METATHESIS
Journal of English Language, Literature, and Teaching
Volume 8, No 1, April 2024
DOI: https://doi.org/10.31002/metathesis.v8i1

Publisher:
English Education Department
Faculty of Education and Teacher Training
Universitas Tidar, Magelang - Indonesia

e-ISSN: 2580-2720 p-ISSN: 2580-2712

O'Speak Version 1.0: A New Tool to Measure Segmental Pronunciation Features

Widya Ratna Kusumaningrum[1], Rolisda Yosintha[2], Boris Ramadhika[3*]

[1,2,3]Universitas Tidar, Magelang, Indonesia

[1]kusumaningrum@untidar.ac.id
[2]rolisda@untidar.ac.id
[3]ggramadhika@untidar.ac.id

*corresponding author

History

Received
7 March 2023

Revised
3 July 2024

Accepted
3 July 2024

Published
9 July 2024

Abstract

The rapid enhancement of technology has made it possible to integrate technology and L2 pronunciation assessment. While the investigation of L2 pronunciation was considered vital in English Language Teaching, assessing pronunciation is granted the least attention. This study attempts to discuss the roles and impacts of O'Speak version 1.0 as an automated pronunciation tool and compare it with human ratings while assessing L2 segmental pronunciation features uttered by Indonesian learners of English. This study aims to pilot an android-based pronunciation test, namely, O'Speak, which was developed using Feuerstein's Mediated Learning Experience principles. Performed under a quasi-experimental research design, this study ran an independent two-sample t-test involving 50 participants. The study showed that there was no statistically significant difference between O'Speak and human ratings in the segmental pronunciation assessment. This indicates that a new tool functions equally with the ability that human rating has. During the study, this study identified some caveats shown by the human rating that leads to its ability to be equal to O'Speak, and these include teaching experience, hallo effect, and rating experience.

Keywords: *automated pronunciation evaluation, segmental features, human rating, o'speak*

DOI: https://doi.org/10.31002/metathesis.v8i1.494

Access article distributed under the terms of the Creative Commons Attribution license (https://creativecommons.org/licenses/by-sa/4.0/)

INTRODUCTION

As language learning has undergone rapid technological changes over time, technologies have been sought by many teachers to improve the educational process (Dermentzi et al., 2016). For example, teachers use technologies to promote students' learning (Manca & Ranieri, 2017). One of the enhanced technologies is Computer Assisted Language Learning (CALL) to Mobile Assisted Language Learning (MALL). Both terms are used interchangeably.

MALL generally refers to the use of mobile devices without being limited to a physical time or a determined time (Kukulska-Hulme & Viberg, 2018). Mobile learning is characterized by its informal, personalized, and spontaneous usage. Mobile devices are not used for substituting existing learning devices. They are used as an extension for learning in a new environment (Miangah, 2012) which can affect their attitudes, motivation, and awareness (Ramadhika et al., 2023). The use of mobile devices is beginning to impact how language learners can engage in activities without depending on access to computers (Kukulska-Hulme, 2009).

In line with this trend, mobile application developers compete to provide an application that can be used as alternative language learning tools for those who want to learn a language, one of which is O'Speak. O'Speak, an android-based application, is a mobile-based test assessing pronunciation, which is developed using Feuerstein's Mediated Learning Experience framework (Kusumaningrum et al., 2020). This application helps students to master the correct way of pronouncing English words. In other words, pronunciation must be taught and learned in accordance to how the words are pronounced correctly.

However, teaching pronunciation is often overlooked particularly by those who start learning English as a second language lately (after school years) are likely to have difficulties in pronouncing intelligible pronunciation (Gilakjani et al., 2011). Pronunciation is seen as a sub-skill of speaking in which the distinction between segmental and suprasegmental features is viewed correctly (Hyman, 2019). Furthermore, there is no argument that English's hegemony means of communication is dynamic in nature; it creates a family of "Englishes," which cannot be claimed by one's own due to the different varieties (Guinto, 2013). As a result, There has to be some kind of complexity related to the construction of language proficiency, in particular to accuracy and fluency (Kuiken, 2023).

Second language pronunciation is a cognitive skill. The main problem for ESL students is that they need to change their conceptual patterns in their first language that they have internalized in childhood (Gilakjani et al., 2011). In that case, this paper attempts to find out how significant is the first prototype of O'Speak (version 1.0) as an automated pronunciation evaluation tool when assessing the segmental features. This study specifically aims to see a difference between teachers who have been teaching pronunciation and have understood how segmental features are measured using the O'Speak Version 1.0 application.

Literature Review

Speaking and Pronunciation

Speaking is said to be the most essential skill to be developed in learning a second or foreign language as a means of effective communication, and yet it is not easy to master. Many language learners face some considerable challenges when dealing with speaking. These constraints might be related to some aspects, namely accuracy and fluency. Accuracy refers to the use of exact and complete form of language when speaking. Gower, Philips, and Walter as cited in (Derakhshan et al., 2016) state that accuracy covers the proper use of vocabulary, grammar, and pronunciation. To facilitate accuracy, learners should also be fluent in speaking in such a way that the listeners can easily understand and thus communication breakdown can be avoided (Hughes, 2011). Hedge as cited in (Drajati, 2018) contends that fluency in speaking is correlated to 1) the acceptable answer

Access article distributed under the terms of the Creative Commons Attribution license (https://creativecommons.org/licenses/by-sa/4.0/)

in the changes of communication; 2) proper use of linking devices; and 3) correct intonation and clear pronunciation.

Given that, pronunciation plays an important role both in aiding accuracy and fluency during spoken communication. Even though learners make some inaccuracies in grammar or vocabulary, they could still communicate effectively when they produce a language in proper pronunciation and intonation. In fact, Hinofotis and Baily (as cited in Gilakjani et al., 2011) find out that during communication, the most severe impair that EFL/ESL learners makes is related to pronunciation, not grammar or vocabulary. Pronunciation is defined as a method of producing certain sounds (Richard and Schmidt as cited in Gilakjani, 2016). Similarly, Yates et al., (2015) describes pronunciation as sound production process that is used to create meaning. Therefore, it can be said that pronunciation concerns how a language is spoken to make meaning.

To ensure that meaning is correctly decoded by listeners, a speaker must pay really close attention to the segmental feature of pronunciation. It deals with phonemes, the smallest unit of a language that differentiate one word from another. Phonemes include vowels and consonants. Vowels are the sounds produced when the breath flows out through the mouth without being obstructed by the teeth, tongue, or lips, while consonants are produced by blocking the air from flowing through the mouth by closing the lips or touching the teeth with the tongue. Errors in this feature might impede communication by slowing down the speed of word recognition (Smith, 2004).

Despite the importance of segmental features of pronunciation, this aspect is often given the least care (Gilakjani et al., 2011; Marzá, 2014; Thornbury, 2007). Most teachers often focus on the suprasegmental features by emphasizing linking, intonation, and stress. Limited time is one of the most common reasons why teachers often neglect teaching this feature in the instructional process. In addition, pronunciation is also often viewed to be linked to intonation and stress only, and thus sacrificing accuracy in phoneme level. Learners also find it difficult to produce the appropriate segmental features in pronunciation to resemble that of native speakers. Mother tongue interference, sound system differences between L1 and L2, inconsistency of English vowels, and influence of spelling on pronunciation are identified as some of the factors affecting the constraints on this matter (Hassan, 2014).

To help improve the condition, teachers should employ materials or tools that can help learners improve their accuracy in the segmental aspect of pronunciation. One of the alternatives that teachers can do is to make use of technology to teach and assess L2 pronunciation. By integrating technology in pronunciation teaching and learning process, teachers can have more options to deliver the materials and learners will have more opportunities to develop their pronunciation skill independently. Many researchers have integrated technology into pronunciation instruction by developing assistive tools for evaluating pronunciation and creating speech nativeness (Ronanki et al., 2012; Tu et al., 2018; Zhang et al., 2020). This automated pronunciation evaluation system has seen tremendous success in both research and commercial settings, especially in L2 learning context (Black et al., 2015).

O'Speak

O'Speak is one example of automated pronunciation evaluation applications. This application is developed by Kusumaningrum et al., (2020) using an android-based operation system. It is designed to facilitate teachers' and students' needs for a simple yet accurate pronunciation assessment tool. Using this application, learners could identify their pronunciation errors, learn vocabulary and sounds, and ultimately enhance their pronunciation performance. O'Speak is developed based on the framework of Feuerstein's Mediated Learning Experience (MLE) (Kusumaningrum et al., 2020). MLE envisions an effective learning process and fosters active interaction between learners and environment through the orderly, practical, and structural exposure as given by a mediator (van Leeuwen & Janssen, 2019).

Access article distributed under the terms of the Creative Commons Attribution license (https://creativecommons.org/licenses/by-sa/4.0/)

O'Speak version 1.0 includes three different levels of difficulty: (i) beginner: the word level of English pronunciation, (ii) intermediate: the sentence level of English pronunciation, and (iii) advanced: the paragraph level of English pronunciation (see Figure 1). The application has two main features (i) my class that allows the interaction with the classroom teacher (ii) practice that could be done independently due to the availability of four different topics: clothes, health, weather, and transportation.

Figure 1. O'Speak version 1.0 in three different levels of difficulty

METHOD

The study was undertaken using a quasi-experimental research design to investigate the significant difference of O'Speak version 1.0 as compared to the human rating in segmental features assessment. The study was carried out over a period of six weeks (November-December 2019).

Participants

The study involved two groups of EFL students at a tertiary institution who were taking the English Phonetics and Phonology Course (N=50; Mage=19.47 years; range=19-20 years): (1) the O'Speak Group (OG) (N=26) and Human Rating Group (HRG) (N=24). This study employed purposive sampling from four classes (N=112) learning English phonetics and phonology course theories. Prior to identifying the intact/experimental groups and control groups, all of the students who were taking the English Phonetics and Phonology course were evaluated based on their average scores on the prior English Voice and Accent course (MGPA_Class1=3.17; MGPA_Class2=3.19) and their pretest score (Mpretest_Class1=68.27; MGPA_Class2=68.50). These two classes were then as the intact/experimental group (henceforth, the O'Speak Group shorten as OG) and as the control group (henceforth, the Human Rating Group, HRG).

Instruments

O'Speak version 1.0

This study used O'Speak version 1.0 to investigate how segmental features are measured. O'Speak is an android-based test that is applicable for evaluating English pronunciation, and it integrates the Feuerstein's Mediated Learning Experience principles (Kusumaningrum et al., 2020). O'Speak version 1.0 includes three different levels of difficulty: (i) beginner for the word level of English pronunciation, (ii) intermediate for the sentence level of English pronunciation, and (iii) advanced for the paragraph level of English pronunciation (see Figure 1). The application has two main menus my class that permitted the interaction with the classroom teachers and practice that were augmented with practice on four different topics such as clothes, health, weather, and transportation. In this study, we focused on isolating the word level of English pronunciation within these four topics.

Access article distributed under the terms of the Creative Commons Attribution license (https://creativecommons.org/licenses/by-sa/4.0/)

Recording

The recording was used to help the EFL teacher document and store the students' pronunciation in the control group (the Human Rating group) prior to the pronunciation assessment.

The pronunciation ratings

The study adopted from training materials and onscreen labels for pronunciation and lexicogrammar judgment (Saito et al., 2015). The onscreen labels do not have any numerical labels or marked intervals but rather a 1000-point sliding scale (see Figure 2).

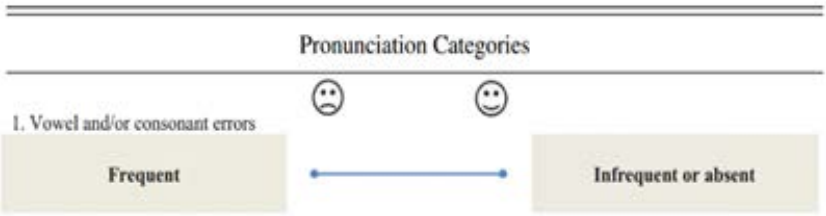

Figure 2. Training materials and onscreen labels for pronunciation and lexicogrammar judgment (Saito et al., 2015)

Data Analysis Procedures

The study employed a reading-aloud strategy for both intact/experimental groups and controlled groups. Both groups were given the same tasks for the pretest and posttest where the participants read the word level of English pronunciation on four different topics. For the O'Speak Group, the O'Speak version 1.0 prototype was shared and installed on their mobile phones. The participants should register their information in the application before it could be used. Meanwhile, for the Human Rating Group, the participants articulated the same words referring to the tasks on O'Speak version 1.0, (see Table 1) and their pronunciation was recorded and rated by an EFL teacher using training materials and onscreen labels for pronunciation and lexicogrammar judgment (Saito et al., 2015). Both groups were given limited time to plan and practice (less than 5 seconds). The results of the pretest and posttest were then evaluated to identify some possible differences between the two. The study used the independent two-sample t-test to see the significant difference between two independent groups between the O'Speak Group and the Human Rating Groups when assessing segmental pronunciation features.

Table 1. Segmental features in isolated words

Topic	Task	Segmental focus
clothes	dress	consonant /s/
	polo shirt	consonant /ʃ/
	handbag	consonant /g/ and vowel /æ/
	sneakers	consonant /s/
	tops	consonant /t/
	singlet	vowel /ɪ/
	jumper	consonant /r/
	coat	diphthong /əʊ/
	vest	consonant /v/
	blazer	consonant /z/

Access article distributed under the terms of the Creative Commons Attribution license (https://creativecommons.org/licenses/by-sa/4.0/)

health	rash	consonant /ʃ/
	sunburn	vowel / ɜː/
	headache	diphthong /eɪ/ and consonant /k/
	diarrhea	diphthong /aɪ/, vowel /ə/ vowel /ə/
	cough	vowel /ɒ/, consonant /f/
	cold	diphthong /əʊ/
	fever	consonant /v/
	broken leg	diphthong /əʊ/ and consonant /g/
	asthma	vowel /æ/ and vowel /ə/
	measles	vowel /iː/ consonant /z/
weather	snowflakes	diphthong /əʊ/
	mild	diphthong /aɪ/ and consonant /d/
	dull	vowel /ʌ/
	damp	vowel /æ/ and consonant /p/
	frost	vowel /ɒ/ consonant /t/
	breeze	consonant /z/
	miserable	vowel /ɪ/ and consonant /z/
	puddles	vowel /ʌ/
	refreshing	consonant /ʃ/
	scorching	consonant /tʃ/
transportation	lorry	vowel / ɒ/
	moped	diphthong / əʊ/ and vowel /e/
	fireboat	triphthong /aɪə/ and diphthong /əʊ/
	submarine	vowel /ʌ/
	bomber	silent consonant /b/
	sled	consonant /d/
	tugboat	vowel /ʌ/
	tube train	consonant /tʃ/
	yacht	consonant /j/ vowel /ɒ/
	handcart	vowel /æ/

FINDINGS AND DISCUSSIONS

Descriptive Statistics

The descriptive statistics allows us to see the normal distribution of the data by examining the average pretest and posttest scores before the parametric inferential statistics were done.

Table 2. Descriptive statistics for the participants' pretest and posttest

Group	Test	N	Lower	Upper	Mean	Std. Deviation	Skewness	Kurtosis
OG	Pretest	26	63	75	68.27	2.85	0.02	0.12
	Posttest	26	70	80	74.35	2.50	0.64	0.46
HRG	Pretest	24	65	75	68.50	2.70	0.27	0.90
	Posttest	24	66	75	69.25	2.21	0.27	0.96

Access article distributed under the terms of the Creative Commons Attribution license (https://creativecommons.org/licenses/by-sa/4.0/)

The descriptive statistics in Table 2 imply that both groups had equal knowledge and understanding before the study was done, with (Mpretest_Class1=68.27; MGPA_Class2=68.50). The investigation on the normal score distribution was seen from their skewness range to understand the distribution of the score and their kurtosis range to see the score's density. As depicted in Table 2, each group's skewness range was between 0.02-0.64, and the kurtosis range was between 0.12-296. These data were interpreted that the data displayed positive skewness with a narrow margin of the normal condition with the mesokurtic shape. In short, these raw data were normally distributed (see Figure 3).

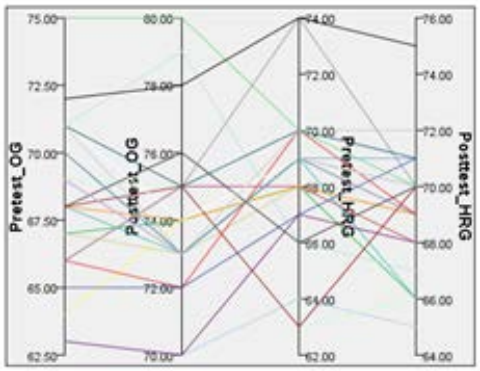

Figure 3. The distribution and density of the pretests posttests scores

O'Speak Consistency

Before carrying out the parametric inferential statistics, the study measured the reliability to see the degree of consistency of the O'Speak in assessing the segmental pronunciation features (consonants and vowels). It was evident that O'Speak had a high-reliability coefficient of Cronbach's Alpha (α=0.779) (see Table 3) with high inter-item correlation (see Table 4).

Table 3. Reliability Statistics for O'Speak Consistency

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.779	.783	2

Table 4. Inter-Item Correlation Matrix for O'Speak Consistency

	Pretest_OG	Posttest_OG
Pretest_OG	1.000	.644
Posttest_OG	.644	1.000

In Table 4, the high correlation between the pretest and posttest scores are showed during the use of O'Speak to measure the segmental pronunciation features (α=0.644). It appears that O'Speak was able to be used to reveal the segmental quality.

Human Rating Consistency

The study involved an EFL teacher as the human ratings, and to see the internal scoring consistency, the intra-rater reliability coefficients (see Table 5) were calculated between two different times, pretest and posttest (see Table 6).

Access article distributed under the terms of the Creative Commons Attribution license (https://creativecommons.org/licenses/by-sa/4.0/)

Table 5. Reliability Statistics for Human Rating Consistency

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.736	.745	2

In table 5, it is identified that the EFL teacher who measured and gave personal judgment to the students with the adapted pronunciation rating scale had a high-reliability coefficient of Cronbach's Alpha (α=0.736). This coefficient was verified with the internal consistency between pretest and posttest scores (see Table 6), which means the rater was reliable enough to evaluate students' segmental quality.

Table 6. Inter-Item Correlation Matrix for Human Rating Consistency

	Pretest_HRG	Posttest_HRG
Pretest_HRG	1.000	.593
Posttest_HRG	.593	1.000

Inferential Statistics

This study ran the independent two-sample t-test to understand the significant difference between O'Speak and human rating in measuring the segmental pronunciation features (see Table 7).

Table 7. Summary of Independent Two-sample T-test

		Levene's Test for Equality of Variances		t-test for Equality of Means				
		F	Sig.	t	df	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference
								Lower / Upper
scores	Equal variances assumed	.201	.656	7.726	48	5.179	.670	3.831 / 6.527

In Table 7, the significant value of Levene's Test for the Equality of Variances was 0.656 with mean difference of 5.179, which was higher than the (α=0.05). This indicates the initial version of O'Speak has the equal ability with the human rating to measure segmental features. Given that understanding, as a prototype that was just developed, we might assume that O'Speak could be used as an alternative to assist student's independent learning and as a tool to assist teachers while measuring segmental features. This means that the tool has several potentials that could be explored in the future. \

Looking into the similar ability between the two, one possible answer to explain why the new tool could have relatively equal ability as compared to human rating is that O'Speak allows the participants to focus on words in isolation. O'Speak recognized their utterance as the input and coded the articulation to the algorithm and gave direct feedback correct/wrong answer. Using the same principles, O'Speak is also equipped with a consistent scoring system, allowing it to have the same capacity as the human rating.

Meanwhile, the human rater's scoring system shows some lack of accuracy and precision, increasing the potential to be biased and chance-error. Some contributing factors such as teaching experience, hallo effect, and rating experience are shown during the study. To have a clear picture, the following table informs how both measurement groups carried out their evaluation across four different tasks (see Table 8).

Access article distributed under the terms of the Creative Commons Attribution license (https://creativecommons.org/licenses/by-sa/4.0/)

Table 8. Summary of Segmental features in isolated words

Segmental focus	**Task**	**Discrimination of the segmental level**	
		O'Speak Group	**Human Rating Group**
consonant /s/ - /z/	dress	0.47	0.44
	sneakers	0.62	0.52
	blazer	0.41	0.59
	measles	0.47	0.41
	breeze	0.45	0.42
	miserable	0.69	0.53
consonant / ʃ/	polo shirt	0.58	0.61
	rash	0.53	0.43
	refreshing	0.54	0.42
consonant /tʃ/	Scorching	0.43	0.31
	tube train	0.52	0.34
consonant /p/ /b/	damp	0.52	0.49
	bomber	0.42	0.4l
consonant /g/ /k/	handbag	0.51	0.47
	headache	0.46	0.53
	broken leg	0.63	0.59
consonant /t/ /d/	tops	0.45	0.43
	mild	0.53	0.47
	frost	0.59	0.41
	sled	0.49	0.38
consonant /f/ /v/	cough	0.51	0.42
	fever	0.53	0.43
	vest	0.49	0.38
consonant /r/	jumper	0.58	0.56
consonant /j/	yacht	0.45	0.50
vowel /æ/	handbag	0.47	0.42
	asthma	0.41	0.39
	damp	0.46	0.34
	handcart	0.51	0.35
vowel /ʌ/	dull	0.45	0.42
	puddles	0.43	0.41
	submarine	0.62	0.46
	tugboat	0.59	0.43
vowel /ɪ/	singlet	0.70	0.56
	miserable	0.68	0.42
vowel /iː/	measles	0.59	0.56
vowel / ɜː/	sunburn	0.51	0.47
vowel /ə/	diarrhea	0.43	0.37
	asthma	0.41	0.34

Access article distributed under the terms of the Creative Commons Attribution license (https://creativecommons.org/licenses/by-sa/4.0/)

vowel /ɒ/	cough	0.65	0.54
	frost	0.57	0.43
	lorry	0.55	0.46
	yacht	0.49	0.47
vowel /e/	moped	0.39	0.38
diphthong /əʊ/	coat	0.42	0.30
	cold	0.49	0.41
	broken leg	0.51	0.29
	snowflakes	0.53	0.34
	moped	0.40	0.34
	fireboat	0.52	0.46
vowel /aɪ/	diarrhea	0.42	0.47
	mild	0.57	0.36
vowel/eɪ/	headache	0.53	0.46
triphthong /aɪə/	fireboat	0.56	0.48

Even though the development of O'Speak as one of automated pronunciation evaluation tools needs to be upgraded, Table 8 informs that O'Speak could elicit a greater degree of how participants had difficulty differentiating between voiced/voiceless labiodental fricatives /v/, /f/ (Dlaska & Krekeler, 2008), voiceless alveolar fricatives /s/, /z/. Meanwhile, the human rater's high-reliability consistency is not always in accordance with the definition of the accuracy of the measurements. Therefore, it confirms some arguments that human rating may have some pitfalls (Saito et al., 2015). In their study, Saito explains more possible factors that lead to lack of precision and accuracy, such as native language interference, rating experience, educational background, teaching experiences, and understanding pronunciation rating scales. Therefore, O'Speak could be utilized as one of the automated pronunciation evaluation tools to assist human rating.

CONCLUSION

The study used a quasi-experimental research design to compare the roles of O'Speak as an automated pronunciation rating tool and human rating to measure segmental pronunciation features. The findings inferred that O'Speak might help teachers recognize some areas of difficulty in English pronunciation and be referred to as a tool to measure the segmental pronunciation features in a more reliable and consistent procedure. The use of human ratings without any automated evaluation tools needs to be carefully considered since some possible conditions and variables affect the measurement that may lead to wrong judgment and interpretation. In order to improve the consistency and reliability of pronunciation metrics, educators and academics interested in this area may think about incorporating automated technologies such as O'Speak into their evaluation procedures. Furthermore, continuous training and calibration for human raters can assist reduce the drawbacks of human assessment and raise the accuracy of evaluations overall.

REFERENCES

Black, M. P., Bone, D., Skordilis, Z. I., Gupta, R., Xia, W., Papadopoulos, P., Chakravarthula, S. N., Xiao, B., Van Segbroeck, M., Kim, J., Georgiou, P. G., & Narayanan, S. S. (2015). Automated evaluation of non-native english pronunciation quality: Combining knowledge- and data-driven features at multiple time scales. Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH.

Access article distributed under the terms of the Creative Commons Attribution license (https://creativecommons.org/licenses/by-sa/4.0/)

Derakhshan, A., Khalili, A. N., & Beheshti, F. (2016). Developing EFL Learner's Speaking Ability, Accuracy and Fluency. English Language and Literature Studies. https://doi.org/10.5539/ells.v6n2p177

Dermentzi, E., Papagiannidis, S., Osorio Toro, C., & Yannopoulou, N. (2016). Academic engagement: Differences between intention to adopt Social Networking Sites and other online technologies. Computers in Human Behavior. https://doi.org/10.1016/j.chb.2016.03.019

Dlaska, A., & Krekeler, C. (2008). Self-assessment of pronunciation. System, 36(4), 506–516. https://doi.org/10.1016/j.system.2008.03.003

Drajati, N. A. (2018). Photovoice: Exploring the Role of Teacher's Question for University Students' Fluency in Speaking Class. JSSH (Jurnal Sains Sosial Dan Humaniora). https://doi.org/10.30595/jssh.v2i1.2160

Gilakjani, A., Ahmadi, S., & Ahmadi, M. (2011). Why is Pronunciation So Difficult to Learn? English Language Teaching. https://doi.org/10.5539/elt.v4n3p74

Gilakjani, A. P. (2016). English Pronunciation Instruction: A Literature Review. In International Journal of Research in English Education.

Guinto, N. L. (2013). Segmental Features of English modeled by selected professors in a state university in the Philippines: Implications in teaching English. International Journal of Research Studies in Language Learning. https://doi.org/10.5861/ijrsll.2013.475

Hassan, E. M. I. (2014). Pronunciation Problems: A Case Study of English Language Students at Sudan University of Science and Technology. English Language and Literature Studies. https://doi.org/10.5539/ells.v4n4p31

Hughes, R. (2011). Teaching and Researching: Speaking. In Teaching and Researching: Speaking. https://doi.org/10.4324/9781315833736

Hyman, L. (2019). Segmental phonology. In The Bantu Languages. https://doi.org/10.4324/9781315755946-4

Kuiken, F. (2023). Linguistic complexity in second language acquisition. Linguistics Vanguard. https://doi.org/10.1515/lingvan-2021-0112

Kukulska-Hulme, A. (2009). Will mobile learning change language learning? ReCALL. https://doi.org/10.1017/S0958344009000202

Kukulska-Hulme, A., & Viberg, O. (2018). Mobile collaborative language learning: State of the art. In British Journal of Educational Technology. https://doi.org/10.1111/bjet.12580

Kusumaningrum, W., Anggraeni, C., & Asmara, R. (2020). Envisaging the Needs for Designing L2 Pronunciation Assessment. https://doi.org/10.4108/eai.12-10-2019.2292208

Manca, S., & Ranieri, M. (2017). Implications of social network sites for teaching and learning. Where we are and where we want to go. Education and Information Technologies. https://doi.org/10.1007/s10639-015-9429-x

Marzá, N. E. (2014). Pronunciation and comprehension of oral english in the English as a Foreign Language class: Key aspects, students' perceptions and proposals. Journal of Language Teaching and Research. https://doi.org/10.4304/jltr.5.2.262-273

Miangah, T. M. (2012). Mobile-Assisted Language Learning. International Journal of Distributed and Parallel Systems. https://doi.org/10.5121/ijdps.2012.3126

Ramadhika, B., Karma, C. P. F., & Yosintha, R. (2023). Washback effects of Google Meet automatic caption feature in assessing students' pronunciation. EduLite: Journal of English Education, Literature and Culture. https://doi.org/10.30659/e.8.2.406-421

Ronanki, S., Bo, L., & Salsman, J. (2012). Automatic Pronunciation Evaluation and Mispronunciation Detection using CMU Sphinx. Proceedings of the Workshop on Speech and Language Processing Tools in Education. 24th International Conference on Computational Linguistics (COLING 2012) .

Access article distributed under the terms of the Creative Commons Attribution license (https://creativecommons.org/licenses/by-sa/4.0/)

Saito, K., Trofimovich, P., & Isaacs, T. (2015). Second language speech production: Investigating linguistic correlates of comprehensibility and accentedness for learners at different ability levels. Applied Psycholinguistics, 37(2), 217–240. https://doi.org/10.1017/S0142716414000502

Smith, R. (2004). The role of fine phonetic detail in word segmentation. Unpublished Doctoral Dissertation, University of Cambridge, UK.

Thornbury, S. (2007). How to teach speaking. Longman.

Tu, M., Grabek, A., Liss, J., & Berisha, V. (2018). Investigating the role of L1 in automatic pronunciation evaluation of L2 speech. Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH. https://doi.org/10.21437/Interspeech.2018-1350

van Leeuwen, A., & Janssen, J. (2019). A systematic review of teacher guidance during collaborative learning in primary and secondary education. In Educational Research Review. https://doi.org/10.1016/j.edurev.2019.02.001

Yates, L., Terraschke, A., Zielinski, B., Pryor, E., Wang, J., Major, G., Radhakrishnan, M., Middleton, H., Chisari, M., & Williams Tetteh, V. (2015). Adult Migrant English Program (AMEP) Longitudinal study 2011-2014: Final Report.

Zhang, L., Zhao, Z., Ma, C., Shan, L., Sun, H., Jiang, L., Deng, S., & Gao, C. (2020). End-to-end automatic pronunciation error detection based on improved hybrid ctc/attention architecture. Sensors (Switzerland). https://doi.org/10.3390/s20071809

Access article distributed under the terms of the Creative Commons Attribution license (https://creativecommons.org/licenses/by-sa/4.0/)