment. It is the process by which individuals organize and interpret their sensory impressions to give meaning to their environment (Judge, 2015). It is categorized into several types and influenced by several factors. Its types are classified into positive and negative perceptions (Irwanto, 2002). If someone has positive perception, he/she will accept and support the perceived thing. When a person has negative perception, he/she will reject the perceived thing.

Furthermore, perception is influenced by some factors: perceiver, target, and situation (Robbins & Judge, 2019). Perceiver is the person who does perception related to attitude, motives, interests, experiences, and expectations. Attitude deals with what he feels. Three kind of it includes cognitive (knowledge about Google Translate Speech-to-Text), affective (feelings from the effects of using Google Translate Speech-to-Text), and behavior (reason for using Google Translate Speech-to-Text). Motive deals with the goal of taking action to achieve something. Interest is related to curiosity. Experience acts as a motive related to the way he/she receives knowledge. While expectation is his/her hope. Target deals with something being perceived in terms of its novelty, motion, sounds, size, background, proximity, and similarity. Situation is the context of the perception process, including time, work setting, and social setting

Method

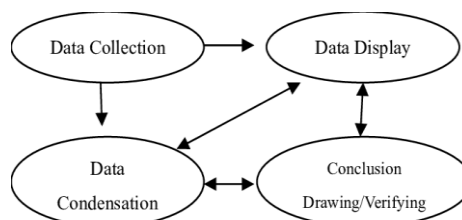
This Research employed qualitative method. Creswell and Poth (2016) state that qualitative research is based on the understanding and tradition of research methodologies that analyze social or human problems. In addition, the research design used was descriptive qualitative that focuses on the many different experiences the respondents get from a phenomenon of existing conditions that happen in the present (Creswell, 2014).

The research conducted in SMK HKTl (Himpunan Kerukunan Tani Indonesia) 2 Purwareja Klampok. The population of the research were X TKJ 5 class of SMK HKTl 2 Purwareja Klampok in academic year 2023/2024 which consists of 35 Students and 7 (seven) students with the lowest English Pronunciation scores selected as the sample.

Data collection techniques used were Observation, Questionnaire, and Interview. Observations made to all participants when the teacher is teaching, questionnaires have distributed to students via the Google form, and interviews will be conducted to students orally to participants in this research.

Furthermore, to analyze the data, the researcher used descriptive qualitative by Miles and Huberman (2014). To be more precise, there are three stages of data analysis proposed by Miles and Huberman (2014) that can be seen in the following picture.

Figure 1 Miles Huberman (2014)



In addition, this research will use triangulation of data collection method. Triangulation of data collection chose by the researcher through multiple sources such as Observation, Interview, and Questionnaire (Creswell, 2012). Furthermore, Sugiyono (2013) also stated that the triangulation of data collection method is a way of checking the validity of data or research findings.

Findings and Discussion

The research of “Students' Perception on The Use of Google Translate Speech-to-text When Learning English Pronunciation (Descriptive Qualitative Research of X TKJ 5 Grade Students at SMK HKTI 2 Purwareja Klampok in Academic Year 2023/2024)” was conducted from August 24th to September 8th 2023. The data were obtained from interview as primary data, questionnaire, and observation.

Interview

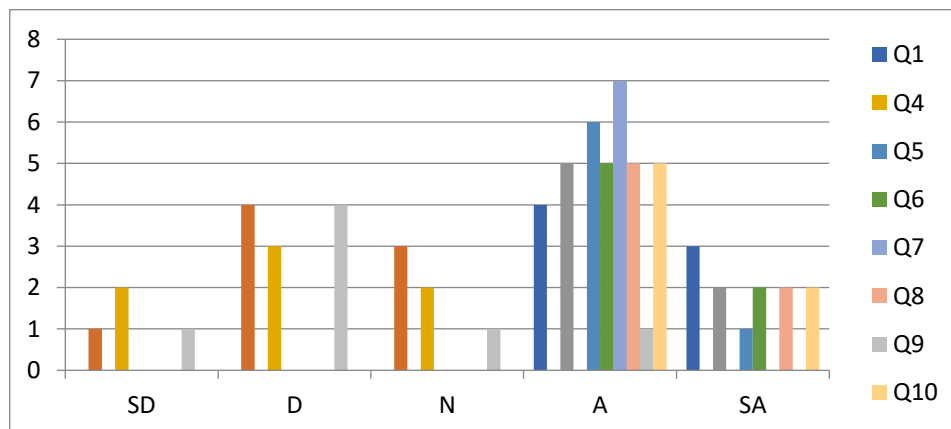
Interviews were conducted directly in class. The is to find out the detail students' perceptions of using Google Translate Speech-to-Text in learning pronunciation. it was done by five focuses. The first focus is about how the students learning pronunciation beforehand. All of the students have taught pronunciation, and almost of them taught by using conventional learning method. Only two students who had taught by speech dictionary such as Google Translate. They said that they felt get bored when learning using conventional learning method. Second focus is on students' perception after using Google Translate Speech-to-Text. They felt happy, comfortable, motivated and enjoy with this learning media. Those all experience happened because it has friendly user-interface and easy to use. Third focus deals with the benefits of Google Translate Speech-to-Text. There are some

benefits including: this media can help the students to produce accurate pronunciation since they can do self-evaluation through the speech-to-text feature, flexible to use anytime and anywhere, has friendly user interface, and consumes less of internet data. Fourth focus is about difficulties. In fact, there was no difficulties faced by the students. Fifth focus is on possibility of future use. All of the students said they will use this media in the future when they get new words that they do not know how to pronounce them.

Questionnaire

The questionnaire was distributed by the researcher to 7 students of X TKJ 5 class through google form. The link of the form given to the leader of the class, then he shared the form to WhatsApp group of X TKJ 5.

Chart 1 The summary of questionnaire result



Based on the chart above, in brief, all the students (samples) agreed that Google Translate Speech-to-Text was enjoyable (100%), easy to get and use (71.4%), proper to be used for learning pronunciation (100%), accurate and sensitive (71.4%), able to show correct and error answer (100%), able to improve pronunciation skill (100%), motivational (100%), flexible and portable (100%), efficient in data usage (71.4%), and new interesting way (100%). To conclude, this learning media is useful and beneficial for learning pronunciation.

Discussion

Subsequent to having presented the discoveries from the research instruments, it is necessary to discuss them. This discussion will be able to provide answers to the research questions. Therefore, it can provide theoretical and practical significances. It will be focused on the perceptions and benefits of the use of Google Speech-to-Text.



The use of Google Translate Speech-to-Text has a variety of usefulness perceived by students. It is interesting and motivated students in learning pronunciation. This finding supported by Prabawati, Asriati, and Asmayanti (2021) stated that an interesting learning media can stimulate students' motivation, feelings, and of course skills. Flexibility and simplicity are also the advantages of using Google Translate Speech-to-Text, of course those all made the students comfortable. This finding is linear with Riyana (2020) stated that audio-based learning media is simpler and does not need lot of preparation and possible to use anywhere. In addition, the students also said that learning media that using technology-based made them enjoy. It is reinforced by statement from Blair and Serafini (2014) explained that students feel comfortable in a dynamic, digital and social world. discussing about possibility of future use, the students revealed that they will use it in the future since it is future proof. It proved by Lieshout and Cardoso (2022) who explained that Google Translate is a multi-purpose tool that can be easily adapted to students' needs, interests and learning styles in learning pronunciation, so that this tool can be used continuously in the future. In other hand, discussing about the benefits, there are several benefits perceived by students. First, it is easy to use since it has friendly user interface. User interface has significant impact with the students experienced, since it is the bridge between students and the software. This statement corroborated by Senevirathne and Manathunga (2021), User interface (UI) design of e-learning is a point of interaction between user and computer software. Second, able to improve the students' pronunciation since it can be used anytime and anywhere without time limitation. Nagai (2007) has the same opinion that basically language learners have to imitate the model's pronunciation repeatedly. Third, able to do self-evaluation. The self-evaluation made them efficient in evaluating their erros. Hincks (2003) explained the obstacle in evaluating pronunciation is determining practical methods for evaluation. so, when students can do this thing, it helped them. Fourth, it is efficient and consumed less of internet data. And it made students easier to use this media. This finding related with Robiyansah (2022) who explained, the selection of effective online media is the key to the success of the learning process. Fifth, since learning media is one of external factor related to students' motivation, Google Translate Speech-to-Text indeed motivated them. The statement proved by Kurniawan (2019) who explained without considering the internal motivation that students may have, as an English teacher, which is an external factor for them, they can create learning activities that are able to attract them to continue practicing speaking, such as using technology-based learning media.

Conclusion

Using a descriptive-qualitative design, this research tried to find out the perception of X TKJ 5 Grades Students on Google Translate Speech-to-Text and the benefits of it. Observation, interview, and questionnaire were used to collect the data with seven students as samples. Descriptive analysis and descriptive analysis were used to analyze the data. The results of data analysis indicate that the students perceived



Google Translate Speech-to-Text in positive ways. Moreover, it was also revealed that Google Translate Speech-to-Text has various benefits.

For the students' perception there were several conclusions. First, since Google Translate Speech-to-Text is technology-based learning media, it is new interesting thing for students which made them enjoy and happy. Second, since it has speech-to-text feature, it able to do self-evaluation to make their pronunciation better. Third, it is very helpful because it has various useful features. Last, the students also said that they will use this in the future when they have new words and do not know how to pronounce it. In other hand, dealing with the benefits, there are some benefits of it, including: First, the students can use Google Translate anywhere and anytime without depended on pronunciation examples from teachers. Second, Google Translate speech to text can show whether their pronunciation correct or not. Third, Google Translate Speech-to-Text consumes less of internet data. This efficiency greatly facilitates them in learning pronunciation. Fourth, this media has friendly user interface that made students feel comfortable. And fifth, this media can motivate students in learning pronunciation.

To further provide significances, the researcher would like to put forward some suggestions. The suggestions will be advantages for teachers, students, and ministry of education. For the teachers, it is expected that if students often practice pronunciation so that they can produce accurate pronunciation and can become good English speakers. In the learning process, it would be better if the students use gadget. because gadgets have benefits that can be used for study such as learning pronunciation, not only as a means of communication and entertainment. For teachers, it would be better if teachers found more interesting ways for students to learn pronunciation such as Google Translate Speech-to-Text or other speech dictionaries. Moreover, it would be better if the teachers give instruction to students to learn pronunciation vocabulary related to their major, which is computers and networks. This is important with a purpose that when they enter the work industry, they have good English provisions related to their competence in the field of technology. For ministry of education, it expected if the ministry of education Inputting various learning innovations into the curriculum, Providing internet data assistance to schools that do not have Wi-Fi or far from internet connections so that students and teachers who are there can be integrated with modern learning, and make a policy regarding the use of gadgets in schools so that students have freedom in learning but still under teacher supervision.

References

- Anderson, Ronald.H. (1994). *Pemilihan dan Pengembangan media Video Pembelajaran*. Jakarta: Grafindo Pers.
- Blair, R., & Serafini, T. M. (2014). Using social media networks to engage students. *Systemics, Cybernetics, and Informatics. Integration of education*, 6(12), 28-31.
- Creswell, J. W. (2009). *Research Design: Qualitative, Quantitative, and Mixed Methods (3rd Edition)*. California: Sage Publication, Inc.
- Creswell, J. W. (2012). *Educational research*. Boston: Pearson.
- Damiati, H. (2007). *The correlation between listening skill and pronunciation accuracy: a case study in the first year of smk vocation high school pupita bangsa ciputat school year 2005-2006*.
- Duong, T. M., Tran, T. Q., & Nguyen, T. T. P. (2021). Non-English majored students' use of English vocabulary learning strategies with technology-enhanced language learning tools. *Asian Journal of University Education*, 17(4), 455-463. .
- Gilakjani, A. P. (2016). English pronunciation instruction: A literature review. *International Journal of Research in English. Journal of English Language Teaching and Learning (JELTL)*, 1-6. <https://doi.org/10.17507/jltr.0806.30>.
- Hidayatullah, M. S. (2018). Improving Students' Pronunciation through Western Movie Media. *Jurnal Bahasa (e-Journal)*, 3(1). <https://doi.org/10.30603/al.v3i1.381>.
- Hikmah, D. (2019). Media for language teaching and learning in digital era. *Journal of English Education and Linguistics (IJoEEL)*, 1(2), 36-41.
- Hincks, R. (2003). Speech technologies for pronunciation feedback and evaluation. *ReCALL*, 8 <https://doi.org/10.1017/s0958344003000211>.
- Irwanto, E. H. (2002). *Psikologi Umum: Buku Panduan Mahasiswa*. Jakarta: PT.Prenhallindo.
- Judge, T. A. (2015). *Perilaku Organisasi edisi ke-16*. Jakarta: Salemba Empat.
- Kemendikbud. (2022). *KEPUTUSAN MENTERI PENDIDIKAN, KEBUDAYAAN, RISET, DAN TEKNOLOGI NOMOR 56/M/2022 TENTANG PEDOMAN PENERAPAN KURIKULUM DALAM RANGKA PEMULIHAN PEMBELAJARAN*. Jakarta: Kemendikbud.
- Kurniawan, I. (2019). *Student's perception on the use of youtube as a learning media to improve their speaking skill*. Bengkulu: Bengkulu University.
- Maulida, H. (2017). Persepsi mahasiswa terhadap penggunaan google translate sebagai media menerjemahkan materi berbahasa Inggris. *Jurnal Saintekom*, 56-66. <https://doi.org/10.33020/saintekom.v7i1.21>.
- Miles, M. B., & Huberman, A. M. . (1994). *Qualitative data analysis: An expanded sourcebook*. California: Sage.
- Miles, Matthew & Huberman, Michael & Saldaña, Johnny. (2014). *Qualitative Data Analysis: A Methods Sourcebook (3rd Edition)*. Arizona: Sage Publications, Inc.
- Nagai, K. (2007). Differences of Pronunciation Practices: A Study of "Repeat with me" and "Repeat after me". *Journal of the Phonetic Society of Japan*, 11(2), 79 https://doi.org/10.24467/onseikenkyu.11.2_79.
- Numonjonov, S. D. (2020). Innovative methods of professional training. *ISJ Theoretical & Applied Science*, 1(81), 747-750. <https://doi.org/10.15863/TAS.2020.01.81.134>.

- Pennington, B. F., & Bishop, D. V. (2009). *Relations among speech, language, and reading disorders. Annual review of psychology*, 283-306. <https://doi.org/10.1146/annurev.psych.60.110707.163548>.
- Prabawati, A., St Asriati, A. M., & St Asmayanti, A. M. (2021). *THE STUDENTS' PERCEPTION OF THE ONLINE MEDIA USED BY TEACHER IN LEARNING ENGLISH. English Language Teaching Methodology*, 169-181. <https://doi.org/10.22219/jpbi.vxiy.xxyy>.
- Pratiwi, Z. F., & Ayu, M. (2020). *The use of describing picture strategy to improve secondary students 'speaking skill. Journal of English Language Teaching and Learning*, 1(2), 38–43 <https://doi.org/10.33365/jeltl.v1i2.603>.
- Riyana, C. (2012). *Media Pembelajaran. Jakarta: Direktorat Jenderal Pendidikan Islam Kementrian Agama* .
- Robbins, S. P., & Judge, T. A. (2013). *Organizational Behavior*. Toronto: Pearson Education Inc.
- Robbins, S.P. and Judge, T.A. . (2019). *Organizational Behavior Updated 18th Edition Global Edition. United Kingdom: Pearson*.
- Roberts & Cooke. (2009). *Authenticity in the adult ESOL classroom and beyond. Tesol Quarterly*, 43(4), 620-642. <https://doi.org/10.1002/j.1545-7249.2009.tb00189.x>.
- Robiyansah, M. (2022). *The Influence Of Online Learning Media Effectiveness, Self-Efficiency And Learning Motivation On Students'economic Learning Outcomes At Sma Negeri 38 Jakarta In The Era Of Pandemic Covid-19. Jurnal Pendidikan Ekonomi, Perkantoran, dan Akuntansi* 3(2), 34-50.
- Sadiman, A.S., Rahardjo, R., Haryono, A., and Rahardjito. (2002). *Media Pendidikan. Jakarta: PT. RajaGrafindo*.
- Sakkir, G., Dollah, S., & Ahmad, J. (2021). *E-learning in covid-19 situation: Students' perception. Journal of Education and Learning Innovation*, 1(1), 9-15.
- Sugiyono. (2017). *Metode penelitian kuantitatif, kualitatif, dan R&D. Bandung: Alfabeta*.
- Sugiyono, D. (2013). *Metode penelitian pendidikan pendekatan kuantitatif, kualitatif dan R&D. Bandung: Alfabeta*.
- Suprayogi, B. E. (2020). *Students' Perspectives Toward News Voiceover Activity in Pronunciation Class. In Twelfth Conference on Applied Linguistics. Jakarta: Atlantis Press* <https://doi.org/10.2991/assehr.k.200406.041>.
- Taka, S. D. (2019). *Teaching Speaking by Using Snake and Ladder Board Game. IDEAS: Journal on English Language Teaching and Learning, Linguistics and Literature*, 7(2). <https://doi.org/10.24256/ideas.v7i2.1021>.
- Walgito, B. (2010). *Pengantar Psikologi Umum Edisi ke-5. Yogyakarta: CV. Andi Offset*.
- Wilga, M. R. (2018). *Teaching foreign language skills. Chicago: University of Chicago Press*.
- Yule, G. (2022). *The study of language. cambridge: university press*.