Available online at https://li02.tci-thaijo.org/index.php/ssstj

Suan Sunandha Science and Technology Journal

Evaluating the Effectiveness of English Speaking Conversation Practice for Communication: A Case Study of Students Enrolled in the English for Communication Course at Rajamangala University of Technology Tawan-ok Using the Random Forest Technique

Piyatida T.Chaisuwan [1], Kulwadee Tanwong [2], Paveena Meteevorakij [3*]

[1]Faculty of Science and Technology, Rajamangala University of Technology Tawan-ok,
43 Moo 6, Bang Phra, Sriracha, Chon Buri 20110, Thailand
[2]Faculty of Physical Education, Srinakharinwirot University,
63 Moo 7, Rangsit-Nakhon Nayok Road, Ongkarak, Ongkarak Nakhon Nayok 26120, Thailand
[3]Faculty of Humanities and Social Sciences, Rajamangala University of Technology Tawan-ok,
43 Moo 6, Bang Phra, Sriracha, Chon Buri 20110, Thailand
***Corresponding author e-mail: paveenameteevorakij@hotmail.com**

Received: *20 July 2023* / **Revised:** *12 October 2023*/ **Accepted:** *14 November 2023*

Abstract

The objectives of this research were twofold: 1) to compare the effectiveness before and after using a dialogue-based English speaking skills training model for effective communication, and 2) to identify guidelines for adjusting the English speaking skills training model using the random forest technique suited to students enrolled in the English for communication course at Rajamangala University of Technology Tawan-ok. The tools used in this research were 1) a dialogue-based English speaking skills assessment questionnaire, and 2) observation of English speaking behaviour for communication. The research findings indicated that the average scores before implementing the speaking skills training model were 3.11, which was considered good. After the training, the average scores increased to 4.30, indicating a significantly higher level of proficiency. The statistical test yielded a t-value of 15.871. Therefore, it could be concluded that the English speaking skills training model significantly enhanced students' English communication abilities at a significance level of .05. Regarding the guidelines for adjusting the English speaking skills training model using the random forest technique, the data classification results demonstrated an accuracy rate of 86.78%, a recall rate of 85.4%, and a precision rate of 83.13%. Additionally, 18 rules were identified as having significant correlations. The adjusted training model focused on various topics, including Greetings, Feelings and Emotions, Health, Asking for and Giving Directions, Weather, and University Life.

Keywords: Efficiency evaluation, Dialogue-based English speaking skills training, Data mining, Random forest technique

1. Introduction

From Article 34 of the ASEAN Charter, which states, 'The working language of ASEAN shall be English,' the English language plays a crucial role as an international language, emphasizing its use in government affairs and private sector businesses. This has been the case for almost ten years since 2012 (Ruksuan, 2022). Despite the awareness that English is a tool for academic success and increasing opportunities in the workforce, the proficiency level of English language communication among Thais remains significantly low, as evident from the results of English language proficiency exams. Thailand's English language skills are ranked lower than other countries in the Asian region. According to the EF English Proficiency Index (2022), which measures language proficiency through online tests with a sample group of at least 400 individuals, English language skills in Thailand are at a very low proficiency level. In terms of English proficiency, Thailand ranks 21st out of 24

countries in the Asian region, according to the EF English Proficiency Index (2022). The factors that contribute to low English language proficiency include individual characteristics, teaching techniques, and national educational policies (Hengsuko, Boonkao, Srihanam, & Manee, 2019). Additionally, there is a research study by Tangpattanakit (2020) that examines the influence of factors on English language communication. It found that basic knowledge, learning strategies, and individual attitudes contribute to English language communication. Therefore, it is necessary to explore teaching methods that enable learners to communicate in English in their daily lives. Teachers have the responsibility to encourage learners to express themselves in the language, utilize language activities as a tool for practice, and incorporate real-life situations to aid in the learning process. The statement is consistent with Rujichom (2018), which stated that communication skills, particularly speaking skills, are essential as they are used more frequently in daily life than other skills. Speaking ability facilitates the acquisition of other skills and contributes to better interpersonal interactions. Additionally, Ur (1996) emphasizes the importance of speaking skills as they demonstrate and reveal the speaker's knowledge and understanding of the language. Indeed, it is necessary to have standardized criteria to assess individuals' speaking skills in order for them to understand their own development. This is why individuals need to be aware of their initial level of English language communication skills to facilitate their appropriate personal development. It enables them to select suitable materials and processes for improvement.

Testing English-speaking skills can indeed improve students' communication abilities. When students are assessed on their speaking skills, it prompts them to actively use and apply their language knowledge. This process helps identify strengths and weaknesses, encourages practice and preparation, and ultimately leads to an enhancement in speaking proficiency through consistent assessment and practice. Therefore, testing English speaking skills for communication should include pre-training and post-training assessments to compare the effectiveness of the media and the process used. This objective measurement should be complemented by statistical methods to analyze whether learners' abilities have significantly improved compared to the traditional approach.One statistical test that can be used is the Independent-Sample t-test, which tests the hypothesis between two independent sample groups. This test determines if there is a significant difference in the means of the sample groups. The preliminary requirement is that the two sample groups must be independent and obtained through random sampling from a population with normally distributed variables. Additionally, the variable being tested should be on an interval or ratio scale. Alternatively, a Paired-Sample t-test can be used to compare data before and after within the same sample group. The preliminary requirement for this test is that there should be one sample group, obtained through random sampling from a population with normally distributed variables. The variable being tested should be on an interval or ratio scale, and the independent variable should be related to the dependent variable (Kanjanawasee, Pitayanon, & Srisukho, 2008). However, the statistical measurements mentioned only provide information about the difference in English communication skills before and after the intervention. If we want to delve deeper into the content that was used to enhance skills and design learning strategies for other groups, researchers may consider utilizing data mining techniques to further enhance the clarity of the research findings and improve the reliability of their applications. For example, the research by Meteevorakij, Chaisuwan, Nubpetchploy, and Tanwong (2020) using decision trees to predict English language communication abilities for business communication based on 21st-century learning skills had an accuracy of 76%. Additionally, there is a study by Huynh-Cam, Chen, and Le (2021) that used decision trees and random forest algorithms to predict comprehensive teaching status. They predicted indicators to forecast performance for strategic planning in teaching management before the start of a new academic term. The study found that the prediction accuracy was 79.99% for random forest and 74.59% for decision trees. This research aligns with the study conducted by Madaan, Kumar, Keshri, Jain, and Nagrath (2021), which compared random forest and decision trees on the same dataset. The conclusion drawn from the results was that the random forest algorithm outperformed decision trees in terms of accuracy. It is a method of organizing information into nodes and branches. Each node represents a feature test. And the results of the test depend on the specified characteristics or independent variables. Building this model involves dividing data into groups based on specified conditions. Continue this process until you have data groups that correspond to the target variables you want to predict. If the amount of data is too small, this can cause the model to be unable to accurately predict data that has not been encountered before and cause overfitting. Although decision trees can be straightforward to interpret, random forest focuses on diverse features and variations that occur, enabling it to capture complex feature patterns more

effectively than decision trees. Additionally, it reduces the risk of overfitting by preventing excessive reliance on training data. This contributes to enhancing prediction accuracy. This perspective is supported by the research of Denisko and Hoffman (2018). Therefore, researchers have adopted these techniques to aid in the analysis of measuring the effectiveness of English conversation training for communication using the random forest technique. Random forest is a data mining technique that combines multiple non-repetitive decision trees to address the inflexibility of individual decision trees and reduce overfitting issues. (Jayaprakash, Krishnan, & Jaiganesh, 2020). Moreover, it is a machine learning technique that can be used for classification, regression, and other tasks. It creates multiple decision trees during training and outputs the class selected by most trees for classification tasks or the mean or average prediction of the individual trees for regression tasks. Random Forest is a popular algorithm due to its simplicity, flexibility, and ability to handle both classification and regression problems. This is consistent with the research conducted by Kaewpanitch (2020), which found that the model developed using the random forest technique provided higher accuracy in predicting the performance of students in English language learning compared to the decision tree technique.

This can help in adjusting the topics of English conversation training for communication to be more suitable for students studying English for Communication at Rajamangala University of Technology Tawan-ok. It also enables instructors to plan their teaching and learning strategies in English language courses for communication that are tailored to future learners.

2. Research Objective

2.1 To compare the effectiveness of a dialogue-based English speaking skills training model for communication purposes before and after its implementation.

2.2 To identify guidelines for adapting the dialogue-based English speaking skills training model, using the random forest technique, to cater to the specific needs of students enrolled in the English for communication course at Rajamangala University of Technology Tawan-ok.

3. Research Scope

3.1 The research scope of this study included the population of 150 students enrolled in the English for Communication course during the first semester of the academic year 2022 at Rajamangala University of Technology Tawan-ok Bangphra Campus. Subsequently, the sample size was calculated using the G*power program version 3.1.9.7. The size of the effect (Effect size) was set 0.5, the significance level α (α err prob) is 0.05 and level of power (Power (1-β err prob)) is 0.95. Due to the lack of knowledge regarding estimated parameters from previous studies, the approach of specifying a standardized effect size was chosen. The medium effect size 0.5 (Cohen, 1988) was selected, resulting in a minimum total sample size of 45 individuals. Therefore, this research determined the sample size to be 50. The study sample consisted of 50 individuals who registered for the English for Communication course during the same semester. The sample group was selected through purposive sampling. This method was used because the researchers need the sample group that had similar basic knowledge of English language and to avoid bias that could arise from different evaluators. Then simple random sampling was used subsequently. The research was conducted at the Bangphra campus (The Office of Academic Support and Registration Rajamangala University of Technology Tawan-ok, 2005).

3.2 The research duration involved conducting the study with students enrolled in the English for Communication course at Rajamangala University of Technology Tawan-ok, Bangphra Campus, during the first semester of the academic year 2022. The study commenced in May 2022 and concluded in September 2022, with a total of 50 participants.

3.3 The research tools used in this study included.

3.3.1 The dialogue-based English speaking skills training model for communication, which comprised six topics: Greeting-Basic, Feeling and Emotion, Weather, Health, University Life, and Asking and Giving Directions. These topics were derived from a referenced textbook by Beatty, Longshaw, and Austin (2020). For each topic, students received materials from the instructor to practice and apply independently.

3.3.2 The observation of English speaking behavior for communication, which encompassed three aspects: Pronunciation, Grammar, and Fluency. These aspects were based on the core lessons and criteria

developed following Clark's framework (Stansfield, 1973). The criteria were divided into four levels, ranging from beginner to advanced proficiency. The levels are as follows:

(1) Pronunciation: speaking or responding with accurate pronunciation.

(2) Grammar: speaking without grammar mistakes and using vocabulary and grammar structures accurately.

(3) Fluency: speaking conversationally with natural fluency.

3.3.3 Comparing the proficiency scores in English speaking skills for communication using the Likert scale criteria of 4 levels.

4. Research Framework

Based on the theories and concepts of Carroll (1980), Ellis and Johnson (1994), individuals with a good knowledge of the English language may still struggle to respond or communicate effectively when asked for directions or in various social situations. This can be considered a failure of the Human Introduction Processing System in language learning, particularly in English. Therefore, learners need both knowledge and practice in using English to develop proficiency, especially in speaking skills. Many students begin learning English in secondary school but are unable to communicate effectively with native speakers, which leads to a lack of confidence and hinders their progress in professional or higher-level learning contexts. Dell Hymes, a sociolinguist from the United States, highlights the significance of communicative competence, which focuses on enabling language learners to use the language appropriately in diverse social contexts. This approach equips learners with language knowledge and skills simultaneously, enabling effective progress in their careers or further studies (Richards & Rodgers, 2002).

For this research, the researchers have integrated both of the aforementioned frameworks and applied them to examine the development of English language communication abilities among students at Rajamangala University of Technology Tawan-ok.

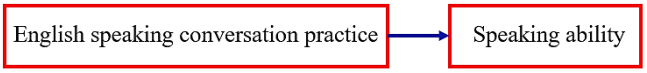

Figure 1. Conceptual Framework.

5. Research Methodology

This research employed a quasi-experimental design, specifically a one-group pretest-posttest design. The researchers outlined the following steps for conducting the study:

5.1 Participants: The study involved 50 students who were enrolled in the English for Communication course during the first semester of the 2022 academic year.

5.2 Process of creating and finding the quality of tools involved the following steps:

5.2.1 Conversation Speaking Exercise:

(1) Reviewed relevant documents and studies on speaking activities for communication purposes, such as "A study of out-of-class English language learning activities of secondary school students in English Program" by Wiengnil (2010) and the book "Practice Makes Perfect: English Conversation Premium, Second Edition" by Yates (2016).

(2) Created an English conversation speaking exercise based on the curriculum and content framework of the English Communication course. The exercise focused on 6 topics: Greeting-Basic, Feeling and Emotion, Weather, Health, University Life, and Asking and Giving Directions.

5.2.2 Created an observation form for English speaking behavior for communication purposes. The form utilized a pretest and posttest format and established assessment criteria based on Clark's evaluation framework.

5.2.3 Assessed the content validity and the measurement and evaluation of the training and observational behavior of English communication skills by experts. Calculated the Index of Congruence (IOC) based on the analysis conducted by three experts. The results showed that the average IOC values for all items were greater than 0.6, indicating that the training materials were appropriate and aligned with the objectives (Thaweerat, 1997).

5.3 Data collection was conducted by the researchers following the sequential steps below:

5.3.1 A pretest was conducted using an oral test format, consisting of 6 topics, to assess English speaking skills for communication. The topics were provided along with vocabulary prompts and applicable sentence structures. The evaluation was performed by two assessors, including one foreign language instructor and one researcher, totaling two assessors.

5.3.2 The researchers conducted the experiment using a training model consisting of 6 dialogue-based topics for English speaking skills in communication. The topics were arranged in the following order: Greeting-Basic, Feeling and Emotion, Weather, Health, University Life, and Asking and Giving Directions. The researchers taught the vocabulary and grammatical structures to the students. The students then applied the vocabulary they had learned into the grammatical structures they had studied. Next, the researchers asked the students to read and translate the created conversation sentences, and the researchers explained the pronunciation, translation and overall meaning of the sentences. Finally, the students were provided suggestions for creating new sentences using the vocabulary and grammatical structures learned. The training emphasized practice and transfer of knowledge, with one topic taught per week, following the order of the aforementioned topics. The process workflow was as follows:

(1) Introduced Topic 1: Greeting-Basic.

(2) Explained the content of the lesson before providing the practice exercises.

(3) The sample group practiced the dialogue exercises for a duration of 1 week.

(4) At the end of the week, the performance of the dialogue practice in that specific topic was assessed.

(5) Continued this process, sequentially moving through all 6 topics without repeating any previous topics.

The sample group had freedom to independently engage in the training process with the assigned lesson content.

5.3.3 The posttest assessment involved using the same conversational English practice materials as the training. Each individual was assessed through an oral test. This test was conducted after the learners had practiced each lesson on their own.

5.3.4 The scores obtained from the pretest and posttest assessments, following Clark's framework, were averaged and compared to evaluate the proficiency in English speaking for communication purposes. The results were then translated using Likert's criteria for assessing levels of speaking proficiency in English.

5.4 In data analysis, the researchers conducted the following steps:

5.4.1 Descriptive statistics: Employed descriptive statistics to analyze the general data.

5.4.2 Compared pretest and posttest scores in English speaking ability using the Paired-Samples t-test and performed a preliminary assumption check:

(1) Examined whether the dependent variable, which is the posttest scores of English speaking ability following the use of the conversational training program, followed a normal distribution.

(2) Examined whether the independent variable, which is the pretest scores of English speaking ability prior to the use of the conversational training program, had a correlation with the dependent variable.

5.4.3 Adjusted the conversational training program for English speaking communication using the random forest technique, which involves combining multiple non-repetitive decision trees to work together.

(1) The process involved understanding the data, defining the attributes, and analyzing the data for the conversational training program for English language with 6 topics and the observation of English-speaking behavior across 3 dimensions.

(2) The Data Preparation Phase involves transforming the data into a format that is compatible with "Nodes" and lists the sample for each layer. During the creation of a decision tree, the objective is to select the attribute/feature with the lowest Gini index as the root node. The Gini coefficient has a range from zero to one. A value of zero represents a system of perfect equality across the population (all nodes are in the same class), while a value of 0.5 indicates distributed sampling. The equal data in the two "value" classes at the end come from samples that satisfy the output condition. The formula for the Gini index is as follows (Dorfman, 1979).

$$G_i = 1 - \sum_{k=1}^{n} p^2 i, k$$

'pi' represents the probability of an object being classified into a specific class.

pi,k denotes the proportion that indicates the ratio of the number of data instances in Node i belonging to Class k.

(3) Design: The random forest model utilizes ensemble learning, which combines multiple approximate value estimation models to reduce bias, variance, and data sensitivity.

(4) Dataset: The "S_Daily English Conver" dataset is processed using bootstrapping. Bootstrapping involves randomly sampling new subsets from the original dataset, replacing and increasing its size. This process helps generate sufficient data for model building. The dataset is then divided into training data and testing data following the principles of "k-fold Cross Validation (k-fold CV: k=10)." To conduct the model, researchers utilize data for training with a split of 90% for each round as training data, and a split of 10% is reserved as testing data for evaluation. This cross-validation technique assists in assessing the model's performance and its ability to generalize.

(5) Pseudo code for optimizing parameters using Grid Search (English Conversation Practice and English Speaking Behavior Observation: ECP_ESB)

Input: Database Sex, Chapter, Observation, Level Skill,
Output: Ranking Chapter
Method: # load sample data
Training-Data= "x_Eng C"
(x_Eng C_train, y_Eng C_train, z_Eng C_train,model)
While Attribute Num>0 Do
For Each Attribute Chap, Do
Sorted Data=Sort Training Data By (sex,)
sexAdvantagen [i] = Define Advantage (sex, SortedData)

```
 procedure division.
  display " " with blank screen.
  perform loop1 varying i from 1 by 1 until i > 5.
  perform loop2 varying i from 1 by 1 until i > 5.
  display "Max : " max.
     stop run.
     loop1.
  display i  " : " with no advancing.
      accept ar(i).
      loop2.
  if ar(i) > max
      move ar(i) to max.
```

End For …..
End While

(6) Preview: decision tree Models are used to predict the next response based on independent learning from each tree, such as the "Greeting-Basic" model shown in the example figure 2.

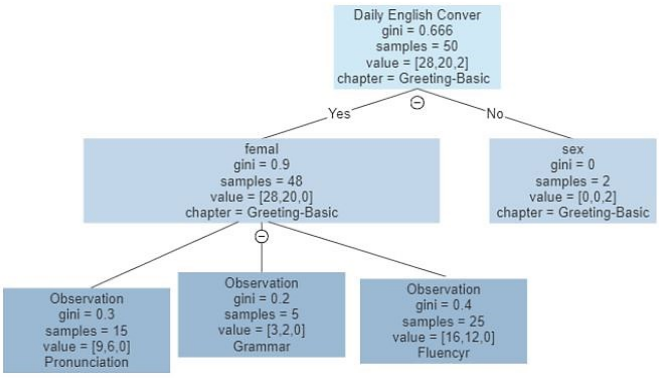

Figure 2. Illustration of the decision tree model "Greeting-Basic" creation.

(7) Evaluate the performance of the model: The accuracy of evaluated classification models is informally defined as the fraction of correct predictions made by our model. However, validity can be defined in the following ways:

$$Accuracy = \frac{Number\ of\ correct\ predictions}{Total\ number\ of\ predictions}$$

For binary classification, accuracy can also be calculated in terms of positive and negative values as follows:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

where TP = True Positives, TN = True Negatives, FP = False Positives and FN = False Negatives. (Table 6)

6. The Results of the Data Analysis

The results of the data analysis, classified according to Objective 1, which aims to compare the effectiveness before and after using the English conversation training, are as follows:

Table 1. English communication abilities before and after using English conversation training.

Average Level	Meaning	Number		Percent	
		Before	After	Before	After
1.00-2.00	Low ability	8	-	16.00	-
2.01-3.00	Medium ability	16	-	32.00	-
3.01-4.00	Good Ability	12	20	24.00	40.00
4.01-5.00	Very good ability	14	30	28.00	60.00
Total		50	50	100.00	100.00

According to Table 1, the finding of English communication abilities before using English conversation training revealed that the majority of students in the sample group had a moderate level of English communication skills prior to the test (32.00%). The next highest category was a very good level (28.00%), followed by a good level (24.00%), and a low level (8.00%) of English communication skills prior to the test.

Moreover, it is evident that more than half of the sample group, specifically 60.00%, achieved a very good level of English communication abilities after the test. Additionally, a substantial portion, 40.00%, attained a good level of English communication abilities after using English conversation training.

Table 2. The average scores before and after using the English conversation training program for each section.

Test Results	Greeting-Basic	Feeling and Emotion	Weather	Health	University life	Asking and Giving Directions
Pretest	4.06	3.58	2.37	3.23	2.20	3.23
Posttest	4.86	4.70	3.96	4.41	3.66	4.20

According to Table 2, it is evident that the sample group of students had the highest average score in the "Greeting-Basic" section (4.06) before using the English conversation training program. The second highest average score was observed in the "Feeling and Emotion" section (3.58), followed by the "Health" and "Asking and Giving Directions" sections (3.23) respectively. The lowest average score was obtained in the "University Life" section (2.20).

Regarding the average scores after using the training program, the highest average score was achieved in the "Greeting-Basic" section (4.86), followed by the "Feeling and Emotion" section (4.70) and the "Health" section (4.41). The lowest average score was observed in the "University Life" section (3.66).

Table 3. The results of the post-training scores in English conversation speaking ability follows a normal distribution.

Test	Skewness		Kurtosis		Kolmogorov-Smirnov	
Results	Statistic	Std. Error	Statistic	Std. Error	Kolmogorov-Smirnov Z	Sig
Posttest	-.467	.337	-1.211	.662	1.343	.054

*p<.05

According to the results from Table 3, the post-training scores in English conversation speaking ability showed a distribution that was close to normal. This conclusion is supported by the values of Zskewness = -1.386, Zkurtosis = -1.829, and Kolmogorov-Smirnov Z = 1.343 (sig = 0.054). Therefore, we can conclude that the variable "post-training scores in English conversation speaking ability" exhibited a distribution that closely resembles a normal distribution (Hair, Black, Babin, & Anderson, 2019).

Table 4. The results of the evaluation show that the scores before and after using the conversational English speaking training program for communication.

Test results	N	Correlation	Sig.
Pair 1 Posttest & Pretest	50	.812	.000

*p<.05

Based on Table 4, it is evident that there exists a significant correlation between the scores before and after implementing the conversational English speaking training program for communication (r = .812, sig = .000). Consequently, considering the information presented in Tables 3 and 4, it can be concluded that the data aligns with the initial assumptions of the statistical analysis utilizing the Paired-Samples t-test (Kanjanawasee et al., 2008).

Table 5. Comparing the average scores before and after using the conversational English speaking training program for communication in the sample group.

Test Results	N	$\bar{x}$	SD	df	t	p
Pretest	50	3.11	0.87	49	15.871	.000
Posttest	50	4.30	0.57			

*p<.05

According to Table 5, it is evident that among the students in the sample group who enrolled in the "English for Communication" course during the 1st semester of the academic year 2022, there is a statistically significant difference in the average scores before and after utilizing the conversational English speaking training program. The calculated test statistic was 15.871, with degrees of freedom equal to 49, and a significance value of .000. Consequently, it can be concluded that following the training in conversational English speaking, there was a substantial enhancement in English language communication skills at a significance level of .05.

Regarding the analysis of data for Objective 2, which aims to establish guidelines for adapting the conversational English speaking training program for effective communication using the random forest technique, the following information is provided:

Table 6. The performance measurement of the conversational English speaking training data classification using the random forest technique.

	Pronunciation	Grammar	Fluency	**Precision**
Pronunciation	842	0	0	82.6%
Grammar	0	809	5	80.1%
Fluency	0	0	904	86.7%
Recall	85.2%	88%	83%	
Accuracy		86.78%		

According to Table 6, the performance measurements for the classification of conversational English speaking training data using the random forest technique are as follows: The accuracy rate was 86.78%, the recall rate was 85.4%, and the precision rate was 83.13%.

Table 7. The classification of conversational English speaking training data is done for each of the 6 topics that have an impact on improving the test scores of students majoring in English for international communication.

Data classification rules	Lesson
If Chap = 1 and Eva_Ob = 1 and Level = 4	Greeting-Basic
If Chap = 1 and Eva_Ob = 2 and Level = 3	Greeting-Basic
If Chap = 1 and Eva_Ob = 3 and Level = 4	Greeting-Basic
If Chap = 2 and Eva_Ob = 1 and Level =3	Feeling and Emotion
If Chap = 2 and Eva_Ob = 2 and Level = 3	Feeling and Emotion
If Chap = 2 and Eva_Ob = 3 and Level = 4	Feeling and Emotion
If Chap = 3 and Eva_Ob = 1 and Level = 2	Weather
If Chap = 3 and Eva_Ob = 2 and Level = 2	Weather
If Chap = 3 and Eva_Ob = 3 and Level = 3	Weather
If Chap = 4 and Eva_Ob = 1 and Level = 3	Health
If Chap = 4 and Eva_Ob = 2 and Level = 2	Health
If Chap = 4 and Eva_Ob = 3 and Level = 4	Health
If Chap = 5 and Eva_Ob = 1 and Level = 2	University life
If Chap = 5 and Eva_Ob = 2 and Level = 2	University life
If Chap = 5 and Eva_Ob = 3 and Level = 2	University life
If Chap = 6 and Eva_Ob = 1 and Level = 3	Asking and Giving Directions
If Chap = 6 and Eva_Ob = 2 and Level = 2	Asking and Giving Directions
If Chap = 6 and Eva_Ob = 3 and Level = 3	Asking and Giving Directions

Data representation for data transformation is as follows:

Chapter (Chap) represents 1= Greeting-Basic, 2 = Feeling and Emotion, 3 = Weather, 4 = Health, 5 = University life, 6 = Asking and Giving Directions

Observation (Eva_Ob) represents 1 = Pronunciation, 2 = Grammar, 3 = Fluency

Level Skill (Level) = Pronunciation, Grammar, and Fluency represents each Observation as follows:

Most 4, Moderate 3, Low 2, Very low 1

From Table 7, we can observe that the data used for learning to classify future data types based on the training set has generated 18 rules with interrelated relationships. These rules are described as follows:

Rule 1: Students who received training in the Greeting-Basic topic and have a pronunciation proficiency level of 4 (able to pronounce correctly) tend to have excellent communication outcomes.

Rule 2: Students who received training in the Greeting-Basic topic and have a grammar proficiency level of 3 (able to use grammar structures correctly with minor errors) tend to have good communication outcomes.

Rule 3: Students who received training in the Greeting-Basic topic and have a fluency proficiency level of 4 (speak conversationally and fluently like a native speaker) tend to have natural and continuous communication outcomes.

Rule 4: Students who received training in the Feeling and Emotion topic and have a pronunciation proficiency level of 3 (occasional pronunciation errors) tend to have communication outcomes with minor pronunciation errors.

Rule 5: Students who received training in the Feeling and Emotion topic and have a grammar proficiency level of 3 (able to use grammar structures correctly with minor errors) tend to have communication outcomes with minor grammar errors.

Rule 6: Students who received training in the Feeling and Emotion topic and have a fluency proficiency level of 4 (speak conversationally and fluently like a native speaker) tend to have natural and continuous communication outcomes.

Rule 7: Students who received training in the Weather topic and have a pronunciation proficiency level of 2 (pronunciation errors that make understanding difficult) tend to have communication outcomes with pronunciation difficulties.

Rule 8: Students who received training in the Weather topic and have a grammar proficiency level of 2 (basic grammar errors) tend to have communication outcomes with minor grammar errors.

Rule 9: Students who received training in the Weather topic and have a fluency proficiency level of 3 (speak conversationally and fluently with occasional hesitations) tend to have natural and continuous communication outcomes with occasional hesitations.

Rule 10: Students who received training in the Health topic and have a pronunciation proficiency level of 3 (occasional pronunciation errors) tend to have communication outcomes with minor pronunciation errors.

Rule 11: Students who received training in the Health topic and have a grammar proficiency level of 2 (basic grammar errors) tend to have communication outcomes with minor grammar errors.

Rule 12: Students who received training in the Health topic and have a fluency proficiency level of 4 (speak conversationally and fluently like a native speaker) tend to have natural and continuous communication outcomes.

Rule 13: Students who received training in the University Life topic and have a pronunciation proficiency level of 2 (pronunciation errors that make understanding difficult) tend to have communication outcomes with pronunciation difficulties.

Rule 14: Students who received training in the University Life topic and have a grammar proficiency level of 2 (basic grammar errors) tend to have communication outcomes with minor grammar errors.

Rule 15: Students who received training in the University Life topic and have a fluency proficiency level of 2 (struggle to speak and attempt to restart at times) tend to have communication outcomes with efforts to restart and some hesitations.

Rule 16: Students who received training in the Asking and Giving Directions topic and have a pronunciation proficiency level of 3 (occasional pronunciation errors) tend to have communication outcomes with pronunciation difficulties.

Rule 17: Students who received training in the Asking and Giving Directions topic and have a grammar proficiency level of 2 (basic grammar errors) tend to have communication outcomes with minor grammar errors.

Rule 18: Students who received training in the Asking and Giving Directions topic and have a fluency proficiency level of 3 (speak conversationally and fluently with occasional hesitations) tend to have natural and continuous communication outcomes with occasional hesitations.

7. Conclusions

Upon analyzing the performance before and after using the English conversation training program for communication purposes, it was discovered that the group of students enrolled in the English for Communication course during the first semester of the academic year 2022 exhibited a statistically significant difference in average scores before and after the training. The t-test resulted in a test statistic of 15.871 with 49 degrees of freedom and a significance level (sig) of .000. Consequently, it can be concluded that the training in English conversation skills for communication led to a significant improvement in English language proficiency at a significance level of .05.

As for the guidelines for adapting the English conversation training program using the random forest technique in the future, the rankings from easiest to most challenging are as follows:

Ranking 1 Greetings: This conversation topic obtained the highest scores in pronunciation, grammar, and fluency.

Ranking 2 Feeling and Emotion: This conversation topic ranked second, with a slight decrease in pronunciation scores.

Ranking 3 Health: This conversation topic ranked third, with decreases in pronunciation, grammar, and fluency scores.

Ranking 4 Asking for and Giving Directions: This conversation topic ranked fourth, with a slight decrease in scores.

Ranking 5 Weather: This conversation topic ranked fifth, with a decrease in scores.

Ranking 6 University Life: This conversation topic ranked last.

These rankings provide guidance on the relative difficulty levels of the conversation topics, which can assist in adjusting the English conversation training program effectively.

8. Discuss the Findings

The English conversation training provided in the book "StartUp" by Beatty et al. (2020) is designed to improve communication skills in personal life, studying, and working. It is suitable for learners at various proficiency levels. The results of pre- and post-training performance clearly demonstrate the positive impact of using English conversation training materials on learners' skills. This aligns with the findings of Srithongkul (2022) research, which showed that blended multimedia e-books in English language courses significantly enhance learning skills at Rajabhat Thonburi University, Samut Prakan.

From studying English conversation training that impact students' test scores using the Random Forest Technique, 18 related rules were derived. Rules 1, 3, 6, and 12, which are related at a high level (Level 4), involve the use of the "Greeting-Basic" exercise to promote pronunciation and fluency, and "Feeling and Emotion" and "Health" exercises to promote fluency. When examining the details of all three exercises, they share common characteristics with sentences used in everyday life. They contain information about health, and the dialogues are short. Furthermore, the structure and vocabulary used are at levels A1-A2 according to the Common European Framework of Reference for Languages (CEFR). This passage indicates that the study aligns with the study conducted by Pomin and Srinonyang (2020) found that the management of English language teaching in Thailand emphasizes practical language usage for daily life, aligning with a natural language learning approach. The teaching approach has shifted from a focus on grammar to an emphasis on communication skills, beginning with listening, speaking, reading, and writing in that order. As a result, students have achieved higher levels of language proficiency based on the CEFR. For the Grammar aspect, it was found that there is a proficiency level of "adequate" (Level 3) based on Rules 1, 3, 6, and 12. These rules are practiced in Greeting-Basic and Feeling and Emotion exercises. The development of English language grammar proficiency has improved from an adequate level to a good level through standardized CEFR skill training.

Additionally, analyzing the content of conversational interactions is essential for developing English speaking skills for everyday communication. This supports the ideas of Sacks, Schegloff, and Jefferson (Supakorn, 2020), who emphasize the importance of conversation analysis for understanding language usage and improving communication skills. Based on the research findings, the conversational topics were ranked according to their average scores. Greetings ranked first due to its resemblance to everyday phrases and the relatively easy structure and vocabulary of the exercises. Feeling and Emotion ranked second as students were familiar with the short dialogues from previous learning experiences. Health followed, incorporating specialized vocabulary while maintaining familiarity. Asking for and giving directions came next, focusing on providing and seeking directions, which students had limited exposure to in the lessons. Weather ranked second to last due to its longer dialogues and more specialized vocabulary. Finally, University Life received the lowest scores because of the lengthy dialogues and the variety of question-answer structures involved. It is recommended that instructors arrange the sequence of topics according to the learners' needs. According to conversational analysis theory, the organization of text sequences plays a crucial role in the effective development of learners' English speaking skills (Klanrit, 2019). Byrne (1997) also noted that teaching speaking skills requires practice in using stable components of language such as pronunciation, grammar patterns, and vocabulary, as well as opportunities for individual expression. Instructors should focus on accuracy in the initial stages and emphasize flexibility at higher levels of learning. This approach will enhance the efficiency of English communication skills in the long run.

9. Suggestion

Suggested recommendations based on the research findings are as follows:

1. The research should be replicated with a sample group studying English for Communication at Rajamangala University of Technology Tawan-ok to gather diverse research results.

2. The insights derived from the 18 correlated rules should be utilized in designing research activities aimed at addressing learning challenges in English subjects that share similar content, ensuring practical and tangible outcomes.

3. It is advisable to develop additional activities that offer students ongoing opportunities to apply English for communication in authentic real-life situations.

Acknowledgements

This research project was supported by the funding from Rajamangala University of Technology Tawan-ok, Bangphra Campus.

References

Beatty, K., Longshaw, R. N., & Austin, H. M. (2020). *StartUp*. Hoboken: Pearson Education.

Byrne, D. (1997). *Teaching oral English* (7th ed.). Harlow: Longman.

Carroll, B. J. (1980). *Testing communicative performance*. London: Pergamon.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). New Jersey: Lawrence Erlbaum Associates.

Denisko, D., & Hoffman, M. M. (2018). Classification and interaction in random forests. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(8), 1690-1692. doi:10.1073/pnas.1800256115

Dorfman, R. (1979). A formula for the Gini coefficient. *The Review of Economics and Statistics*, *61*, 146-149. doi:10.2307/1924845

EF English Proficiency Index. (2022). *Is English enough? But the ranking results can't tell! Thai English skills rank at the bottom of the rankings.4 Asia*. Retrieved from https://www.smartsme.co.th/content/248794

Ellis, M., & Johnson, C. (1994). *Teaching Business English*. Oxford: Oxford University Press.

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate data analysis* (8th ed.). Hampshire: Cengage Learning.

Hengsuko, C., Boonkao, K., Srihanam, N., & Manee, W. (2019). Current situation and factors affected English language communication skill of Thai people. *MBU Education Journal: Faculty of Education Mahamakut Buddhist University*, *7*(1), 155-160.

Huynh-Cam, T. T., Chen, L. S., & Le, H. (2021). Using decision trees and random forest algorithms to predict and determine factors contributing to first-year university students' learning performance. *Algorithms*, *14*, 318. doi:10.3390/a14110318

Jayaprakash, S., Krishnan, S., & Jaiganesh, V. (2020). Predicting students academic performance using an improved random forest classifier. In *2020 International Conference on Emerging Smart Computing and Informatics (ESCI)* (pp. 238-243). Pune, India. doi:10.1109/ESCI48226.2020.9167547

Kaewpanitch, C. (2020). *Using data mining techniques to analyze most characteristics that affect the performance of undergraduate students learning English* (Master's thesis, National Institute of Development Administration, Bangkok). Retrieved from https://as.nida.ac.th/wp-content/uploads/ 2022/08/6210426022 _ชิดชนก.pdf

Kanjanawasee, S., Pitayanon, T., & Srisukho, D. (2008). *Selection of appropriate statistics for research* (5th ed.). Bangkok: Chulalongkorn University Printing House.

Klanrit, P. (2019). Conversation analysis: An approach to teaching English speaking skill. *Faculty of Humanities and Social Sciences Review*, *7*(1), 33-51.

Madaan, M., Kumar, A., Keshri, C., Jain, R., & Nagrath, P. (2021). Loan default prediction using decision trees and random forest: A comparative study. *IOP Conference Series: Materials Science and Engineering*, *1022*, 012042. doi:10.1088/1757-899X/1022/1/012042

Meteevorakij, P., T.Chaisuwan, P., Nubpetchploy, S., & Tanwong, K. (2020). Performance measurement forecasting equations ability to study English for business communication based on the 21st century learning skills by creating decision tree. *Journal of Educational Measurement*, *37*(102), 202-218.

Pomin, T., & Srinonyang, P. (2020). Applying Web-based instruction to EFL for undergraduate students to improve the English language ability based on the reference levels of the Common European Framework of Reference for Languages. *Journal of Modern Learning Development*, *5*(5), 212-223.

Richards, J., & Rodgers, T. (2002). *Approaches and methods in language teaching* (2nd ed.). New York: Cambridge University Press.

Rujichom, K. (2018). English communication life skills of Thailand people in the 21th century. *Journal of Arts Management*, *2*(3), 199-210.

Ruksuan, P. (2022). Development of English innovation and guidebook based on communicative approach for working preparation about hospitality industry concerning hotel and accommodation services in ASEAN Community for Maehongson Community College students. *The Academic Journal: Faculty of Humanities and Social Sciences Nakhonsawan Rajabhat University*, *9*(1), 210-221.

Srithongkul, W. (2022). The development of Multimedia Electronic Book in English for Study Skills Subject of students at Dhonburi Rajabhat University Samutprakarn. *Sisaket Rajabhat University Journal*, *16*(1), 89-100.

Stansfield, C. (1973). Review of foreign language testing: Theory and practice, by John L. D. Clark. *Hispania*, *56*(3), 746-747. doi:10.2307/339288

Supakorn, S. (2020). A conversation analytic study of classroom interactional competence. *LEARN Journal: Language Education and Acquisition Research Network*, *13*(2), 15-40.

Tangpattanakit, J. (2020). Influence of factors affecting English communication with motivation as a central variable and anxiety as a controlling variable. *RMUTL Journal of Business Administration and Liberal Arts*, *8*(2), 111-126.

Thaweerat, P. (1997). *Social science research* (7th ed.). Bangkok: Charoenphon Printing.

The Office of Academic Support and Registration Rajamangala University of Technology Tawan-ok. (2005). *English for Communication Academic Year 2022*. Retrieved from http://regis.rmutto.ac.th/registrar/home.asp

Ur, P. (1996). *A course in language teaching: Practice and theory.* Cambridge: Cambridge University Press.

Wiengnil, A. (2010). *A study of out-of-class English language learning activities of secondary school students in English Program* (Master's thesis, Chulalongkorn University, Bangkok). Retrieved from http://cuir.car.chula.ac.th/bitstream/123456789/30421/1/atitaya_wi.pdf

Yates, J. (2016). *Practice makes perfect: English conversation* (2nd ed.). Tennessee: McGraw-Hill Education.