Indonesian Journal of English Language Teaching and Applied Linguistics

Vol. 8(2), November 2023
www.ijeltal.org
e-ISSN: 2527-8746; p-ISSN: 2527-6492

The Use of Artificial Intelligence to Promote Autonomous Pronunciation Learning: Segmental and Suprasegmental Features Perspective

Senowarsito[1], Sukma Nur Ardini[2]

[1] *Universitas PGRI Semarang, e-mail: senowarsito@upgris.ac.id*
[2] *Universitas PGRI Semarang, e-mail: sukmanurardini@upgris.ac.id*

ARTICLE INFO

Keywords:
Artificial Intelligence, Autonomous learning, higher education, segmental features, suprasegmental features

DOI:
http://dx.doi.org/10.21093/ijeltal.v8i2.1452

ABSTRACT

The study aimed at investigating the effects of autonomous pronunciation learning using AI as well as the experiences of autonomous pronunciation learning using AI by higher level students. Explanatory sequential mixed-method research using both quantitative and qualitative methods was employed within thirty-two students from Universitas PGRI Semarang's first-year students serving as the sample. Assessments, interviews, and an evaluation of instructional materials were used as the instruments. Through pre- and post-testing, quantitative analysis was used to evaluate the students' pronunciation proficiency. Quantitative data analysis was done using SPSS. However, a qualitative analysis was used to review the interview. To bolster the findings of the tests, it was descriptively examined. After the treatments using an AI based application named ELSA, there was a significant correlation between the use of AI and autonomous pronunciation learning. However, ELSA has certain shortcomings. It appears to be primarily concerned with segmental than suprasegmental features. Only intonation is available from among all the features offered to practice suprasegmental features. While students found it difficult to emphasize words, there is no other practice for suprasegmental qualities. In reality, the ELSA website states that its curriculum covers core English skills such as word stress, intonation, rhythm, listening, and conversation. As a result, the ELSA creator may take this criticism into consideration as they continue to improve their product. It implies that the creator is responsive to the concerns or suggestions of their customers or users, which can contribute to the ongoing development and success of the product.

How to cite:
Senowarsito, S. & Ardini, S.N. (2023). The Use of Artificial Intelligence to Promote Autonomous Pronunciation Learning: Segmental and Suprasegmental Features Perspective. *Indonesian Journal of English Language Teaching and Applied Linguistics, 8*(2), 133-147

1. Introduction

Artificial Intelligence (AI) in English as a Foreign Language (EFL) learning has gained significant attention and is being increasingly integrated into language education (Jiang, 2022; Layali & Al- Shlowiy, 2020). The field of AI technologies, which is flourishing, is ideal for use in education. A book by Tuomi (2018), stated that,

> "AI is now often called "the next electricity." The transformative impact of general-purpose technologies, like AI, however, becomes visible only gradually, when societies and economies reinvent themselves as users of new technologies. Technological change brings social and cultural change that is reflected in lifestyles, norms, policies, social institutions, skills, and the content and forms of education."

It highlights the idea that AI has the potential to be as transformative for society. AI is poised to bring about significant changes in how we work, live, and interact with the world. When it comes to pronunciation, technology has played a role in transforming how we interact with language and communication. AI has been employed in various ways to assist individuals in learning and improving their pronunciation skills in a foreign language or even in their native language (Rusmiyanto, et al., 2023; Huang & Moore, 2023). Researchers and developers have leveraged AI-driven technologies to create tools and applications that provide personalized feedback and guidance for pronunciation improvement.

The study is based on earlier research that was conducted by the researchers and focused on Indonesian high school and higher education students' pronunciation proficiency (Al-Fisyar et al., 2020; Ardini et al., 2016; Rifqiyah et al., 2021; Senowarsito & Ardini, 2019). Before proceeding to a critical assessment, the current study provides an examination into the utilization of a specific, hyped-up, and well-liked AI. Zhang & Yin (2009) contend that pronunciation teaching should start in the primary grades since children absorb the sounds system more quickly and effectively than adults due to their age. In reality, this urgency didn't receive as much attention or support.

Tahereen (2015) asserts that pronunciation is ignored right away because it is a talent that is rarely practiced in both the classroom and outside of it in many countries. Additionally, in Indonesia, where absence of a pronunciation background causes teachers to place a greater emphasis on the value of grammar and writing proficiency. While the authority of native English teachers as the protector of accuracy and intelligibility still remains powerful, learners are increasingly being taught by non-native English speakers in English language instruction (Phuong, 2021). Additionally, the majority of bilinguals' use of English as a second language has an impact on their first languages (Gashimov, 2023). The educators should be aware, though, that non-native English speakers frequently mispronounce terms during instruction, such as "focus," which they sound as /fokjus/, 'determine' as /ditə(r)main/, and 'target' as /tardʒet/. As a result, they are not always the best judges of whether students' pronunciation is understandable or not. In brief, since their mother tongue, Bahasa Indonesia, functions in this way, non-native English teachers—especially those from Indonesia—remain to read literally as they see. Due to the significance of pronunciation as a component of English, non-native English teachers need to follow the pronunciation of native English speakers as a

model. If not, they will mislead the children and cause them to pronounce words incorrectly for the rest of their life. As a result, this problem should not be ignored.

On the other hand, higher education students must study independently since they are said to be digital natives because they are literate in using technology in the same way that they breathe. Regarding this issue, AI technology is one that experts say has great potential for the future of education (Chassignol et al., 2018). Higher education students now have the opportunity to master pronunciation on their own through directed and supervised practice. In light of this context, the goal of this study is to investigate the effects of autonomous pronunciation learning using AI as well as the experiences of autonomous pronunciation learning using AI by higher level students.

2. Literature Review

2.1 Autonomous Pronunciation Learning to Higher Education

Autonomous learning is how students take control of their learning responsibilities such as establishing learning objectives, arranging learning content, choosing methods and techniques, monitoring the process, and evaluating the result (Holec & Council of Europe., 1979). It is suitable with the current developed technologies in education. Some studies had discovered that it was found useful to use digital technology as tools for promoting pronunciation learning autonomously (Banafa, 2008; Carlet & De Souza, 2018; Kruk, 2012; Sariani et al., 2022). When Banafa (2008) looked into pronunciation practice, she found that using computers was extremely beneficial since they offered a secure environment for learning. According to Kruk's research, the experimental group that was given computer-based work was able to do better on tests than the control group and showed more autonomy. In contrast, the perception of vowel sounds considerably improved in genuine and non-word contexts in Carlet and Souza's study. However, despite receiving autonomous instruction, there was no increase in output, and it was likely because of the limitation of the study. Reversely, a study from Sariani et al claimed that incorporating mobile-based artificial intelligence, such as AI-powered pronunciation apps, can enhance learners' autonomy in pronunciation classes. However, some teachers may still be uncertain about maximizing the use of technology in pronunciation classes

2.2 Artificial Intelligence (AI) in Higher Education

AI is one of the many innovations brought about by the rapid advancement of technology. Popenici & Kerr (2017) define artificial intelligence (AI) as a computational system that can-do human-like functions like learning, adapting, synthesizing, self-correcting, and using data for intricate processing tasks. Its system is designed to function independently to varied degrees. AI refers to computers that carry out cognitive functions, including learning and problem-solving, that are typically performed by human minds (Baker et al., 2019). Keeping those concepts of AI in mind, there are numerous human functions that AI may assist with, especially in the realm of education. Additionally, the AI application can, in a sense, be seen as a workable option for the process of individualized education due to the automated aid in relation to the learning process (Ocaña-Fernández et al., 2019). According to Luckin et al. (2016), there are three categories of AI software applications already available in education: personal tutors, intelligent support for collaborative learning, and intelligent virtual reality. Luckin et al. (2016) also discuss the potential of AI for higher education. Each of them has a

unique function that has the potential to be employed in the process of teaching and learning. According to Yang, (2000) cited in Xie (2020), the adaptability of autonomous learning materials can eliminate the constraints of time and space, promote students' independent learning skills and attitudes, modify the various experiences, interests, abilities, and needs of individual learners, and then adjust the learning theme, speed, and path. This is especially true that students become independent in the learning process.

A software for learning English pronunciation called ELSA Speak: English Learning App (English Language Speech Assistant) uses artificial intelligence to recognize speech. It focuses on learning crucial English conversations and phrases rather than just the forty-four English sounds. It is appropriate for the needs of students enrolled in higher education who are getting ready to enhance their careers rather than working for a living. The reason to choose ELSA is because it focuses on practicing English through enjoyable language games that include essential English abilities including pronouncing words correctly, emphasizing keywords, maintaining rhythm and intonation, listening attentively, and conversing. A thorough report on the user's pronunciation strengths and shortcomings will be provided. By going to https://play.google.com/store/apps/details?id=us.nobarriers.elsa, anyone can access it. According to Becker & Edalatishams, (2019), who came to the same conclusion as the website, ELSA is an app that has made significant advancements in the field of AI for practicing the pronunciation of individual sounds. In line with that, Kholis (2021) stated that ELSA Speak utilized speech recognition technology to provide feedback and practice opportunities for students to improve their pronunciation skills. Yet, the research did not give further evaluation for the improvement of ELSA app itself.

2.3 Segmental and Suprasegmental Features

Understanding both segmental and suprasegmental features is important for effective communication and language acquisition. They play a significant role in language perception, production, and comprehension. The components are both present in phonology and cannot be distinguished. According to Senowarsito et al. (2019), segmental features are utterance characteristics that can be separated from an utterance, like consonants and vowels. However, individuals cannot correctly pronounce something using only segmental features; they also require loudness, stress, and a specific level of highness and lowness of pitch. Suprasegmental features are those that appeared along with segmental traits. Suprasegmental features, also known as prosodic features, go beyond individual sounds and extend over syllables, words, or phrases (Britannica, 2020; Zhang & Zhang, 2022). Furthermore, according to Morley (1999), the employment of suprasegmental pronunciation characteristics like stress, rhythm, and intonation goes beyond simply enhancing meaning to actually creating meaning. Because of this, certain suprasegmental elements can be emphasized to improve pronunciation. This is consistent with Derwing & Rossiter (2002) assertion that the transition from contrastive segmental teaching, which included little pair drilling, to a "more global approach" highlighted suprasegmental regions that had the greatest impact on comprehension. They note that this is changing as a result of research suggesting that native speakers' comprehension of non-native speech is closely related to their prosody.

3. Research Methodology

3.1 Research Design

Explanatory sequential mixed-method research using both quantitative and qualitative methods was used to create this research. A mixed-method research design is a process for gathering, analyzing, and "mix-ing" both quantitative and qualitative methodologies in a single study or a series of studies to comprehend a research problem, according to Meissner et al. (2011), quoted in Creswell (2012). The researchers used an explanatory sequential design that could be used to complete both quantitative and qualitative data in support of the experimental design. According to Fraenkel & Wallen (2012), the major goal of an explanatory sequential design is to collect quantitative data, which will then be followed by qualitative data to hone the conclusions from the quantitative data. In an explanatory sequential mixed method design, the researchers first gather quantitative data, which is followed by the collection of qualitative data to support the findings of the quantitative data. Further, W. J. and Creswell & Creswell (2018) described how the researchers' interpretation of this mixed technique lead them to interpret the findings in the study's discussion section, which took the form of quantitative findings for the first phase and qualitative findings for the second phase. Figure 2 below shows the explanatory sequential mixed methods design.

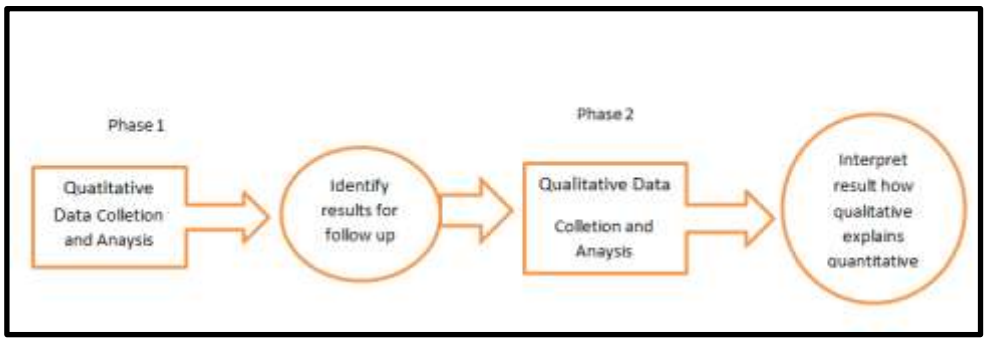

Figure 1: The explanatory sequential mixed methods design
(W. J. and Creswell & Creswell, 2018)

3.2 Participants

A population is a collection of people who share a particular trait (Creswell, 2012). Population refers to the larger group to which the findings are intended to be applied (Fraenkel & Wallen, 2012). In the academic year 2021–2022, 148 students were enrolled in the first semester of the English Education Study Program at Universitas PGRI Semarang. According to Creswell (2012), a sample is a subset of the target population that the researcher intends to analyze in order to draw general conclusions about the target population. According to Fraenkel & Wallen (2012), a research study sample is the population from which information is gathered. For this study, 32 students from Universitas PGRI Semarang's first-year students served as the sample, which was taken from one class. Purposive sampling was the method used to choose the sample. According to Creswell (2018), purposive sampling as a key element of qualitative research designs to achieve specific research objectives and generate valuable insights.

3.3 Instruments

A set of tools for gathering data is a crucial component of performing research. This study included assessments, interviews, and an evaluation of instructional materials as its

instruments. One technique to assess students' abilities, particularly in language study, was through the test. Pre-test and post-test were the two types of tests used in this study. The researchers presented forty-four English sounds in the pre-test. Five words are used to symbolize each sound. The students received treatment following the pre-test. The treatments involved selecting a pronunciation application from thousands of AI-based applications. It was completed between September 2021 and January 2022 in nine sessions. Every student is required to submit a report and supporting documentation for the method they used to independently learn how to pronounce the forty-four sounds using the program of their choice. The final test was a post-test in which the researchers used different words to present the same forty-four English sounds as the pre-test. To assess the student's pronunciation, the researchers performed an oral test.

Instrument validation is a crucial step in research, ensuring that the tools used to collect data are reliable and measurable. The current research used pilot testing to validate the instrument used. The researchers conducted a pilot study with a small sample of participants who were similar to the target sample. At last, the researchers used the result of the pilot study to make necessary revisions, especially to some of the forty-four English sounds.

Parametric and non-parametric tests are examples of tests. The design is seen in the parametric test for the big population. Additionally, a non-parametric test is made for smaller groups in which there is no clear population. The researchers employed the non-parametric test of pronunciation tests to get the data. The data were supported by interviews to validate the statistics and provide qualitative information. It concentrated on studying how the kids were able to learn pronunciation on their own, what application they used to do so, and what steps they took along the way.

3.4 Data Analysis Procedures

Since this study was using a mixed-method, it was dealing with two types of data which are quantitative and qualitative. Those data were collected using the instruments and after that, the data were analyzed differently. Through pre- and post-testing, quantitative analysis was used to evaluate the students' pronunciation scores using SPSS. However, a qualitative analysis was used to review the interview about the experiences of autonomous pronunciation learning. To bolster the findings of the tests, it was descriptively examined.

4. Findings

The primary goals of the current research were to investigate the significant effect of autonomous pronunciation learning using AI which was done quantitatively using SPSS, and to investigate the experiences of autonomous pronunciation learning using AI by higher level students which was qualitatively described. As the explanatory sequential mixed methods used in this research, the data display went initially with quantitative data then followed by the qualitative ones.

The first finding was quantitative data analyzed by SPSS which bore out with the hypotheses of whether the applications of AI based gave a significant effect to the students' pronunciation proficiency (Ho) or they did not give any significant effect to the students' pronunciation proficiency (Ha). The analyzed data were tabulated below.

Table 1: Paired Samples Statistics

		Men	N	Std. Deviation	Std. Error Mean
Pair 1	Pre-Test (X)	66.06	32	5.382	.951
	Post-Test(Y)	85.63	32	5.983	1.058

Table 2. Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	Pre-Test (X) & Post-Test(Y)	32	.557	.001

Table 3. Paired samples test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	Pre-Test (X) - Post-Test(Y)	-19.563	5.376	.950	-21.501	-17.624	-20.585	31	.000

As shown in Table 1, the paired samples statistics found that the mean of pre-test was 60.06 and post-test was 85.63. Table 2 is paired sample correlation which displays correlation between pre-test and post-test which obtained sig of 0.001 < from 0.05. It means there is a significant correlation between the tests. Meanwhile, Table 3 is paired t-test which provided the average difference from pre-test and post-test which is 19.563 and the t-column shows the result of t-count which is 20.585 with a P-Value of 0.000.

The second finding was qualitative findings to support the quantitative data. In-depth interview was conducted after the students fulfilled the tests to bolster the findings of the tests. Figure 2 depicts the interview results.

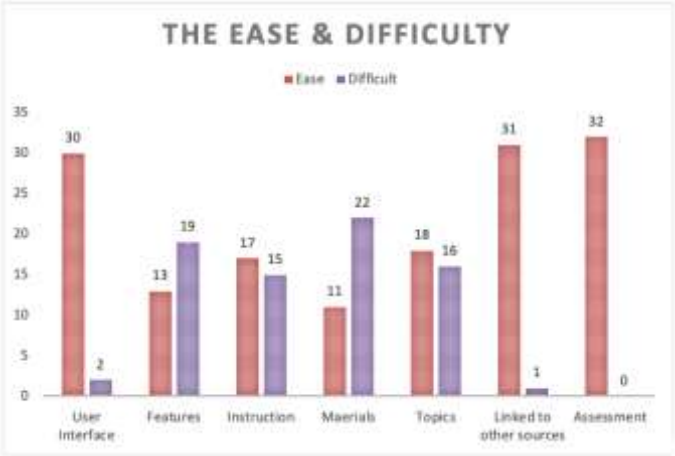

Figure 2: The experiences of autonomous pronunciation learning

Figure 2 displayed the students' experience in learning pronunciation using ELSA SPEAK application as the realization of autonomous learning. The interview included the user interface, features, instruction, materials, topics, linked to other sources, and assessment with the scale of easy and difficult. 30 students said the user interface was easy to use and the other 2 students said it was difficult. The features of the application were considered easy to 13 students, while 19 students said they were difficult. 17 students gave easy thought to the instruction whereas 15 students gave difficult thought. Materials got 11 perception of easy from the students, while 22 got difficult perception. Further, 18 students said the topics were easy while 16 students said they were difficult. Regarding to the ease and difficult linkage or connection from the application to another sources, 31 students said easy and 1 student said difficult. The last, the assessment of the application was considered easy for all of the students.

To support the more detail data, the screenshots of the autonomous pronunciation learning experiences are displayed in the Figure 3 and Figure 4.

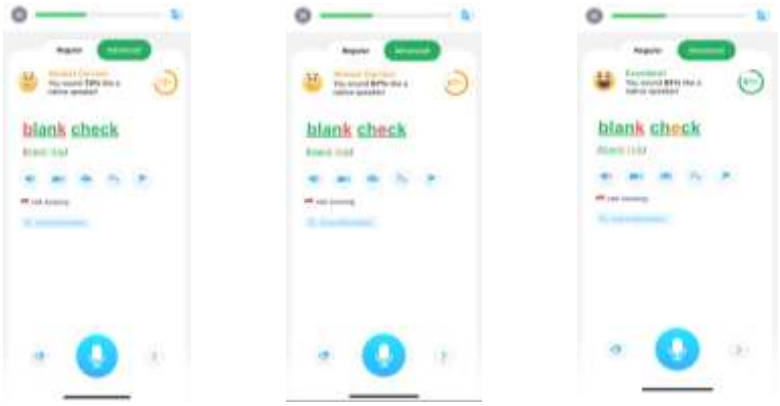

Figure 3: Sample of double consonant practice

Figure 3 was the sample of the process when the student learnt double consonant practice. It displayed three screenshots which means the students had three tries; two were almost correct and the final try was succeed. The first try got 73% correctness where the error was only in the first word. But when the student got the second try, the score was going lower (67%) where the error was on both words. From those three screenshots, it seems the student got difficulty in pronouncing double consonant /ŋk/ and /e/.

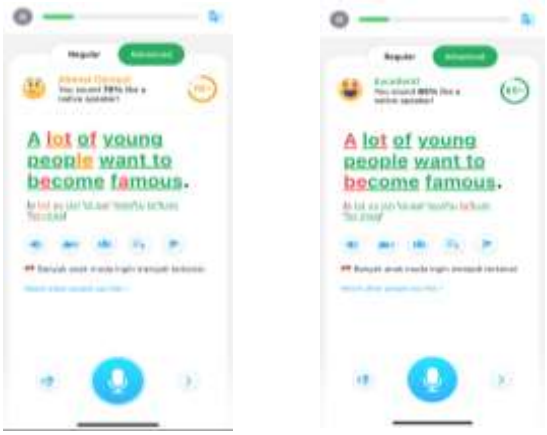

Figure 4. Sample of longer phrase practice

Meanwhile, Figure 4 was the sample of the process when the student learnt longer phrase practice. It displayed two screenshots which means the students had two tries; one scored 70% correctness and the other was 85%. From those two screenshots, it seems the student got difficulty in pronouncing consonants /l/, /t/, /b/, vowels /e/, /ə/, and diphthong /ei/.

5. Discussion

The aims of the study were investigating the effects of autonomous pronunciation learning using AI as well as the experiences of autonomous pronunciation learning using AI by higher level students. The SPPS was running to calculate the effectiveness and to proof the hypotheses whether the applications of AI based gave a significant effect to the students' pronunciation proficiency or it did not. After the treatments using AI based application named ELSA, it was figured out that the mean of post-test (85.63) was higher than pre-test (60.06). Meanwhile, the correlation obtained sig of 0.001 < from 0.05 which means there is a significant correlation between the use of AI and autonomous pronunciation learning. The average difference from pre-test and post-test was 19.563 and the t-column showed the result of t-count which was 20.585 with a P-Value of 0.000. To state the difference is significant, the f P-Value needs to be < 0.05. Therefore, the results claimed that the use of AI at higher education has a significant effect to the students' pronunciation proficiency.

The interview and screenshots findings were applied in attempt to explore deeper information regarding the experiences of autonomous pronunciation learning. Accordingly, the question leads to the ease and difficulty of using the application which classified into seven queries adopted from the important aspects of SLA related to autonomous pronunciation learning, including 1) segmentation and suprasegmentation are used, 2) lessons are aligned based on requirements analysis, and 3) demographic data such learner background and/or first languages are considered (Lee et al., 2015). The first query was about the user interface including the shape, color, and writing are designed as attractive as possible. 30 students said that ELSA application is colorful, interesting, eye-catching compared to other similar apps. The second dealt with features including topics and skills, levels, reports, and others. 19 students stated too many options led the users to confusion. In this phase, the students needed guidance from the lecturer. The next discovered about instruction and 17 students said that it is clear and understandable. Materials coverage including segmental and suprasegmental features became the next aspect which showed 22 students argued that the app explored mostly on segmental features, while suprasegmental features was only about intonation. The topics including food, entertainment, technology, culture, daily conversations were interesting topics based on 18 students since it was a day-to-day basis conversation. It was linked to other sources including YouTube was agreed by 31 students since they accessed videos such as Rachel Pronunciation and many others. The last is assessment including percentage and detailed report which is positively agreed by 32 students. The assessment was done within sounds.

However, the interview discovered the thought that 30 out of 32 students think that pronunciation is pretty hard and tricky. The more vary the English sounds, the more struggle they experienced to pronounce words correctly. As displayed in Figure 2, on the first try he got 73% correct for he was struggle with sound /b/ and /k/ in word 'blank' but then became lower to 67% on the second try since he added the incorrectness to the second word that was 'check'. Then, on the third try, he made it by getting 83% incorrectness even though the

application showed the same incorrectness. In the learning process, the students also struggled with several things. The first was phonetic symbols which is the most troublesome one. As most of the EFL students, learning them for the first time led so much confusion and needed extraordinary memory dealing with those symbols along with the sounds. However, by using ELSA application, students could access deeper comprehension and practices about diphthongs, the flap sound /r/, dropping consonants, how to make differences within similar sounds such as /s/, /ʃ/, /z/, consonants clusters, 'TH' sounds, joining sounds, and so on. As Figure 1 displayed that the app explored mostly on segmental features than suprasegmental features, another issue was occurred; that is the imbalance between work on the segmental and suprasegmental levels to improve comprehensibility and fluency. This statement is also in line with the studies from Anderson-Hsieh et al. (1992); Derwing & Munro (1997); Derwing & Rossiter (2002); Kang et al. (2010); Reinders & Pegrum (2015) which claimed that suprasegmental features are mostly affected to have accent reduction than segmental ones.

Problematic exercise choices for the segmental and suprasegmental nervous systems are lacking (Becker & Edalatishams, 2019). The body of research on pronunciation instruction demonstrates unequivocally that instruction that emphasizes suprasegmental qualities enhances intelligibility and fluency more than education that emphasizes segmental aspects (e.g., Derwing & Rossiter (2002); Kang et al. (2010). In reality, according to Kang et al. (2010), as long as the suprasegmental features are fairly accurate, "audience can tolerate large inaccuracies in the pronunciation of consonants and vowels" (p. 555). The importance of suprasegmental traits above segmental features has been demonstrated in other research as well (Anderson-Hsieh et al., 1992; Derwing & Munro, 1997). Reinders & Pegrum (2015) approach for evaluating MALL applications has two criteria that link app design to teaching strategies, although there are many more. ELSA pays little attention to suprasegmental features that may be most beneficial for accent reduction.

Other study, Baker et al. (2019), also stated that accuracy of AI can be missed. It is explained that human decision-making can be opaque and biased in its own way, and is often an alternative to AI decision-making. However, Indonesian students are okay with the accuracy of accent since they are not going to focus solely on it. This is supported by Ur (1984) that EFL educators should persuade their students that studying pronunciation is not about acquiring a native accent, but rather about teaching them how to talk clearly and comfortably with other speakers. Indonesian students are more concerned by the score percentage they get every lesson. Furthermore, this method proves to enhance higher students' pronunciation as long as there is an obligation to report the progress every week. It will keep improve their skill even though every student has different usage frequency.

They can practice pronunciation by listening and imitating the provided audio. Most of the applications also provide phonetic symbols with the audio in each symbol. It can be the tool to understand more about English pronunciation sounds. The students could recognize the sounds and the symbols because of the provided audio. For example, vowel sounds and double vowels. Those two sounds are the sounds that most of the students couldn't distinguish. But after learning with EPA (English Pronunciation Application), a lot of students testified that they could distinguish them now. There are more sounds that the students could understand and distinguish better after learning with English Pronunciation Application. The students also could understand better where to put emphasis in pronouncing words since it is also provided in application. All of the features provided in

application made the students can practice pronunciation in English like native speakers. The students' learning process is also aided by the clear application of lessons. The application was not difficult for the students to use. As the application can be accessed offline, the students can spend their leisure time to learn from it. In that way, they can study autonomously everywhere and anywhere.

Another thing that made students struggled to learn pronunciation was the word emphasis. English words have different place to emphasis in each word which obviously different from Indonesian. Hence, the students often got confused where to put the emphasis. ELSA gives them chance to imitate directly from the natives to make them get used to English intonation. However, emphasis belongs to one of elements in suprasegmental features. There are more such as stress, length, rhythm, tone, etc. ELSA provides one feature which is actually including all of those features but it is not specifically tutored and explored in the app. It is called intonation. Thus, students need guidance in this phase.

From the result, the researchers also found they love using ELSA because even though the premium version is more complete than free, but the free version is enough for beginners. This application also can detect the pronunciation error made by the students. After analyzing the error, it also gives feedback so students can correct their pronunciation to practice more.

Within seconds of speaking a phrase, you'll be able to: (1) hear how you sound compared to a native speaker; (2) get pointers on how to improve; (3) receive feedback on other sounds via IPA through the Advanced Feedback feature; and (4) watch a video that demonstrates how to produce those difficult sounds. This app uses "proprietary artificial intelligence" (AI), including automated voice recognition (ASR), as stated on their website, to provide feedback on users' pronunciation accuracy.

The result of the current study support what has been conveyed by Luckin et al. (2016) where AI has so much potential for higher education, mostly the Industrial Revolution era 4.0. This era has tremendous impacts on human life, one of them is in education sector. Higher education is critical to preparing future generations to be ready for the AI-filled work environment that awaits them. On the other hand, Hidayati & Husna (2020) found that learning through applications more enjoyable. By using application, the students showed a considerable potential for being independent learners. Since students only considered some aspects of online learning enjoyable while the other portions were not, the study did not evaluate online learning as effective or ineffective. Therefore, this current study is one step ahead that learning through application is not only enjoyable and giving contribution to autonomous learning but also appointing to particular course and providing many sources of the application which become the novelty of this study.

6. Conclusion

It is not simple to teach pronunciation in higher education in Indonesia. This study reports the effects of autonomous pronunciation learning using AI as well as the experiences of autonomous pronunciation learning using AI by higher level students. The results of both quantitative and qualitative which had been discussed on the preceded sub-section convinced the researchers to encourage the educators, especially English teachers and lecturers, to use English Pronunciation Application in enhancing students' pronunciation

proficiency in every course they have. The results of this study also became solution of the thought belongs to the students that learning pronunciation is hard, moreover, when most of them are not familiar using the application with AI based. In attempt to support the Fourth Revolution Industry where the development of science and technology are rapidly increased, the application of AI in every course that educators have is very recommended.

The features provided in the application allow the students to train the students pronouncing English words correctly and engagingly. They provide easy access since it already has clear instructions. This application also can support learning pronunciation autonomously because it can be accessed anytime and anywhere. The lecturers are responsible as facilitator. Therefore, this activity is advisable to be applied in the hybrid learning.

However, ELSA, the AI employed in this study, has certain shortcomings. When we look at its features, ELSA appears to be primarily concerned with segmental features rather than suprasegmental ones. Only intonation is available from among all the features offered to practice suprasegmental features. While students were finding it difficult to emphasize words, there is no other practice for suprasegmental qualities. In reality, the ELSA website states that its curriculum covers core English skills such word stress, intonation, rhythm, listening, and conversation. This criticism might be taken into consideration by ELSA's creator as they attempt to improve the course.

References

Anderson-Hsieh, J., Johnson, R., & Koehler, K. (1992). The Relationship Between Native Speaker Judgments of Nonnative Pronunciation and Deviance in Segmentais, Prosody, and Syllable Structure. *Language Learning*, *42*(4), 529–555. https://doi.org/10.1111/J.1467-1770.1992.TB01043.X

Ardini, S. N., WL, M. Y., & Ouwpoly, N. L. (2016). Error Analysis of Phonetic Fossilization Uttered by English Department Students University of PGRI Semarang. *Lensa: Kajian Kebahasaan, Kesusastraan, dan Budaya*, *6*(1), 1-8. https://jurnal.unimus.ac.id/index.php/lensa/article/view/1917

Baker, T., Smith with Nandra Anissa, L., Sheehan, K., Ward, K., Waters, A., Berditchevskaia, A., Van Den Berg, C., Campbell, N., Candsell, O., Casasbuenas, J., Cinnamon, J., Copeland, E., Duffy, E., Hannon, C., John, J., Grant, J., Klinger, J., Latham, M., Macken, C., … Ward-Dyer, G. (2019). *Educ-AI-tion Rebooted? Exploring the future of artificial intelligence in schools and colleges*. www.nesta.org.uk

Becker, K., & Edalatishams, I. (2019). ELSA Speak Accent Reduction [Review]. *Pronunciation in Second Language Learning and Teaching Proceedings*, *10*(1), 434. https://www.iastatedigitalpress.com/psllt/article/id/15397/

Britannica, T. Editors of Encyclopedia (2020, February 28). *suprasegmental*. *Encyclopedia Britannica*. https://www.britannica.com/topic/suprasegmental

Carlet, A., & De Souza, H. K. (2018). Improving L2 Pronunciation Inside and Outside the Classroom: Perception, Production and Autonomous Learning of L2 Vowels. In *Ilha do Desterro* (Vol. 71, Issue 3, pp. 99–123). Universidade Federal de Santa Catarina. https://doi.org/10.5007/2175-8026.2018v71n3p99

Chassignol, M., Khoroshavin, A., Klimova, A., & Bilyatdinova, A. (2018). Artificial Intelligence Trends in Education: A Narrative Overview. *Procedia Computer Science*, *136*, 16–24. https://doi.org/10.1016/J.PROCS.2018.08.233

Creswell, J. (2012). *Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research*. http://nuir.nkumbauniversity.ac.ug/handle/20.500.12383/985

Creswell, W. J. and, & Creswell, D. J. (2018). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches, 5th Edition: by John W. Creswell and J. David Creswell, Los Angeles, CA: SAGE, 2018, $38.34, 304pp., ISBN: 978-1506386706: Journal of Electronic Resources in Medical Libraries: Vol 19, No 1-2*. https://www.tandfonline.com/doi/full/10.1080/15424065.2022.2046231

Derwing, T. M., & Munro, M. J. (1997). Accent, Intelligibility, and Comprehensibility: Evidence from Four L1s. *Studies in Second Language Acquisition*, *19*(1), 1–16. https://doi.org/10.1017/S0272263197001010

Derwing, T. M., & Rossiter, M. J. (2002). ESL Learners' Perceptions of their Pronunciation Needs and Strategies. *System*, *30*(2), 155–166. https://doi.org/10.1016/S0346-251X(02)00012-X

Fraenkel, J. R., & Wallen, N. E. (2012). *How to Design and Evaluate Research in Education*.

Gashimov, E. (2023). Interference in a Language and Culture Communication. *Allure Journal*, *3*(1), 56–63. https://doi.org/10.26877/allure.v3i1.14156

Hidayati, T., & Husna, F. (2020). *Learning English from Home during the Covid-19: Investigating Learners' Experience for Online and Autonomous Learning*. *6*(2). http://dx.doi.org/10.313

Holec, Henri., & Council of Europe. (1979). *Autonomy and Foreign Language Learning.* 53.

Huang, G., & Moore, R. K. (2023). Using Social Robots for Language Learning: Are We There Yet? *Journal of China Computer-Assisted Language Learning*, *3*(1), 208–230. https://doi.org/10.1515/jccall-2023-0013

Hussein Banafa, F. (2008). *Effects of IT on Pronunciation Impact of the Internet and Interactive Multimedia on English Pronunciation of Arab College Students Studying in the United States*. VDM Verlag.

Jiang R (2022) How does Artificial Intelligence Empower EFL Teaching and Learning Nowadays? A review on Artificial Intelligence in the EFL Context. Front. Psychol. 13:1049401. Doi: 10.3389/fpsyg.2022.1049401

Joan Morley. (1999). New Developments in Speech Pronunciation Instruction. As We Speak. *Newsletter of the TESOL Speech, Pronunciation, and Listening Interest Section*, 1–5.

Kang, O., Rubin, D., & Pickering, L. (2010). Suprasegmental Measures of Accentedness and Judgments of Language Learner Proficiency in Oral English. *The Modern Language Journal*, *94*(4), 554–566. https://doi.org/10.1111/J.1540-4781.2010.01091.X

Kholis, A. (2021). Elsa Speak App: Automatic Speech Recognition (ASR) for Supplementing English Pronunciation Skills. *Pedagogy: Journal of English Language Teaching, 9*(1), 01-14. doi:10.32332/joelt.v9i1.2723

Kruk, M. (2012). Using Online Resources in the Development of Learner Autonomy and English Pronunciation: The Case of Individual Learners. *Journal of Second Language Teaching & Research*, *1*(2), 113–142. https://pops.uclan.ac.uk/index.php/jsltr/article/view/28

Layali, K., and Al-Shlowiy, A. (2020). Students Perceptions of E-learning for ESL/EFL in Saudi Universities at Time of Coronavirus: A Literature Review. *Indonesian EFL Journal*. 6, 97–108. doi: 10.25134/ieflj.v6i2.3378

Lee, J., Jang, J., & Plonsky, L. (2015). The Effectiveness of Second Language Pronunciation Instruction: A Meta-Analysis. *Applied Linguistics*, *36*(3), 345–366. https://doi.org/10.1093/APPLIN/AMU040

Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. (2016). *Intelligence Unleashed: An Argument for AI in Education*. http://oro.open.ac.uk/50104/1/Luckin%20et%20al.%20-%202016%20%20Intelligence%20Unleashed

Meissner, H. I., Creswell, J. W., Klassen, A. C., Clark, V. L. P., & Smith, K. C. (2011). *Best Practices for Mixed Methods Research in the Health Sciences*.

Ocaña-Fernández, Y., Valenzuela-Fernández, L. A., & Garro-Aburto, L. L. (2019). Inteligencia Artificial y Sus Implicaciones en la Educación Superior. *Propósitos y Representaciones*, *7*(2). https://doi.org/10.20511/pyr2019.v7n2.274

Penny Ur. (1984). *A Course in Language Teaching: Practice and Theory.* Cambridge University Press.

Phuong, T. T. H. (2021). Who Should Teach English Pronunciation? Voices of Vietnamese EFL Learners and Teachers. *Journal of Asia TEFL*, *18*(1), 125–141. https://www.dbpia.co.kr/journal/articleDetail?nodeId=NODE10649594

Popenici, S. A. D., & Kerr, S. (2017). Exploring the Impact of Artificial Intelligence on Teaching and Learning in Higher Education. *Research and Practice in Technology Enhanced Learning*, *12*(1). https://doi.org/10.1186/s41039-017-0062-8

Reinders, H., & Pegrum, M. (2015). *Supporting Language Learning on the Move. An evaluative framework for mobile language learning resources.* https://www.researchbank.ac.nz/handle/10652/2991

Rifqiyah, A., Ardini, S. N., & Kusumo, A. B. P. (2021). English pronunciation application as A Media to Improve Students' Pronunciation: The Effectiveness. *Linguistics and Education Journal*, *1*(1), 74–85. https://doi.org/10.17576/3L-2019-2502-06

Rusmiyanto, R., Huriati, N., Fitriani, N., Tyas, N., Rofi'i, A., & Sari, M. (2023). The Role of Artificial Intelligence (AI) In Developing English Language Learner's Communication Skills. *Journal on Education*, *6*(1), 750- 757. https://doi.org/10.31004/joe.v6i1.2990

Sariani, Miladiyenti, F., Rozi, F., Haslina, W., & Marzuki, D. (2022). Incorporating Mobile-based Artificial Intelligence to English Pronunciation Learning in Tertiary-level Students: Developing Autonomous Learning. *International Journal of Advanced Science Computing and Engineering*,4(3), December 2022, pp. 220-232.

Senowarsito, & Ardini, S. N. (2016). *English Phonology (For Learners of English as a Foreign Language). Universitas PGRI Semarang*.

Senowarsito, & Ardini, S. N. (2019). Phonological Fossilization of EFL Learners: The Interference of Phonological and Orthographic System of L1 Javanese. *The Southeast Asian Journal of English Language Studies*, *3*(2), 74–85. https://doi.org/10.17576/3L-2019-2502-06

Tahereen, T. (2015). Challenges in Teaching Pronunciation at Tertiary Level in Bangladesh. *International Journal of English Language & Translation Studies*, *3*(1), 9–20. http://www.eltsjournal.orghttp://www.eltsjournal.org

Tiara, A. F., Ardini, S. N., & Nugrahani, D. (2020). Blended Learning in Pronunciation Classroom for Higher Education: Students' Perception. *E-Structural (English Studies on Translation, Culture, Literature, and Linguistics)*, 3(02), 145-156. DOI: https://doi.org/10.33633/es.v3i02.4414

Tuomi, I. (2018). *The Impact of Artificial Intelligence on Learning, Teaching, and Education. Policies for the future, Eds*. Cabrera, M., Vuorikari, R & Punie, Y., EUR 29442 EN, Publications Office of the European Union, Luxembourg, ISBN 978-92-79-97257-7, doi:10.2760/12297, JRC113226.

Xie, Z. (2020). Effectiveness of Autonomous Learning Materials for Students during the COVID-19 Pandemic: A Case Study of the Daxie Second Elementary School in Ningbo, Zhejiang, China. *Science Insights Education Frontiers*, *6*(1), 613–624. https://doi.org/10.15354/sief.20.or023

Yang. (2000). *Self-study textbook design manual* (pp. 3–5). Psychological Publishing House.

Zhang, F., & Yin, P. (2009). *A Study of Pronunciation Problems of English Learners in China*.

Zhang, H., & Zhang, L. (2022). Introducing Chinese Linguistics: A Handbook for Chinese Language Teachers and Learners. Studies in Chinese Language and Discourse. John Benjamins Publishing Company. https://benjamins.com/content/home