

The Influence of Karaoke Singing on Students' Pronunciation Accuracy in Reading Aloud

Yadi Kusmayadi

Midwifery Study Program, STIKes Budi Luhur Cimahi, INDONESIA

yadikusmayadi1812@gmail.com

Article Information

Received:
14 November 2025

Revised:
11 December 2025

Accepted:
20 January 2026

Abstract

Technological advancement in information and communication allows learners to have more exposure to English. However, learners with a non-English background, such as Indonesians, have distinctive problems in speaking English, one of which is pronunciation. Integrating karaoke singing into English language teaching and learning is believed to influence learners' language acquisition, especially pronunciation significantly. This study aims to find the influence of karaoke singing on students' pronunciation accuracy in reading aloud. This study used a quantitative method through one experimental group involving 23 participants with 10 meetings of intervention of karaoke singing, a pre-test, and a post-test. The pretest and post-test were conducted using ELSA Speak application where the participants were required to read aloud 16 sentences. The participants' scores were analyzed using univariate and bivariate analysis. The later analysis through t-test provides a p-value of 0.0000 which is lower than the alpha value of 0.05, it leads to the conclusion that there is an influence of karaoke singing on students' pronunciation accuracy in reading aloud. Hence, it is suggested that teachers or lecturers apply karaoke singing intervention during the class activity to enhance students' pronunciation accuracy.

Keywords: *ELSA Speak, English Pronunciation, Implicit Learning, Karaoke Singing, Reading Aloud*

Cite this paper:

Kusmayadi, Yadi. (2025). The Influence of Karaoke Singing on Students' Pronunciation Accuracy in Reading Aloud. *JOINT: Journal of Inspirational English Language Teaching*, 3(2), 137-145.
<https://doi.org/10.33830/joint.v3i2.13807>

INTRODUCTION

The era of industrial 4.0 and society 5.0 have brought us to a new setting of environment which is becoming more global and digital from time to time (Angelianawati, 2019). Recently, English proficiency not only offers some benefits but also becomes necessary. Almost all the recent information, technology, and innovation are delivered or proposed in English. To compete in the recent global competitive market, every person needs to be able to access, understand, and apply the information described and explained in English. Furthermore, in a

more interconnected business atmosphere, the ability to communicate in English fluently becomes the key point of fostering international cooperation.

In academic sector, English works as a lingua franca (Anggraini, Sari, & Setiawati, 2023). Academic journals, conferences, and online courses mainly use English as their language which makes it obligatory for students, teachers, and researchers to master the language to have access to knowledge and also to contribute to the global academic world. The digital platform, dominated by English-based websites, social media, and online resources, emphasizes that English language proficiency is important. To browse and surf around those platforms, each individual needs to have good English proficiency.

More educational institutions integrate English into classroom activities and scientific publications. Thus, it is believed that by mastering English, individuals get more opportunities in terms of their education and career. However, learning English is not as easy as flipping hands in reality. According to Ahmad (2023), one of the five factors becoming difficulties for students to learn to speak English is pronunciation. The significant difference between what is written and the way how to pronounce it has become a distinctive obstacle for Indonesian students in learning English. Moreover, the students' pronunciation is highly affected by their local language. Furthermore, Jahara and Abdelrady (2021) stated that pronunciation problems encountered by non-native speakers were mostly sound reproduction. By highlighting the issue above, some students tend to choose not to speak due to lack of confidence to produce the correct English pronunciation (Khoiriyah, Waris, & Juhansar, 2019).

To highlight the challenges of pronunciation in English language teaching and learning, especially in terms of problems in producing sound among students, an effective teaching methodology is essential to cope with the problem. Harmer (2015) stated that pronunciation is an essential aspect of language learning and it needs to be integrated into all levels of the teaching and learning process. Thus, Harmer proposes a balanced approach that prioritizes fluency and accuracy in pronunciation. However, Harmer (1998) also mentioned that the goal of teaching pronunciation is not to make the students sound like native speakers, but to help them be intelligible. Additionally, Harmer (2015) suggests that using visual aids in teaching pronunciation as well as encouraging students to listen to native speakers and imitate their pronunciation is very important. In this era, due to the advancement of technology, the teachers, especially those who teach English as a foreign language, are forced to utilize and incorporate technological-based learning media to the classroom activity (Santiana, et al., 2024). The technology gives a wide array of option for teachers to develop or cultivate learning media widely available on the internet or in the software application.

Improving English pronunciation can be approached in various ways. One particular method is incorporating music into study routine. Because songs create a relaxed and engaging atmosphere, they allow learners to more easily master elements like rythm, word stress, and sentence structure. Beyond just the sounds, music helps students absorb new vocabulary and grasps the deeper meanings behind the lyrics (Panggabean & Batubara, 2022). Furthermore, Brewster and Girard (2004) mentioned that carefully selected songs, rhymes, and chants for language teaching context can offer a rich source of authentic input which cover the linguistic



resource, the psychological/affective resource, the cognitive resource, and the cultural and social resource.

Karaoke singing, when it is incorporated into language teaching and learning, can address effectively the aspects that are highlighted by Harmer. Embedding karaoke into language learning or combining both karaoke singing and language learning, gives a unique and engaging approach towards pronunciation practice such as following: first, karaoke singing hones the learners' fluency by encouraging them to continuously produce language they learn and drills them rhythmically. The repetitive nature of the song, especially in the refrain part, allows the learners to practice their pronunciation patterns and improve their overall fluency, as well as enjoy the singing. The second, karaoke singing requires the students to maintain their accuracy in terms of the pitch and rhythm which are more into musical aspect, and also intonation and lyrics which are more into the language aspects. Both aspects provide an assist to learners in refining their quality of pronunciation and also in developing an accurate sound production.

Karaoke singing also offers a valuable opportunity for students to imitate the native speakers they hear from the song. Thus, by listening and singing along with the native speakers, in this case, the vocal of the song, the learners can add more exposure to the authentic pronunciation of particular words and develop their ability to mimic the target language sounds. Karaoke singing also provides a relaxing pronunciation exercise, without fear of judgment from other learners regarding their ability in pronunciation. The fun and interactiveness of karaoke singing can also motivate learners to practice pronunciation regularly and enjoy the process of language learning without feeling forced to do so.

Furthermore, according to Karsono (2019), integrating music and songs in language teaching and learning processes emits several positive outcomes. First, music and songs can promote language acquisition in learners. It is beneficial for the enhancement of their language acquisition. Second, music and songs can assist teachers in establishing a more creative and imaginative learning environment. By combining various activities involving these elements, teachers can stimulate students' creativity and critical thinking skills. Third, music and songs give a new approach to assisting the development of different language proficiencies. They can be used to hone language proficiencies such as listening, speaking, reading, and writing in an interactive and fun way. Fourth, embedding music and songs in language teaching and learning promote cultural awareness. By exploring music and songs from different cultures, students will be able to acquire a deeper comprehension and appreciation of diverse traditions and values. And the last, music and songs can provide an innovative and interesting way for language instruction. By incorporating the elements above, teachers can establish a more dynamic and stimulating learning environment.

Therefore, this karaoke integration in the class is the means of adding more exposure to students as the stimulus-response process. Stimulus-response learning is evident in the acquisition of the sound system of a foreign language in which, through a process of conditioning and trial and error, the learner makes closer and closer approximations to native-like pronunciation. Simple lexical items are, in one sense, acquired by stimulus-response



connections; in another sense, they are related to higher-order types of learning (Brown, 2000). Regarding this karaoke integration into learning language, in addition, some studies also mentioned that singing is one of the effective ways of improving pronunciation. Karsono (2019), in his study, revealed that karaoke activities significantly affected the tenth graders' English pronunciation. Finally, considering the prospective reason of the underlying discussion, this study wanted to test and prove the hypothesis of whether or not karaoke singing has an impact on students' pronunciation by measuring their reading-aloud proficiency.

METHOD

This study employed an experimental design with a pretest and post-test format which focused solely on an experimental group without a control group. The participants consisted of 23 students of the Midwifery study program Diploma in STIKes Budi Luhur Cimahi. This study was conducted over approximately three months, starting from August 23, 2023, to November 18, 2023 or twelve meetings which include the pretest at the beginning and also a post-test at the end of the sessions.

To assess the participants' English pronunciation accuracy before and after the treatment, the pretest and post-test were administered by using a mobile application called ELSA Speak. This ELSA which stands for English Language Speech Assistant is a unique and interactive application that is specifically designed to enhance the English communication skills of its users. This platform can identify pronunciation mistakes and offer improvement through detailed feedback and suggestions using Artificial Intelligence (AI) (Elsani et al, 2023 in Sari et al. 2023).

The ELSA Speak application is highly recommended for students to improve their speaking skills because it provides several facilities and features, including speech recognition technology, a custom, tested curriculum, a free online dictionary, pronunciation and intonation checker, and free assessment exams throughout the program (Karim et al., (2023). The application also incorporates a feature that allows the users to participate in an diagnostic test which enables them to test their speaking performance. This assignment test consists of reading aloud 16 stages of comprising sentences provided by the ELSA Speaking application. These sentences were carefully selected to represent a variety of grammatical structures, vocabularies, and pronunciation challenges such as intonation, articulation, and fluency commonly encountered by English language learners. By engaging with these sentences, users were exposed to a broad range of English-speaking contexts and had the opportunity to practice their pronunciation, intonation, and fluency.

For each stage, users need to press the mic icon and read the provided sentence aloud, after that, they need to press the 'Submit & Continue' button to move forward to the next stage. After completing the 16 stages, after pressing the 'Submit & Continue' button, the app's user interface will bring the users to their score page which is shown in percentage score, and their proficiency level which is shown as beginner, intermediate, and advanced. The result score page also consists of the detailed score for three different aspects of speaking: pronunciation,



intonation, and fluency, however, this study only took the overall score shown in the score page.

Meanwhile, the ten meetings of treatment were in the form of karaoke singing which was integrated into the English for Midwifery course. The treatment took fifteen minutes before the class ended. This treatment also worked as the 'cooling down' phase of the whole teaching and learning process of that English for Midwifery course. The treatment was in the form of karaoke singing as a method of drilling to make the students familiarize themselves and get used to English pronunciation. The songs used were three songs with different paces: slow, medium, and rather fast pace. The first song was 'One Twenty-One Guns' from a band called Green Day. It is a slow-paced song that was used in treatment meetings one to three. The next song was a song called 'Something Just Like This', a rather average pace song from Coldplay which was used in treatment meeting four to six, and the last was a song called 'Memory' from a band called Maroon Five for the fast pace song which was used during the rest of the treatment meetings.

The procedure of treatment in each meeting required the students to sing along to the songs mentioned earlier. The song itself was shown on a screen using a projector, and the song was taken from YouTube. The karaoke singing drill was in two different steps. In step 1, the students sang by using the music video consisting of the song with the vocals from the band and the lyrics itself. Meanwhile, in step 2, the students sang by using the karaoke version of the music video which only showed the lyrics and the music, without the vocal (minus one).

FINDINGS AND DISCUSSION

The pretest scores are the ELSA Speak students' overall scores before the intervention of karaoke singing, which exhibited a range of 26 to 70 scores. The lowest pretest score was 26 indicating a relatively low level of English pronunciation proficiency, while the highest pretest score of 70 suggested a higher level of English pronunciation proficiency. On the contrary, the post-test scores which reflect the students' ELSA Speak overall score after having the karaoke singing intervention demonstrate a range of 41 to 76 scores. Initially observed, the lowest post-test score of 41 is higher than the lowest pretest score which suggests that the intervention had a positive influence on the students who initially struggled with their pronunciation. The highest post-test score of 76 surpasses the highest pretest score which indicates there is significant improvement in pronunciation skills for some participants.

With a sample size of 23 participants, the study provided a moderately sized dataset for analysis. This sample size allowed for a reasonable analysis of the distribution of pretest and post-test scores. The observed range of pretest score which ranges from 26 to 70, and post-test scores which range from 41 to 6, indicates that the sample included students with varying levels of initial English pronunciation proficiency. This presence of a diverse range of scores within the samples is a positive aspect of the study. It shows that the findings are not solely influenced by a small group of students with similar levels of proficiency. Instead, the results are likely to be more representing the broader population of English language learners. This diversity also



strengthens the internal validity of the study as it reduces the likelihood of confounding variables that affect the outcomes.

The variety and gap between the lowest and the highest in pretest scores also have several implications. The first implication is heterogeneity. This heterogeneity may affect the effectiveness of the intervention or the treatment. The second implication is the individual differences. Understanding this individual difference is essential in developing or designing the intervention or treatment that is required to meet the expected outcome. The last implication is defining further statistical analysis which covers the univariate which aims to describe the characteristics of each variable and bivariate analysis to identify the relationship between two variables (Sugiyono, 2013).

The pretest and post-test scores were later analyzed using the univariate analysis through the SPSS application. This analysis was aimed to give an understanding of individual variables. The following table shows the result of the univariate analysis.

Tabel 1. Students Pronunciation Accuracy Before and After the Intervention of Karaoke Singing

| Variable | Mean | SD | Min-Max | 95% CI |
|---------------------|-------|--------|---------|---------------|
| Before Intervention | 47.70 | 12.256 | 26-70 | 42.40 – 53.00 |
| Adter Intervention | 57.30 | 9.842 | 41-76 | 53.00 – 61.56 |

Based on Table 1 above, the result of the univariate analysis shows that the students' pronunciation accuracy in reading aloud before the intervention of karaoke singing was 47.70 with a 95% confidence interval score ranging from 42.40 to 53.00. The standard deviation was 12.256 with a minimum score was 26 and a maximum score was 70. After being given the karaoke singing intervention, the average score became 57.30 with a 95% confidence interval score ranging from 53.05 to 61.56 with a minimum score of 41 and the maximum was 76.

The gap in the mean score indicates that there is an increasing trend and it means that there is a positive indication that the intervention or treatment was successful in enhancing the students' pronunciation accuracy in reading aloud. Meanwhile, the pre-intervention standard deviation score of 12.256 shows that there is a significant variation in scores among the students before the interventions. The decreasing score of standard deviation on post-intervention into 9.842 indicates that the variation among the participants is getting smaller which means that the treatment or intervention was effectively spread in balance among the students which is also indicated by the reduction of the minimal and maximal range of the pretest and post-test score. The increasing score of 95% confidence interval ranging from 42.40 to 53.00 before the intervention and 3.05 to 61.56 after the intervention also indicates that the intervention has a significant impact on students' pronunciation accuracy in reading aloud.

Although the result of the univariate analysis shows a positive result, further analysis was conducted to confirm the finding. Therefore, the bivariate analysis was conducted. This type of analysis was also meant to find the relationship between two variables. The normality test in this analysis showed a significant score of 0.576 on the pretest and 0.0146 on the post-test. Those scores show that the distribution of the data is normal, therefore, the further analysis



used was bivariate through t-test analysis using the SPSS application and gave the following score in the table below.

Tabel 2. The Influence of Karaoke Singing on Students' Pronunciation Accuracy in Reading Aloud

| Variable | Mean | SD | Mean Rank | P Value |
|---------------------|-------|--------|-----------|---------|
| Before Intervention | 47.70 | 12.256 | | |
| Adter Intervention | 57.30 | 9.842 | 9.60 | 0.000 |

Based on Table 2 above, the presentation of the average score of students' pronunciation accuracy before and after getting the karaoke intervention is still similar to the dataset shown in the univariate analysis which is 47.70 for the pretest and 57.30 for the post-test. Two new datasets are added to the table, with a mean rank score of 9.60 and a p-value of 0.0000. The 9.60 mean rank score shows that the students' pronunciation accuracy in reading aloud after getting intervention is higher than that before the intervention, thus, there is an increase in the score pronunciation. The p-value of 0.0000 indicates or supports that the student's performance in pronunciation after having the intervention is significantly different from the pronunciation score before having the intervention. Therefore, the statistical analysis above shows that there is an influence of karaoke singing on students' pronunciation accuracy in reading aloud. This is in line with the claim of Karsono (2019) that the integration of music and songs in teaching and learning can promote language acquisition, as well as help teachers create the environment of teaching and learning.

In English language teaching and learning, the concept of stimulus and response is essential. The concept reflects the dynamic interaction between both teachers and students. The stimulus triggers responses from the students which covers the teaching and learning process. Meanwhile, the drilling method involves repetition in practice to let the learners learn by themselves and reinforce self-correction. Both concepts, stimulus and response and drilling method, are brought into the class in terms of karaoke singing, where the stimulus is the song and the response is the learners' singing activity. This approach not only turns the sometimes-tedious drilling method into a more enjoyable and beneficial one for the learners by exposing pronunciations implicitly. That process of karaoke singing was repeated during the 10 intervention or treatment meetings, 10 to 15 minutes before the English class was dismissed. This was meant to familiarize and make the students used to singing in English and surely in line with the behaviourism theory Brown (2000) mentioned that language teaching practices likewise for many years were influenced by a behaviouristic tradition. The resulting score of students' pronunciation using the ELSA Speak application shows increasing performance.

Furthermore, by using karaoke singing, students were unaware that they were in the middle of learning pronunciation, as they eagerly sang to only make themselves happy, and this is called implicit learning. Implicit learning refers to learning without intention. It is also defined as learning without awareness—and is seemingly ubiquitous in daily life (Frensch & Runger, 2003). In implicit learning, a permanent modification of the organism's activity



resulting from its interactions with the external environment in the absence of intention and awareness about the learning materials and methods (Nejati, 2022).

CONCLUSION

Statistical analysis of the study confirms that karaoke singing significantly enhances students' pronunciation accuracy during reading aloud tasks. The bivariate t-test yielded a p-value of 0.0000, below the alpha = 0.05 threshold, leading to the rejection of the null hypothesis in favor of the alternative hypothesis. Beyond quantitative data, the effectiveness of this intervention is rooted in implicit learning, a process where acquisition occurs without conscious effort. This approach not only fosters deep engagement and motivation but also aligns with Harmer's (2015) pedagogical standards, which advocate for the seamless integration of pronunciation across all proficiency levels to ensure comprehensive language development.

From practical standpoint, karaoke serves as a multifaceted instructional tool that simultaneously addresses fluency, accuracy, and imitation. By requiring students to match the rhythmic and melodic patterns of a song, it forces continuous language production and adherence to native-like pitch and intonation. Furthermore, karaoke provides a low-stakes, supportive environment where learners can mimic authentic models and experiment with sounds without the typical anxiety of formal drilling. By bridging the gap between enjoyment and technical practice, karaoke transforms pronunciation instruction into a dynamic, interactive experience that encourages regular practice and long-term skill retention.

REFERENCES

Ahmad, Sulaiman. (2023). *Analisis Siswa dalam Berbicara Bahasa Inggris di MAS Mulia Balai*. JIP (Jurnal Ilmu Pendidikan) 1, 1044-1050. <https://jip.jln.org/index.php/pendidikan/article/view/110>.

Angelianawati, L. (2019). *Being an English Teacher in Industrial Revolution 4.0: an Overview about Roles, Challenges, and Implications*. Jurnal Dinamika Pendidikan 11(3):307. DOI: 10.33541/jdp.v11i3.896

Anggraini, Y., Sari, I. P., & Setiawati. (2023). *Enhancing English language learning through mind mapping: A case study of students' engagement and effectiveness*. J-Shelves of Indragiri (JSI), 5(1), 98–110. DOI: <https://doi.org/10.61672/jsi.v5i1.2660>

Brewster, J., Ellis, G., and Girard, D. (2002). *The Primary English Teacher's Guide*. England: Penguin English.

Brown, H. Douglas. (2000). *Principles of Language Learning and Teaching* - 4th ed. New York: Pearson Education.

Elsani, Emilia., Salasbila, Runi., Putra, Muhammad Fikri Indika., Nabila, Naba Khilma., & Nahartini, Desi. (2023). *The Effect of Using Elsa Speak App for First-Semester Students' English Speaking Proficiency*. Edukatif: Jurnal Ilmu Pendidikan 5(6): 2644-2655. DOI: <https://doi.org/10.31004/edukatif.v5i6.5976>

Frensch & Runger. (2003). *Implicit Learning*. Current Directions in Psychological Science 12(1):13-18 DOI:10.1111/1467-8721.01213

Harmer, Jeremy. (1998). *How to Teach English*. Essex: Pearson Education Limited.



Harmer, Jeremy. (2015). *The Practice of English Language Teaching*. London: Pearson.

Jahara, Syed Farhat & Abdelrady, Abbas Husein. (2021). *Pronunciation Problems Encountered by EFL Learners: an Empirical Study*. AWEJ (Arab World English Journal) 12(4): 194-212. DOI: <https://dx.doi.org/10.24093/awej/vol12no4.14>

Karim, S.A., Hamzah, A.Q.S., Anjani, N.M., Prianti, J., & Sihole, I.G.(2023). Promoting EFL Students' Speaking Performance through ELSASpeak: An Artificial Intelligence in English Language Learning. JOLLT Journal of Languages and Language Teaching, 11(4), pp.655-668. DOI: <https://doi.org/10.33394/jollt.v%vi%i.8958>

Karsono. (2019). *The Effect of Using Karaoke Activities on Students' English Pronunciation*. JET (Journal of English Teaching) 5(3):179. DOI: 10.33541/jet.v5i3.1311

Khoiriyah, H., Waris, A. M., & Juhansar. (2019). *The Students' Achievement in Pronouncing English Song Using Smule Application*. IEFLJ (Indonesian EFL Journal) 5(1): 45-57. <https://journal.uniku.ac.id/index.php/IEFLJ/index>

Nejati, Vahid. (2022). *Principles of Cognitive Rehabilitation*. Science Direct DOI: <https://doi.org/10.1016/C2021-0-03320-4>

Panggabean, Hanna Maria & Batubara, Lika Silvia. (2022). *Utilizing a Karaoke App in Learning English Pronunciation: a Case Study of EFL College Students*. IJEAL (International Journal of English and Applied Linguistics) 2(1): 103-111. DOI: [doi.org/ijean.v2i1.1432](https://doi.org/10.33394/ijean.v2i1.1432)

Santiana, et. al. (2024). *Yuk Belajar Bahasa Melalui Media dan Sumber Belajar di Era Society 5.0*. Tulungagung: Akademika Pustaka. Retrieved from https://www.researchgate.net/publication/378234494_YUK_BELAJAR_BAHASA_MELALUI_MEDIA_DAN_SUMBER_BELAJAR_DIERA_SOCIETY_50

Sugiyono. (2013). *Metode Penelitian Kuantitatif, Kualitatif, dan R&D*. Bandung: Penerbit Alfabeta.

