ISSN: 2289-7844 | e-ISSN: 0127-9750
DOI: https://doi.org/10.37134/jictie.vol11.2.1.2024

Journal of ICT in Education (JICTIE)
Volume 11, Issue 2, Pages 1-17, 2024

Research Article

Enhancing English Pronunciation Learning in Primary Education Through Natural Language Processing: A Quantitative Study

Mishalini Chandran[1], Shir Li Wang[1,2]*, Sumayyah Binti Dzulkifly[1], Theam Foo Ng[3], Amr S. Ghoneim[4]

[1]*Software Engineering and Smart Technology Department,*
[2]*Data Intelligent and Knowledge Management (DILIGENT),*
Faculty of Computing and Meta-Technology (FKMT), Sultan Idris Education University, Perak, Malaysia;
mishalini.chandran95@gmail.com, shirli_wang@meta.upsi.edu.my, sumayyah.dzul@meta.upsi.edu.my
[3]*Centre for Global Sustainability Studies (CGSS), Universiti Sains Malaysia, Penang, Malaysia; tfng@usm.my*
[4]*Faculty of Computers and Artificial Intelligence, Helwan University, Helwan, Egypt; amr.ghoneim@fci.helwan.edu.eg*

Received: 1 July 2024; Revised: 29 August 2024; Accepted: 15 September 2024; Published: 15 October 2024

**corresponding author*

Abstract

Learning English pronunciation is crucial for primary school pupils to enhance their communication skills, yet many struggle. This study evaluates the effectiveness of a natural language processing (NLP) tool in improving English pronunciation among primary school pupils by measuring motivation, confidence, and performance. Conducted in a primary school in Johor, the research used a randomised controlled trial (RCT) design with 30 Year 3 pupils, divided into two groups of 15 each: control and treatment. The English teacher assessed motivation, confidence, and performance using a rubric-based evaluation, with a Cronbach's Alpha value of 0.932, indicating high reliability. Paired t-tests were conducted to determine if the differences in motivation, confidence, and performance between the two groups were statistically significant. For motivation, t-values ranged from -1.871 to -1.468 for the traditional method and -6.548 to -13.229 for the NLP method. For confidence, t-values ranged from -1.871 to -1.468 for the traditional method and -11.500 to -7.135 for the NLP method. For performance, t-values ranged from -1.871 to -1.000 for the traditional method and -6.548 to -13.229 for the NLP method. With $\alpha = 0.05$, the mean differences in motivation, confidence, and performance were significantly different for the NLP method (all p-values < 0.05) but not for the traditional method (all p-values > 0.05). High partial η^2 values, ranging from 0.474 to 0.827, suggest that the teaching methods account for a significant proportion of variance in the dependent variables. With $\alpha = 0.10$, multivariate analysis of variance (MANOVA) assessed the effectiveness of the teaching methods. The combined dependent variables significantly differed based on teaching methods, Pillai Trace = 0.95, $F(24, 5) = 3.72$, $p < 0.074$, partial $\eta^2 = 0.95$. In conclusion, the study demonstrates that the NLP tool enhances English pronunciation learning by significantly improving motivation, confidence, and performance among Year 3 pupils.

Keywords: natural language processing, English pronunciation learning, primary education, educational technology, randomised controlled trials

INTRODUCTION

Malaysia's rich linguistic and ethnic diversity underscores the critical role of language in everyday communication and expression (Abdullah et al., 2021; McNamara et al., 2017). Among the fundamental aspects of language, pronunciation significantly influences how words are articulated. While traditional English language instruction in Malaysian primary schools adheres to conventional methods, modern pedagogical strategies have emerged, catering to individual pupil needs through interactive, collaborative, and activity-based approaches (Zhang et al., 2020).

Incorporating these contemporary teaching techniques, especially for young learners, can revolutionize English-language learning, making it more engaging and enjoyable while enhancing pupils' motivation and critical thinking skills, which are vital for their academic advancement (Santosa et al., 2019; Zolkipli et al., 2023). The intersection of education and technology is clear, with Artificial Intelligence (AI) playing a pivotal role in transforming learning experiences. Among AI technologies, Natural Language Processing (NLP) is prominent, influencing various areas of our lives, including language translation, spam detection, advanced search engines, and conversational systems (Ghafar et al., 2023; Azman et al., 2024). NLP is the field where machines learn to understand and communicate with human language — written or spoken. Consequently, the integration of NLP tools into education, particularly for English pronunciation learning, is rising (Liu et al., 2019).

This study explores the influence of teaching methods, specifically using an NLP tool (namely, the English Language Speech Assistant ELSA), on the motivation, confidence, and performance of Year 3 pupils in English pronunciation learning. The research aims to assess the impact of the NLP tool on these three aspects among primary school pupils. The study focuses on Year 3 pupils aged 9, underscoring the significance of early language learning for better adaptability and knowledge retention (Myles, 2017). The ultimate goal of this research is to gain insights into the motivation, confidence, and performance levels of Year 3 pupils using an NLP tool for English pronunciation learning, thereby contributing to the ongoing evolution of language education in Malaysia.

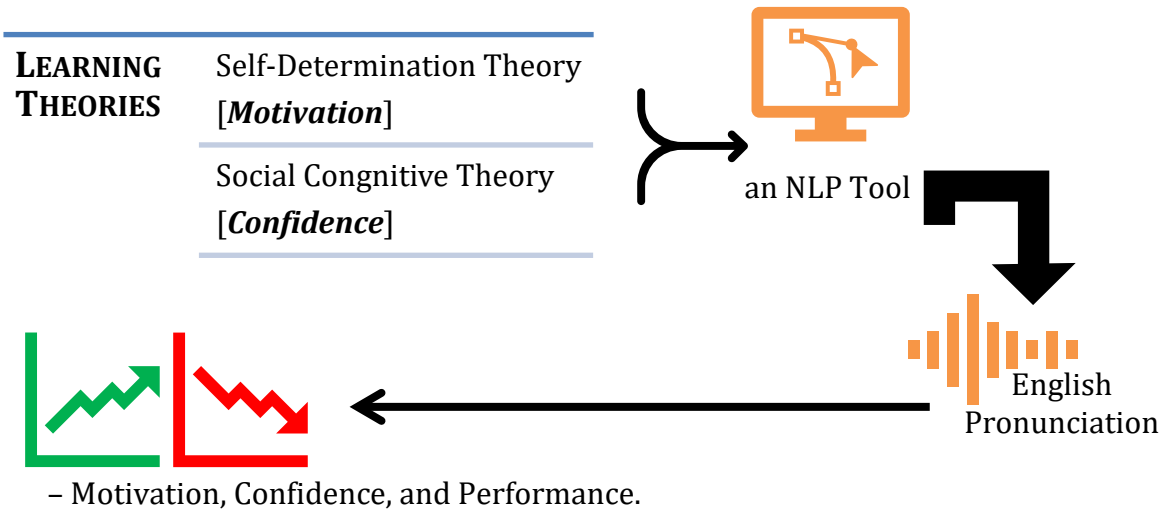

Figure 1: The conceptual framework of the conducted study

This study utilises the Self-Determination Theory (SDT) as its primary framework to understand the dynamics of motivation as shown in Figure 1. SDT explores motivation by focusing on human needs and well-being within societal contexts. The development of confidence is explored through the lens of the Social Cognitive Theory (SCT), which underscores the importance of self-efficacy (Devi et al., 2017).

The research hypothesises a link between using the NLP tool and increased pupil engagement and autonomy. This aligns with the principles of autonomy, competence, and relatedness emphasised in SDT, potentially leading to heightened motivation. The NLP tool's features of active participation, modelling, and feedback are expected to strengthen learners' confidence, given the tool's strong focus on interactive learner engagement and effective communication.

The expected result is improved performance, as SDT proposes that motivation is tied to increased learner engagement, persistence, and determination, especially when learners experience autonomy, competence, and social connections. Confidence is also expected to positively impact performance, with beliefs in self-efficacy and observational learning encouraging learners to apply their skills, take risks, and reach higher proficiency levels (Devi et al., 2017).

LITERATURE REVIEW

This literature review explores the intersection of educational theories, language acquisition, and the role of Natural Language Processing (NLP) in education. We delve into fundamental educational theories, specifically Self-Determination Theory (SDT) and Social Cognitive Theory (SCT). We also examine language learning challenges and discuss NLP's impact on education, language acquisition, and pronunciation enhancement.

Educational Theory: Self-Determination and Social Cognitive Theories

In educational theory, Self-Determination Theory (SDT), developed by Deci and Ryan in 1985, emphasizes autonomy, competence, and relatedness as key motivators within societal contexts. Conversely, Social Cognitive Theory (SCT) investigates environmental influences on behaviour, highlighting concepts like self-efficacy and observational learning (Manjarres-Posada et al., 2020). These theories provide valuable insights into motivation and confidence within educational settings. SDT focuses on autonomy, competence, and relatedness, while SCT examines environmental impacts and underscores self-efficacy and observational learning (Goldie, 2016; Bozkurt & Ataizi, 2015; Manjarres-Posada et al., 2020).

The Language Learning Challenge

Proficient language learning, particularly in English, is essential for Malaysian pupils (Nor et al., 2019). Pronunciation proficiency directly influences effective communication (Sa-ih, 2017). Low proficiency levels can lead to frustration, and teaching methods significantly impact pupil engagement (Sa-ih, 2017). Motivation and confidence play pivotal roles in language learning, with intrinsic motivation and confidence levels determining overall success (Dincer & Yesilyurt, 2017; Hasbullah et al., 2022).

Addressing language anxiety is crucial for nurturing confidence and sustaining motivation (Sa-ih, 2017; Nor et al., 2019; Nasir, 2014; Dincer & Yesilyurt, 2017). To overcome these challenges, integrating technology with traditional classroom methods is recommended (Annamalai et al., 2023; Samuri et al., 2016).

English Pronunciation Learning

Research on English pronunciation learning for Year 3 pupils reveals several effective approaches and challenges. Fluent speakers are identified by their ability to maintain good intonation, avoid unnecessary pauses, and speak clearly and effortlessly (Reswari, 2018). In this study, the sample consisted of 53 pupils from two classes, Class 3 Pugal and Class 3 Keerthi, who are all 9 years old and studying in Year 3.

Engaging pupils at this age is crucial, as early interest in learning English lays a strong foundation for future language proficiency (Rahberdi & Yusupovich, 2022). Additionally, parents generally support early language learning, recognizing its long-term benefits (Zhou, 2020). These pupils are capable of understanding instructions and participating effectively in research studies, making them ideal subjects for investigating pronunciation learning strategies**.**

NLP in Education

The integration of Natural Language Processing (NLP) into education offers promising solutions to various challenges. NLP techniques, including text analysis, grammar checkers, and linguistic software, provide valuable resources for enhancing language learning (Alhawiti, 2014). Moreover, NLP facilitates efficient content access and processing (for both educators and learners), particularly in language learning contexts (Alhawiti, 2014).

NLP in Language Learning and Pronunciation

Within language learning, NLP contributes to a deeper understanding of cognitive and psychological elements crucial for successful language acquisition (Tyagi & Singh, 2019). NLP also enables educators and learners to access multimedia content in English (Alhawiti, 2014). Notably, within the domain of pronunciation improvement, Computer-Assisted Pronunciation Tools (CAPT) and speech recognition technologies are gaining prominence.

Research shows that tools like Rosetta Stone greatly improve pupil motivation and engagement in pronunciation learning (Yuliani et al., 2023). However, while chatbots can meet psychological needs, they struggle with providing emotional support and accurate information (Annamalai et al., 2023).

English Language Speech Assistant (ELSA): An AI-powered NLP Tool

Among NLP tools designed for pronunciation enhancement, the English Language Speech Assistant (ELSA) stands out as a leading AI-powered platform. ELSA provides learners with feedback on intonation and pronunciation (Radar & Meunier, 2021), offering a variety of lessons and assessments tailored to different proficiency levels. Its contribution to English language learning makes it a valuable resource for learners seeking to improve their pronunciation skills.

METHODOLOGICAL APPROACH AND EXPERIMENTAL SETUP

This section contains 4 subsections which are research design, study design, assessment design and research execution.

Research Design: Quantitative Methods and Randomized Controlled Trials

This research adopts a quantitative approach, meticulously designed to collect accurate and reliable samples for statistical analysis (Queirós et al., 2017). By emphasizing objectivity, quantitative analysis enables quantifiable estimates of variables and inferences drawn from population samples. The data collected will be primary, allowing for rigorous numerical examination of research findings — a common choice in language testing and assessment research (Rahman, 2020).

Figure 2 depicts our chosen research method which is the Experimental Research Method. As explained by Rogers and Revesz (2020), this approach establishes causal relationships between independent and dependent variables. Specifically, our research design employs Randomised Controlled Trials (RCTs). RCTs are robust quantitative designs used to evaluate the effectiveness of interventions or treatments. By comparing outcomes between an experimental group (receiving the intervention) and a control group (not receiving the intervention), RCTs aim to establish causality.

Randomly assigning individuals (participants) to treatment or control conditions minimizes bias and ensures comparability (Gegenfurtner & Ebner, 2019). In the context of education, RCTs serve as valuable tools for evaluating outcomes related to interventions, teaching strategies, curricula, and other educational programs (Odom, 2021).

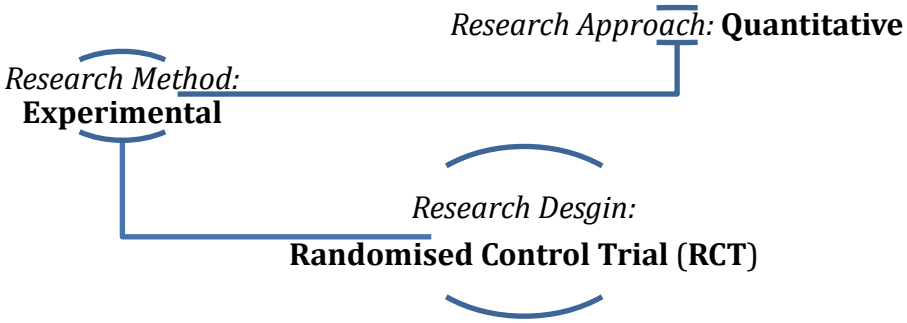

Figure 2: The overall research approach

Study Design: Participant Selection, Sampling Strategy, and Ethical Compliance

Our study focuses on primary school pupils learning English, specifically targeting 9-year-old Year 3 pupils from SJK Tamil Pasir Gudang — a Tamil National Type School with a total pupil population of 351. Igniting pupils' interest in learning English from an early age is critical in foreign language education, aligning with parental encouragement for early foreign language acquisition (Rahberdi & Yusupovich, 2022; Zhou, 2020).

To ensure representativeness, we employ a probability sampling method, specifically the Randomised Controlled Trials (RCT) design. As demonstrated in Figure 3, we randomly selected 15 pupils from the total population — from two Year 3 classes: Class 3 Pugal and Class 3 Keerthi — to form the control and treatment groups, respectively. The selection process is entirely random, without specific criteria, ensuring sample randomness, as Andrade (2021) advocates. All Year 3 pupils are eligible for participation,

emphasising the absence of exclusion criteria. Ethical integrity is paramount, and parental consent is obtained through informed consent forms, safeguarding the rights and well-being of participating pupils.

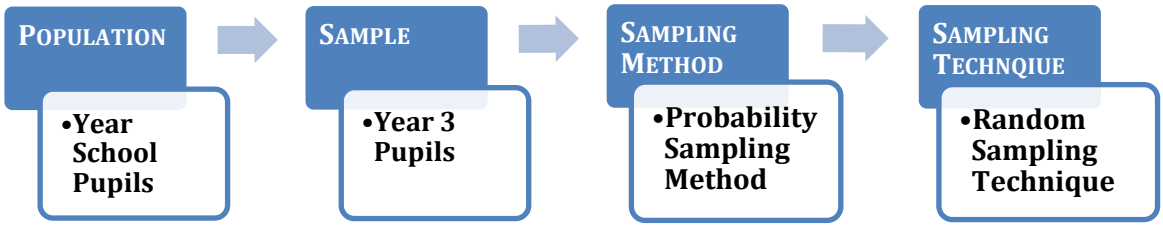

Figure 3: The population and sample

Assessment Design: Rubric Development and Reliability Testing

The assessment in this research is based on a rubric developed with reference to the Standards-Based Curriculum and Assessment Document (DSKP). According to the DSKP English Year 3 guidelines, there are eight criteria that Year 3 pupils should meet by the end of the year. Our evaluation rubric aligns with content standard 3.1, which involves recognising words in both linear and non-linear texts using knowledge of letter sounds, and content standard 3.2, which focuses on understanding a variety of linear and non-linear print and digital texts using appropriate reading strategies.

These contents are detailed on pages 30 and 31 of the DSKP English Year 3 document. Our developed rubric assesses the confidence, motivation, and performance levels of Year 3 pupils in English pronunciation. The teacher will be referring to the Performance Standards Guide for Reading Skills from page 32 of DSKP of English Year 3 for the pupil's evaluation, as shown in Table 1.

The rubric employs a 4-point Likert scale, with the following ratings: Excellent (4 points), Good (3 points), Fair (2 points), and Poor (1 point). Additionally, the teacher will refer to the Performance Standards Guide for Reading Skills from page 32 of the DSKP English Year 3 when evaluating pupils. Although the guide outlines six performance levels, our evaluation rubric utilises the first four scales.

The research is conducted over three weeks, and selecting only the first four performance levels for the Likert scale is appropriate. These levels provide a sufficient range to assess and capture the progression in reading skills without overwhelming the participants or diluting the data with excessive granularity. The reliability of the rubric is tested using the Cronbach's Alpha Test.

Rubric-Based Assessment Validation: The Cronbach Alpha Test

As Aithal & Aithal (2020) suggested, the Cronbach Alpha coefficient is a standard measure to assess internal consistency. A questionnaire is considered to have sufficient internal consistency if Cronbach's alpha value is 0.70 or higher. Conversely, a Cronbach's alpha score below 0.70 indicates weak internal consistency, implying a lack of correlation among the questions. To validate our rubric, we tested it on a random sample of 15 pupils from SJK Tamil Pasir Gudang. The Cronbach's Alpha result is presented in Table 2. With a total of 24 items — eight criteria each for assessing motivation, confidence, and performance — the obtained Cronbach's Alpha value is 0.932. This high value suggests that the evaluation rubric demonstrates reliability.

Table 1: The performance standards guide for reading skills

Performance Level	Descriptors for Reading Skills	Notes
1	- Hardly able to understand the main idea, specific information, and details of short simple texts even with a lot of support from the teacher. - Hardly shows an ability to guess the meaning of unfamiliar words and use dictionary skills even with a lot of support from the teacher. - Hardly shows the ability to read A1 short simple fiction or non-fiction print and digital texts of interest even with a lot of support from the teacher.	Requires support to achieve the curriculum target. (A1 Mid)
2	- Understands the main idea, specific information, and details of short simple texts with a lot of support from the teacher. - Able to guess the meaning of unfamiliar words and use dictionary skills with a lot of support from the teacher. - Able to read A1 short simple fiction or non-fiction print and digital texts of interest with much support from the teacher.	On track to achieve the curriculum target. (A1 Mid)
3	- Understands the main idea, specific information, and details of short simple texts adequately. - Able to guess the meaning of unfamiliar words and use dictionary skills adequately. - Reads A1 short simple fiction or non-fiction print and digital texts of interest adequately.	Achieves expectations of the curriculum target. (A1 Mid)
4	- Displays a good understanding of the main idea, specific information, and details of short simple texts. - Shows good ability to guess the meaning of unfamiliar words and use dictionary skills. - Shows good ability to read A1 short simple fiction or non-fiction print and digital texts of interest.	Working towards exceeding expectations of the curriculum target. (A1 High)

Table 2: Cronbach's Alpha test result for validating the rubric-based assessment

Reliability Statistics		
Cronbach's Alpha	**Cronbach's Alpha Based on Standardized Items**	**Number of Items**
0.932	0.937	24

Research Execution: Participant Grouping, Data Collection, and Teaching Approaches

Data collection for this study takes place at SJK Tamil Pasir Gudang, a primary school located in the Johor Bahru district of Johor state. The study spans three weeks and involves two Year 3 classes — specifically, Class 3 Pugal and Class 3 Keerthi. We randomly select 15 pupils from each class, evenly dividing them into control and treatment groups to ensure equal participant numbers. Our selection process prioritises pupils with comparable academic achievements to minimise potential bias.

The research process begins by providing clear instructions and guidance to the one and only English teacher of both classes. Consent forms will be distributed to the parents of the selected pupils. Upon receiving parental consent, the pupils will be grouped accordingly with the assistance of their teacher. The evaluation rubric, designed for assessing pupils, will be provided to the teacher. Evaluations will occur during the first and third weeks of the research period.

Next, we implement distinct teaching methods for English pronunciation. The control group will receive instruction using traditional methods without digital technology. In contrast, the treatment group will utilise the NLP Tool — specifically, the English Language Speech Assistant (ELSA) — for English pronunciation learning. Both groups will participate in one-hour lessons each week. The treatment group will practice English pronunciation with ELSA's assistance, while the control group will rely on conventional classroom instruction without ELSA.

Statistical Data Analysis and Interpretation: A MANOVA Approach with SPSS

Following data collection, we will analyse the data using a rubric-based evaluation as our assessment tool. For this purpose, we will employ the Statistical Package for the Social Sciences (SPSS) software. Our chosen method for data analysis is Multivariate Analysis of Variance (MANOVA). MANOVA assesses the significance of the impact of one or more independent variables on two or more dependent variables. In our research, we have two independent variables (with and without the NLP tool) and three dependent variables (motivation, confidence, and performance level). Given this setup, MANOVA is an appropriate statistical approach to test our hypothesis.

EXPERIMENTAL OBSERVATIONS AND RESEARCH FINDINGS

This section contains 6 subsections which are participation demographics and group formation, homogeneity test, pupil t-test, multivariate test, Tests of Between-Subjects Effects and discussion.

Participant Demographics and Group Formation

We surveyed a total of 30 participants, drawn from two Year 3 classes at SJK Tamil Pasir Gudang: Class 3 Pugal and Class 3 Keerthi. Notably, age was not a differentiating factor, as all respondents fell within the same age group. Additionally, the educational level was consistent across both classes, verified

through Levene's Test of Equality of Error Variances. Specifically, the control group comprised fifteen pupils from Class 3 Pugal, while the treatment group consisted of fifteen pupils from Class 3 Keerthi as shown in Table 3.

Table 3: Statistics of gender

Gender		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Male	13	43.3	43.3	43.3
	Female	17	56.7	56.7	100.0
	Total	30	100.0	100.0	

Homogeneity Test: Variance Analysis of Control and Treatment Groups

As outlined by Supena et al. (2021), the homogeneity test assesses whether there is heterogeneity or homogeneity in variance between the control and treatment groups. When data are homogeneous, both groups exhibit similar variance. We consider the variance of the data groups to be equal if the significance value (sig) exceeds 0.05. The results of the homogeneity variance test are presented for eight motivation, confidence, and performance criteria.

Examining Variance Homogeneity in Motivation

- For all motivation variables during the pretest (pretest1 to pretest8), $p > 0.05$. Consequently, we do not reject the null hypothesis of equal population variances. Thus, we can infer that the population variances are equal.
- Similarly, for all motivation variables during the posttest (posttest 1 to posttest 8), $p > 0.05$. Again, we do not reject the null hypothesis of equal population variances. Therefore, the population variances remain equal.

Overall, the motivation variables do not violate the homogeneity of variance assumption required for a t-test. Given this equality of variances within all groups for both pretest and posttest, we can proceed with a t-test to compare motivation levels.

Examining Variance Homogeneity in Confidence

- For all confidence variables during the pretest (pretest1 to pretest8), $p > 0.05$. Consequently, we do not reject the null hypothesis of equal population variances, affirming their equality.
- Similarly, for all confidence variables during the posttest (posttest 1 to posttest 8), $p > 0.05$. Again, we do not reject the null hypothesis of equal population variances. Hence, the population variances remain equal.

Notably, the confidence variables satisfy the homogeneity of variance assumption needed for a t-test. Given this consistency of variances across all groups for both pretest and posttest, we can proceed with a t-test to compare confidence levels.

Examining Variance Homogeneity in Performance

- For all performance variables during the pretest (pretest1 to pretest8), $p > 0.05$. Consequently, we do not reject the null hypothesis of equal population variances. Thus, we can infer that the population variances are equal.

- Similarly, for all performance variables during the posttest (posttest 1 to posttest 8), $p > 0.05$. Again, we do not reject the null hypothesis of equal population variances. Therefore, the population variances remain equal.

Overall, the performance variables satisfy the homogeneity of variance assumption required for a t-test. Given this equality of variances within all groups for both pretest and posttest, we can proceed with a t-test to compare performance levels.

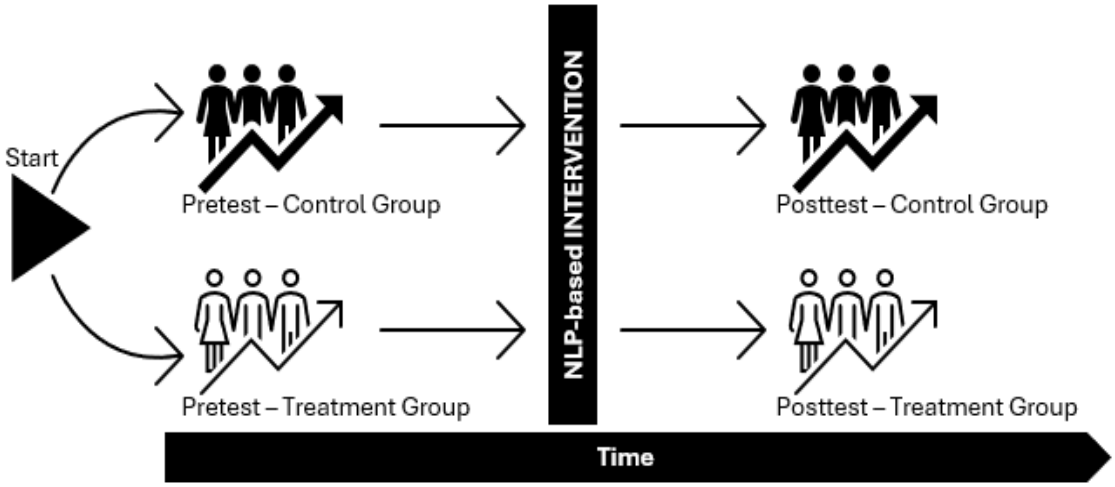

Figure 4: The experimental approach to evaluation (a randomised controlled trial)

Pupil t-Test: Statistical Significance of Control vs. Treatment Differences

We evaluated the effectiveness of NLP tools versus traditional teaching methods by comparing pretest and posttest scores across three dependent variables: motivation, confidence, and performance. Specifically, t-tests were conducted to determine whether statistically significant differences existed between the post-test results of the control group (traditional method) and the treatment group (NLP tools) as shown in Figure 4.

Comparing Motivation Levels: Control Group vs. Treatment Group

This section is divided into two parts: the pretest and posttest of the control group and the pretest and posttest of the treatment group, focusing on the motivation aspect.

a) Pretest and Posttest of Control Group

We observe that the p-values are greater than 0.05 for all criteria. Consequently, we do not reject the null hypothesis, which asserts no significant difference between the pretest and posttest motivation aspect for the control group. In other words, the difference in means between the two conditions is not statistically significant, and it is likely to have occurred due to chance. Therefore, we conclude that the mean difference between the two conditions is insignificant concerning the motivation aspect within the control group.

b) Pretest and Posttest of Treatment Group

We find that the significance level (α) remains at 0.05, but the p-values are now less than 0.05 for all criteria. Consequently, we reject the null hypothesis for the treatment group. The difference in means between the pretest and posttest motivation aspect is significant, and it is highly unlikely to have occurred merely due to chance. Therefore, we conclude that the mean difference between the two conditions is significantly different for the treatment group regarding motivation.

Comparing Confidence Levels: Control Group vs. Treatment Group

Similarly, this section is divided into two parts: the pretest and posttest of the control group, and the pretest and posttest of the treatment group, focusing on the confidence aspect.

a) Pretest and Posttest of Control Group

We find that the p-values are greater than 0.05 for all criteria. Consequently, we do not reject the null hypothesis for the control group. The difference in means between the pretest and posttest confidence aspect is not statistically significant, suggesting that it likely occurred due to chance. Therefore, we conclude that the mean difference between the two conditions is not significant regarding confidence within the control group.

b) Pretest and Posttest of Treatment Group

We note that the significance level (α) remains at 0.05, but the p-values are now less than 0.05 for all criteria. Consequently, we reject the null hypothesis for the treatment group. The difference in means between the pretest and posttest confidence aspect is significant, indicating that it is highly unlikely to have occurred merely due to chance. Therefore, we conclude that the mean difference between the two conditions is significantly different for the treatment group concerning confidence.

Comparing Performance Levels: Control Group vs. Treatment Group

This section is also divided into two parts: the pretest and posttest of the control group, and the pretest and posttest of the treatment group, focusing on the performance aspect.

a) Pretest and Posttest of Control Group

We observe that the significance level (α) is 0.05, and the p-values are greater than 0.05 for all criteria. Consequently, we do not reject the null hypothesis for the control group. In other words, the difference in means between the pretest and posttest performance aspects is not statistically significant. This

difference likely occurred due to chance. Therefore, we conclude that the mean difference between the two conditions for the control group regarding performance is not significantly different.

b) Pretest and Posttest of Treatment Group

We find that the significance level (α) remains at 0.05, but the p-values are now less than 0.05 for all criteria. Consequently, we reject the null hypothesis for the treatment group. The difference in means between the pretest and posttest performance aspects is significant, indicating that it is highly unlikely to have occurred merely due to chance. Therefore, we conclude that the mean difference between the two conditions is significantly different for the treatment group concerning performance.

Multivariate Test

To explore differences in pretest and post-test results between the control and treatment groups across motivation, confidence, and performance aspects, we conducted a one-way Multivariate Analysis of Variance (MANOVA) with the α value of 0.10. This analysis involved Year 3 pupils with and without treatment.

Our evaluation of assumptions related to homogeneity of variance yielded satisfactory results. Employing the Pillai Trace criterion, we found that the combined dependent variables significantly differed based on teaching methods (Pillai Trace = 0.95, $F(24, 5) = 3.72$, $p < 0.074$, partial $\eta^2 = 0.95$). Multivariate test results displayed Pillai's Trace, Wilks' Lambda, Hotelling's Trace and Roy's Largest root.

The choice of Pillai's trace is appropriate, as it performs well with small sample sizes and remains robust even when assumptions are violated. Notably, Wilks' Lambda and Hotelling's Trace are more sensitive to unequal covariances in smaller samples, while Roy's Largest root tends to result in Type I errors and should be avoided (Olson, 1976).

The significance value of the Pillai Trace is 0.074, which is less than 0.10. It should be noted that controversy surrounds using a 0.10 α value in MANOVA instead of the commonly used 0.05 α value, as some research suggests that it may lead to unacceptably low power under specific circumstances (Sheehan-Holt, 1998; Sheehan, 1995). However, Lynch (2018) demonstrated that both 0.05 and 0.10 α levels can yield satisfactory results in osteometric pair-matching analysis.

Additionally, novel techniques for merging p-values in MANOVA can maintain the nominal Type I error rate even when data are missing arbitrarily (Finch, 2016). As a result, this study's settings, as well as any possible effects on Type I error and power, may influence the choice of α level in a MANOVA. In summary, our findings indicate a significant difference between the pretest and posttest results of the control and treatment groups across motivation, confidence, and performance aspects.

Tests of Between-Subjects Effects

We employed Tests of Between-Subjects Effects to assess differences in pupil pronunciation across three critical aspects: motivation, confidence, and performance, comparing the control and treatment groups. These tests focus on comparing means for each dependent variable (motivation level, confidence level, and performance level) among groups exposed to varying levels of the independent variable (English pronunciation learning with or without NLP).

When the significance value falls below the α value of 0.10 it indicates a significant difference between pupil pronunciation and the three aspects. Our analytical results reveal that the treatment group and the control group exhibit significantly different mean scores for the dependent variables (performance, confidence, and motivation).

The strong statistical significance is evident from the extremely low p-values (Sig.) for each dependent variable, all significantly below the α value of 0.05.

- The F-values for motivation score differences (Diff_Motivation1 to Diff_Motivation8) range from 29.729 to 97.447, with associated p-values of 0.000. This suggests that employing NLP tools as part of the teaching method significantly influences pupil motivation.

- Similarly, the confidence differences' F-values (Diff_Confidence1 to Diff_Confidence8) range from 33.353 to 84.700, with p-values of 0.000. The teaching method significantly impacts pupil confidence.

- Regarding performance, the differences in scores (Diff_Performance1 to Diff_Performance8) yield p-values of 0.000, with F-values ranging from 25.200 to 84.700. This underscores the substantial effect of the teaching strategy on pupil performance.

The large effect size, indicated by high partial eta squared values (ranging from 0.474 to 0.827), suggests that the chosen teaching method can account for a significant proportion of the variance in the dependent variables. In summary, statistically significant findings across all measured features underscore the considerable impact of the teaching methodology on pupils' motivation, confidence, and performance.

Discussion

The research objectives centred on investigating the impact of Natural Language Processing (NLP) tool usage on motivation, confidence, and performance in English pronunciation learning among primary school pupils. Here, we delve into the findings and their implications:

- The findings reveal distinct patterns between pretest and posttest results, particularly regarding motivation. In the pretest phase, both the control and treatment groups exhibited similar motivation levels. This initial homogeneity may be attributed to consistent teaching methods before introducing the NLP tool. However, posttest results revealed a significant positive impact of NLP tool usage on motivation. Significance values for all motivation criteria in both groups fell below the 0.05 threshold. This indicates that after the intervention with the NLP tool, Year 3 pupils in the treatment group (Class 3 Keerthi) experienced a substantial increase in motivation compared to their counterparts in the control group (Class 3 Pugal).

- These findings underscore the pivotal role of autonomy, competence, and relatedness, as SDT emphasises, in fostering motivation among pupils. The NLP tool, ELSA, provided an environment where pupils could exercise autonomy over their learning, experience competence through feedback and self-monitoring, and establish relatedness by engaging with interactive learning materials. These factors collectively contributed to a significant boost in motivation among Year 3 pupils in the treatment group.

- Turning to the discussion of confidence, pretest results showed no significant differences in confidence levels between the control and treatment groups. Both groups started with comparable

confidence in their English pronunciation abilities. However, post-test results favoured the treatment group. Significance values for all confidence criteria in both groups were below 0.05, indicating a significant increase in confidence among Year 3 pupils in the treatment group compared to the control group.

- These findings align with SCT principles, particularly observational learning and self-efficacy. The NLP tool, ELSA, allowed pupils to listen to native pronunciations and imitate accents and intonations, fostering observational learning. Furthermore, the continuous feedback loop in NLP tools enhanced self-efficacy, as pupils saw measurable improvements in their pronunciation. This, in turn, significantly boosted their confidence levels.

- NLP tools significantly improved English pronunciation performance among Year 3 pupils. This improvement can be attributed to the combined effects of increased motivation and confidence due to NLP tool usage. Pupils who felt motivated and confident demonstrated enhanced pronunciation skills. These findings emphasise the potential of NLP tools like ELSA in enhancing language learning outcomes.

- Tests of between-subjects effects revealed p-values below 0.10 for motivation, confidence, and performance criteria when comparing the pretest and posttest results of the control and treatment groups. This indicates that the differences in pupils' motivation, confidence, and performance are significant, leading to the rejection of the null hypothesis H_0. Additionally, the multivariate analysis confirmed a significant difference (p-value < 0.10) using Pillai's Trace.

Thus, the NLP-based intervention positively impacted motivation, confidence, and performance levels in English language pronunciation learning among Year 3 pupils.

CONCLUSION

In summary, our research underscores the positive impact of Natural Language Processing (NLP) tools on English pronunciation learning among Year 3 pupils. By comparing NLP tool usage with traditional teaching methods, we observed significant motivation, confidence, and performance improvements. The role of SDT in language learning is evident, with NLP tools providing an environment that fosters pupils' motivation.

NLP tools like ELSA empower pupils by giving them control and enjoyment over their learning process. These tools offer competence-building opportunities through feedback and trial-and-error learning, enhancing motivation. The real-time scoring and feedback provided by NLP tools such as ELSA contribute to pupils' satisfaction and motivation as they track their progress in pronouncing.

Similarly, SCT is crucial in fostering confidence in English language learning among Year 3 pupils. NLP tools like ELSA enable pupils to listen to authentic pronunciations, allowing them to imitate accents and intonations confidently. Moreover, developing self-efficacy is essential for pupils' confidence in English pronunciation. Self-efficacy beliefs, both in pupils and educators, instil the confidence to imitate observed behaviours, reinforcing the positive impact of NLP tools.

Beyond motivation and confidence, NLP tools also significantly enhance pupils' performance in English pronunciation. The challenges faced by pupils and teachers in English language learning, particularly pronunciation, are well-documented. Technology is pivotal in overcoming these challenges in the contemporary education landscape. As our research shows, NLP tools represent a promising solution to

improve English teaching and learning.

FUTURE DIRECTIONS

This study underscores the potential of NLP tools to revolutionise language education, focusing specifically on English pronunciation among primary school pupils. Pan (2020) suggests that leveraging technology to enhance pupils' learning experiences, especially for out-of-class language study, while increasing their willingness to learn is critical in modern education.

To advance our understanding, future research should encompass various types of schools in Malaysia, increase sample sizes across different age groups, and conduct longitudinal studies to gauge the consistency of improvements in English pronunciation (that is, assess the consistency of improvements in English pronunciation over time). Additionally, exploring various NLP tools beyond ELSA for evaluating English pronunciation can provide a more comprehensive understanding of the potential and limitations of different technologies in language learning. In doing so, we can continue to advance our knowledge and contribute to more effective and engaging English language education for pupils.

ACKNOWLEDGEMENT

The authors would like to extend our gratitude to the management of SJK Tamil Pasir Gudang for permitting us and providing the necessary support to conduct our research with their Year 3 pupils. Their cooperation was invaluable in facilitating the successful completion of this study.

CONFLICTS OF INTEREST

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this article.

AUTHOR CONTRIBUTIONS

Mishalini Chandran: Conceptualization, Investigation, Original draft preparation. **Shir Li Wang:** Writing-Review and Editing, Supervision, Project Administration. **Sumayyah Binti Dzulkifly:** Data curation, Writing-Review and Editing. **Theam Foo Ng:** Writing-Review and Editing. **Amr S. Ghoneim:** Visualization, Writing-Review and Editing.

DECLARATION OF GENERATIVE AI

During the preparation of this work, the authors used ChatGPT to enhance the clarity of the writing. After using ChatGPT, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

DATA AVAILABILITY STATEMENT

All generated data included in this article is with proper references. The data supporting this research's findings are available from the corresponding author upon request.

REFERENCES

Abdullah, A. A., Awang, M. D., Kadir, M. N. A., Fadzil, A. F. M., & Hamzani, D. H. (2021). Ethics and civilization in Malaysian multiracial society. *International Journal of Academic Research in Business and Social Sciences, 11*(12). https://doi.org/10.6007/ijarbss/v11-i12/11940

Aithal, A., & Aithal, P. S. (2020). Development and validation of survey questionnaire and experimental data: A systematic review-based statistical approach. *International Journal of Management, Technology, and Social Science, 5*(2), 233–251.

Alhawiti, K. M. (2014). Natural language processing and its use in education. *International Journal of Advanced Computer Science and Applications, 5*(12). https://doi.org/10.14569/ijacsa.2014.051210

Andrade, C. (2021). The inconvenient truth about convenience and purposive samples. *Indian Journal of Psychological Medicine, 43*(1), 86–88.

Annamalai, N., Eltahir, M. E., Zyoud, S. H., Soundrarajan, D., Zakarneh, B., & Salhi, N. R. A. (2023). Exploring English language learning via Chatbot: A case study from a self-determination theory perspective. *Computers and Education Artificial Intelligence, 5*, 100148. https://doi.org/10.1016/j.caeai.2023.100148

Azman, N. S., Rahmatullah, B., Tamrin, K. F., & Qahtan, Y. M. (2024). A preliminary study on tech-based health advisory online system: The case study of UPSI computing students. *AIP Conference Proceedings, 2750, 040006*. https://doi.org/10.1063/5.0148931

Bozkurt, A., & Ataizi, M. (2015). English 2.0: Learning and acquisition of English in the networked globe with a connectivist approach. *Contemporary Educational Technology, 6*(2), 155–168.

Devi, B., Khandelwal, B., & Das, M. (2017). Application of Bandura's social cognitive theory in the technology-enhanced, blended learning environment. *International Journal of Applied Research, 3*(1), 721–724.

Dincer, A., & Yesilyurt, S. (2017). Motivation to speak English: A self-determination theory perspective. *PASAA: Journal of Language Teaching and Learning in Thailand, 53*, 1–25.

Gegenfurtner, A., & Ebner, C. (2019). Webinars in higher education and professional training: A meta-analysis and systematic review of randomized controlled trials. *Educational Research Review, 28*, 100293.

Goldie, J. G. S. (2016). Connectivism: A knowledge learning theory for the digital age. *Medical Teacher, 38*(10), 1064–1069.

Hasbullah, N. H., Rahmatullah, B., Mohamad Rasli, R., Khairudin, M., & Downing, K. (2022). Google Meet usage for continuity and sustainability of online education during pandemic. *Journal of ICT in Education*, 9(2), 46–60. https://doi.org/10.37134/jictie.vol9.2.4.2022

Liu, X., Xu, M., Li, M., Han, M., Chen, Z., Mo, Y., & Liu, M. (2019). Improving English pronunciation via automatic speech recognition technology. *International Journal of Innovation and Learning, 25*(2), 126. https://doi.org/10.1504/ijil.2019.097674

Manjarres-Posada, N., Onofre-Rodríguez, D. J., & Benavides-Torres, R. A. (2020). Social cognitive theory and health care: Analysis and evaluation. *International Journal of Social Science Studies, 8*, 132.

McNamara, D. S., Allen, L., Crossley, S., Dascalu, M., & Perret, C. A. (2017). Natural language processing and learning analytics. In C. Lang, G. Siemens, A. F. Wise, & D. Gasevic (Eds.), *Handbook of learning analytics* (pp. 93–104). Society for Learning Analytics Research.

Myles, F. (2017). Learning foreign languages in primary schools: Is younger better? *Languages, Society & Policy*. https://doi.org/10.17863/CAM.9784

Nasir, A. (2014). An analysis of relationship between English language anxiety, English language interest and English language achievement. *Journal of Language and Linguistic Studies, 5*(2), 78–90.

Nor, K. M., Razali, M. M., Talib, N., Ahmad, N., Sakarji, S. R., Saferdin, W. A. A. W., & Nor, A. M. (2019). Pupils' problem in learning English as a second language among MDAB pupils at UITM Malacca. *International Journal of Humanities, Philosophy, and Language, 2*(7), 01–12.

Odom, S. L. (2021). Education of pupils with disabilities, science, and randomized controlled trials. *Research and Practice for Persons with Severe Disabilities, 46*(3), 132–145.

Pan, X. (2020). Technology acceptance, technological self-efficacy, and attitude toward technology-based self-directed learning: Learning motivation as a mediator. *Frontiers in Psychology, 11*, 564294.

Queirós, A., Faria, D., & Almeida, F. (2017). Strengths and limitations of qualitative and quantitative research methods. *European Journal of Education Studies, 3*(9), 369–387.

Radar, L., & Meunier, F. (2020). Text-to-speech and Dys Vocal tools: How can they enhance the reading skills of dyslexic pupils in L2? *System, 91*, 102264. https://doi.org/10.1016/j.system.2020.102264

Rahberdi, R., & Yusupovich, I. V. (2022). Problems in learning English in different classes in secondary schools. *Thematics Journal of English Language Teaching, 6*(1), 15–22.

Rahman, M. S. (2020). The advantages and disadvantages of using qualitative and quantitative approaches and methods in language "testing and assessment" research: A literature review. *Journal of Education and Practice, 11*(4), 110–119.

Ghafar, N. A., Rahmatullah, B., Razak, N. A., Muttallib, F. H. A., Adnan, M. H. M., & Sarah, L. L. (2023). Systematic literature review on digital courseware usage in Geography subjects for secondary school students. *Journal of ICT in Education*, *10*(1), 26-39. https://doi.org/10.37134/jictie.vol10.1.3.2023

Reswari, G. P. A. (2018). English learning difficulties for multilingual pupils: A case study of an Indonesian pupil in learning English. *Culturalistics: Journal of Cultural, Literary, and Linguistic Studies, 2*(3), 56–65.

Rogers, J., & Revesz, A. (2020). Experimental and quasi-experimental designs. In *Routledge Handbook of Research Methods in Applied Linguistics* (pp. 164–176). Routledge.

Sa-ih, R. (2017). *The problems faced by Thai pupils at University of Muhammadiyah Malang in learning speaking* [Doctoral dissertation, University of Muhammadiyah Malang]. Universitas Muhammadiyah Malang Library. https://onesearch.id/Record/IOS4109.35797?widget=1&institution_id=136

Samuri, S.M., Abdul Ghani, H., Rahmatullah, B., & Ab Aziz, N. S. (2016). Sistem sokongan keputusan untuk menilai dan memantau prestasi guru: Kajian rintis di SMK Bachok, Kelantan. *Journal of ICT in Education, 3*, 55–72.

Santosa, J., Figueiredo, A. S., & Vieira, M. (2019). Innovative pedagogical practices in higher education: An integrative literature review. *Nurse Education Today, 72*, 12–17. https://doi.org/10.1016/j.nedt.2018.10.003

Supena, I., Darmuki, A., & Hariyadi, A. (2021). The influence of 4C (constructive, critical, creativity, collaborative) learning model on pupils' learning outcomes. *International Journal of Instruction, 14*(3), 873–892.

Taherdoost, H. (2016). Sampling methods in research methodology: How to choose a sampling technique for research. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3205035

Tyagi, G., & Singh, R. (2019, March). Implementing CALL system using natural language processing tools. In *Proceedings of 2nd International Conference on Advanced Computing and Software Engineering (ICACSE)*.

Yuliani, S., Khulaifiyah, K., & Idayani, A. (2023). Investigating pupils' motivation on the use of Rosetta Stone in learning English pronunciation. *AL-ISHLAH Jurnal Pendidikan, 15*(2), 2433–2440. https://doi.org/10.35445/alishlah.v15i2.3160

Zhang, A., Olelewe, C. J., Orji, C. T., Ibezim, N. E., Sunday, N. H., Obichukwu, P. U., & Okanazu, O. O. (2020). Effects of innovative and traditional teaching methods on technical college pupils' achievement in computer craft practices. *SAGE Open, 10*(4), 2158244020982986.

Zhou, Y. (2020). The influence of family on children's second language learning. *Journal of Language Teaching and Research, 11*(4), 599–605. https://doi.org/10.17507/jltr.1104.06

Zolkipli, N. Z., Rahmatullah, B., Samuri, S. M., Árva, V., & Pranoto, Y. K. S. (2023). 'Leave no one behind': A systematic literature review on game-based learning courseware for preschool children with learning disabilities. *Southeast Asia Early Childhood Journal, 12*(1), 79–97. https://doi.org/10.37134/saecj.vol12.1.7.2023