

## GENDER AND SCORES IN TOEFL-LIKE INFERENCE-TYPE LISTENING COMPREHENSION TESTS

Dewi Mustika Ningrum<sup>1</sup>, Lalu Jaswadi Putera<sup>2</sup>, Arafiq<sup>3</sup>, Amrullah<sup>4</sup>

University of Mataram

dewirukia.17@gmail.com

### Abstract

*This present study aims to identify, describe, and analyze the students' gender and scores in answering inference-type questions of the TOEFL-like listening comprehension tests. It sought to find answers to two research questions: (1) What is the students' ability in answering inference-type questions of the TOEFL-like listening comprehension tests based on gender; and (2) What are the percentages in answering the four types of inference questions i.e. functions, idiomatic language, contrary meaning, almost negative in the short dialogues part of the TOEFL-like listening comprehension test. Data were gathered from the TOEFL-like listening comprehension test results of 115 6th semester students of English Education Program. They were then analyzed using quantitative ex post facto method. 20 male and 20 female students were randomly chosen as the research samples which were grouped based on each gender. Data of the inference-type questions were taken from the short dialogues part of the Listening Comprehension section test. 5 inference-type questions were found in the test. The inference questions included 2 items about function of suggestion (Phillips, 2001:57), 1 item about idiomatic expressions (Phillips, 2001:70), 1 item about almost negative (Phillips, 2001:49), and 1 item about contrary meanings (Phillips, 2001:63). The results show that the mean score range of the male group 34.0 and the female group 36.0. Using Arikunto's classification, the mean scores are converted into "Poor" category (30.6–50.5 points) which means both male and female groups share the same level of ability in answering inference-type questions of the TOEFL-like listening comprehension tests. In regards to the percentages of students answering the types of inference questions, both male and female gender groups scored poorly in all of them. In functions of suggestion type, only 32% female and 27.5% male students could answer this type of question, meaning that 67% of female and 73% of male students failed. In idiomatic language type, only 45% of male and 40% of female students managed to answer question of this type, meaning that 55%-60% of them did also fail. In almost negative type, only 40% of female and 35% of male students succeeded, meaning that between 60%-65% of them did not succeed. In contrary meaning type, only 35% of both male and female students managed to score, meaning that 65% of them failed. This indicates that there is no gender gap or insignificant gap in the students' ability in answering the inference-type questions of the TOEFL-like listening comprehension test based on gender. However, the results picture a worrying trend among university students since 53%-73% of male and female students show poor performance in answering inference questions in the TOEFL-like listening comprehension test. Thus, this study suggests that students regardless of their gender develop the ability and strategy in answering inference-type questions of the TOEFL-like listening comprehension test be it through self-directed or teacher-directed learnings.*

**Keywords:** TOEFL, MBKM, Listening Comprehension; Inference Questions; Gender

## INTRODUCTION

Indonesia's latest curriculum for higher education commonly known as the *Merdeka Belajar-Kampus Merdeka* (MBKM) policy, equivalent to "freedom to learn-freedom campus", requires students to possess the 21<sup>st</sup> century learning skills known as the 4Cs (critical thinking, creative, collaborative, and communicative) (Putera, 2020). In teaching English as a foreign language context, being communicative is of paramount importance since having such ability guarantees the students' success when giving and receiving different kinds of information in communication (Remache, 2016) as well as to convey or share ideas and feelings effectively without misinterpretation or misleading others (Halawah, 2005; Suter *et al.*, 2009). Communicative abilities are related with having both good receptive skills and productive skills. Receptive skills are for listening and reading while productive skills are for speaking and writing. They are fundamental competencies in English language learning.

Among the four competencies needed to have good communicative skills, listening is one of the most crucial ones that students should develop since it is the most often used language ability in everyday life. Several studies showed that humans spend more than 40% of our everyday communication listening, 35% speaking, and just 9% writing (Miller, 2003 in Walker, 2014; Graham *et al.*, 2015; Gilakjani, 2016). Rost (in Hien, 2015) also stated that listening as foreign language learning is of a great importance since it provides the language input that plays crucial role in students' language communicative development especially for higher education students.

As an essential component of spoken communication, listening takes part in the interactive process in which an individual adopts the roles of speaker and listener through a verbal and nonverbal component. People must hear different types of English often and constantly in order to speak appropriately, meaningfully, and spontaneously. It indicates that humans cannot learn a language without hearing since listening offers language input.

The ability to develop or not develop during the process of learning a foreign language is influenced by, one of them, the gender factor. Several other factors such as physical factors, psychological factors, experience factors, attitude factors, motivation factors, environmental factors, and one's role in society factors also play role (Tarigan, 2015). As one of the influential factors in the development of listening ability in learning a foreign language, male and female genders in terms of their physical differences can be a

significant source of variation among language learners. Gender gives influence to the listening ability because the listening styles between male and female learners tend to be unique, in which men are more objective, active, hardhearted, analytic, rational, stubborn, dominant, neutral, instructive, independent, easier to recognize their needs, emotionally well-controlled compared to women who tend to be more subjective, passive, sympathetic, diffusive, sensitive, easy to influence, easy to succumb, receptive, dependent, and emotional. For this reason, many researchers suggested that gender must be considered seriously by the language teacher since it has an impact, to some extent, on language learning development (Yazdani & Ghafar, 2010 as cited in Abusaeedi *et al*, 2015; Nasri *et al*, 2018).

Listening is one of skills used in the material or section of TOEFL tests. TOEFL (Test of English as a Foreign Language) is a test used to assess the English proficiency of non-native English speakers. TOEFL is a standardized test which is designed to measure the English language ability of non-native speakers to enroll in Universities of America, as stated by Sharpe (2013). According to the Educational Testing Service (2013), also known as ETS, TOEFL scores are accepted by about 9000 institutions and colleges in over 130 countries. Many universities consider that this exam, which consists of academic questions and assignments, is acceptable for admissions choices. As a result, students must take the TOEFL in order to assess their level of English understanding. Prior to the thesis examination, students from all faculties are expected to take the TOEFL before registering their thesis examination. Each faculty and department establishes its own minimal score for students except for English Education department that requires its students to have higher score than other departments' i.e. 500.

Studies on gender performance among university students and their ability differences have become the main focus of several studies, not only related with TOEFL-like listening comprehension but also other competences such as in TOEFL-like reading comprehension (Destiyanti *et al*, 2021). Many university students had difficulties in dealing with the inference-type questions (Alavinia & Sameei, 2012) that have caused failures to both genders. The difficulties are caused by the students' lack of pragmatic competence. In other words, students have problems to comprehend the "implicit" answers in the dialog and make a conclusion from the inferential expressions (Mahmud, 2014; Arifuddin, 2015). Some solutions to the difficulties were highlighted by Arifuddin *et al*. (2014). They

proposed that students must improve their pragmatic competence in order to perform well in the TOEFL-like listening comprehension tests.

In line with the background of the study above, this study aims to analyze and observe the student's gender and scores in answering inference-type questions of TOEFL-like Listening Comprehension test as well as to find out the percentages of students who correctly answered the types of inference questions in the TOEFL-like Listening Comprehension tests based on gender.

## **RESEARCH METHODS**

### **Research Design**

This study employed quantitative ex post facto research design (Giuffre, 1997 as cited in Edmonds *et al*, 2016; Peterson, 2018). It is defined by Kerlinger (1964 in Kerlinger *et al*, 2000) as a research in which the independent variable(s) has already occurred and in which the researcher begins with the observation of a dependent variable(s) in retrospect for their possible relations to, and effect on, the dependent variable(s). This type of study frequently makes use of previously obtained data (Simon & Goes, 2013). This *ex-post-facto* research design was chosen to save time and money.

### **Population, Samples, and Sampling Technique**

The population and samples of this present study were the 6<sup>th</sup> semester students of the English Education Program of the University of Mataram that were selected using purposive sampling technique. The population was all 115 students/ testtakers who had previously taken the TOEFL-like test. From the overall 115 students, 40 students were chosen as the research samples using purposive sampling technique. They were then grouped based on gender that included 20 males and 20 females.

### **Data Collection**

The current research data were collected from three data sources using ex post facto method. The three data source were: (1) the students' scores of the TOEFL-like Listening Comprehension test recapitulated in Excel format, (2) paper-based TOEFL-like test kits for Listening Comprehension section, and (3) TOEFL-like Listening Comprehension test audio file. All of the three data sources were collected with permission

from the test administrator. The researcher team came to the administration office in person and handed the letter of permission to gather the necessary data.

## **Research Instruments**

The data of this *ex post facto* research were taken from the test result and additional documents needed. Ahmad (2019) defines a test as a series or exercise designed to examine the skill, knowledge, attitude, intelligence, ability, or talent of individuals or groups. Sugiyono (2015) defines documentation as documents authored and photographed by someone and utilized to gain information. This current research employed the test results of the TOEFL-like listening comprehension administered prior to the data collection.

### **1. Test**

The test from which this *ex post facto* data were taken from is the TOEFL-like test administered on December 14, 2019 at Gedung Kuliah Bersama building dome. The test were in three sections: Listening Comprehension, Structure and Written Expressions, and Reading Comprehension. The Listening Comprehension section consists of 50 question items. 5 out of 50 items were questioning about inferential questions i.e. questions numbers 5, 7, 12, 22, and 28. The 5 inference questions comprises 25% of the total of Listening Comprehension questions in the test. The types of inferential questions were asking about Function of Suggestion, Idiomatic Language, Contrary Meaning, and Almost Negative.

### **2. Documents**

Additional documents were taken from the students' score recapitulation of the TOEFL-like Listening Comprehension test in Excel format and the TOEFL-like Listening Comprehension test kits (paper-based question sheets, answer keys, and test audio file).

## **Data Analysis**

In conducting the data analysis, the researchers followed the qualitative data analysis methods i.e. data collection, data reduction, data display, and conclusion drawing. In data collection, researchers documented all of the TOEFL-like Listening Comprehension test kits and score results needed to answer the research questions in the current study.

In data reduction, we selected the inference-type questions in the short dialogues part of the TOEFL-like Listening Comprehension test and found 5 items that questioned about inference questions, classified them into 4 types of inference questions i.e. function of suggestion (2 items), idiomatic language (1 item), contrary meaning (1 item), and almost negative (1 item), grouped the students/testtakers based on gender (20 males and 20 females), counted the frequency of correct/incorrect answers in each group by gender and types of inference questions, summed up the scores of each group, calculated the mean scores of each group.

To display the data, we made bar chart diagrams to simplify the students' level of ability and the students' scores in answering the type of inference questions of the TOEFL-like Listening Comprehension test based on gender.

In conclusion drawing phase, we categorized the level of abilities of each group using Arikunto's 7-level of classification i.e. "Excelent" (90.6-100), "Very good" (80.6-90.5), "Good" (70.6-80.5), "Fairly good" (60.6-70.5), "Fairly" (50.6-60.5), "Poor" (30.6-50.5), "Very poor" (0-30.5).

## **FINDINGS AND DISCUSSION**

Upon employing data analysis, this study documented some interesting findings related to the proposed research questions and gender issues in answering inference questions of the TOEFL-like Listening Comprehension test. Each of the finding is elaborated in the following sections.

### ***1. Inference-type Questions in the Short Dialogues Part of the TOEFL-like Listening Comprehension Test***

Table below shows each of the five question items together with its type of inference found in the TOEFL-like Listening Comprehension test.

**Table 1.** Items of Inference Questions in the TOEFL-like Listening Comprehension Test.

No.	Question	Type of Inferences
1	Man : "I just got my third parking ticket this week" Woman : "Why don't you try putting more money in the parking meter when you park your car" Narrator: "What does the woman suggest that the man do?"	Functions: Suggestion
2	Woman : "Professor Mitchell's lecture certainly went on and on for quite some time." Man : "I thought he was never going to finish." Narrator: "What does the man imply about the lecture?"	Idiomatic Language
3	Man : "I need for you to work on these new accounting reports." Woman : "But I scarcely have time to finish the ones I already have." Narrator: "What does the woman imply?"	Almost Negative
4	Man : "I need to buy some stamps." Woman : "Then you'd better get to the post office quickly because it closes at 5 o'clock." Narrator: "What can be inferred from the conversation?"	Functions: Suggestion
5	Man : "How do you think you did on the literature exam that you had this morning." Woman : "I really wish I could take it over again." Narrator: "What does the woman imply?"	Contrary Meaning

Table above shows that Questions 1 and 4 are asking about functions of suggestion, Question 2 idiomatic language, Question 3 almost negative, and Question about contrary meaning. Functions of suggestion consist of 2 items, idiomatic language 1 item, contrary meaning 1 item, and almost negative 1 item.

## ***2. What Is the Students' Level of Ability in Answering Inference-type Questions of the TOEFL-like Listening Comprehension Test Based on Gender?***

This section highlights some important findings related to the first research question proposed in the present study i.e. “What is the students’ level of ability in answering inference-type questions of the TOEFL-like Listening Comprehension test based on gender?”. The findings are shown in the following table.

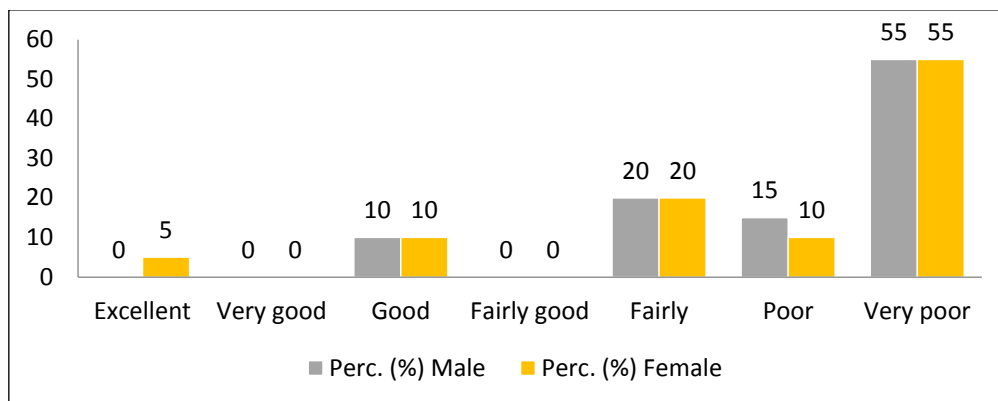
**Table 2.** Gender and Scores in Answering Inference-type Questions of the TOEFL-like Listening Comprehension Test Based on Arikunto’s 2013 Classification.

No.	Standard Score	Classification	Freq. (#)		Perc. (%)	
			Male	Female	Male	Female
1	80.6-90.5	Very good	0	0	0.0	0.0
2	60.6-70.5	Fairly good	0	0	0.0	0.0
3	90.6-100	Excellent	0	1	0.0	5.0
4	70.6-80.5	Good	2	2	10.0	10.0
5	30.6-50.5	Poor	3	2	15.0	10.0
6	50.6-60.5	Fairly	4	4	20.0	20.0
7	0-30.5	Very poor	11	11	55.0	55.0
Total			20	20	100.0	100.0

As we can see, the table above shows that:

1. Neither male nor female groups belong in “fairly good” and “very good” categories in the test. In other words, both groups scored zero (0.0).
2. 5% students in the female group achieved “excellent” score, beating their male counterparts. Despite that, the percentage is not so significant, 1 out of 20.
3. Smaller number of female students (10%) scored “poor” in the test compared to their male counterparts (15%) with a 5% gap in the same category.
4. Both male and female groups shared equal level of ability in the categories of “Fairly” (15% each), “Good” (10% each), and “Very poor” (55% each).
5. “Very poor” category, the very bottom category in the classification, records the highest percentage in both male and female groups with 55%, respectively.





**Figure 1.** Students' Level of Ability in Answering Inference-type Questions of the TOEFL-like Listening Comprehension Test.

To get a clearer picture of the students' level of ability in answering inference-type questions of the TOEFL-like Listening Comprehension test, see the pie chart on Figure 1.

Furthermore, data in Table 3 below show the mean, the highest, and the lowest scores of students' in answering inference-type questions of the TOEFL-like Listening Comprehension test.

**Table 3.** Gender and the Average, the Highest, and the Lowest Scores in Answering Inference-type

Questions of the TOEFL-like Listening Comprehension Test.

Score	Male	Female
Mean	34.0	36.0
	Poor	Poor
Highest	80.0	100.0
	Good	Excellent
Lowest	0.0	0.0
	Very poor	Very poor

The table shows that:

1. The mean scores in male and female groups are 34.0 and 36.0, respectively. Despite having a 2.0 score gap, both genders belong in equal level of ability i.e. "poor" category.

2. The highest score in answering inference-type questions of the TOEFL-like Listening Comprehension test is achieved by the female group with 100 points and 80.0 for the male group. The former belongs in “Excellent” category while the later belongs in “Good” category.
3. The students’ lowest score in answering inference-type questions of the TOEFL-like Listening Comprehension test is 0.0 (zero point zero) both in female group (20%) and male group (15%).

**3. What are the Percentages of Students’ Correct Answers on the Types of Inference Questions of the TOEFL-like Listening Comprehension Test Based on Gender?**

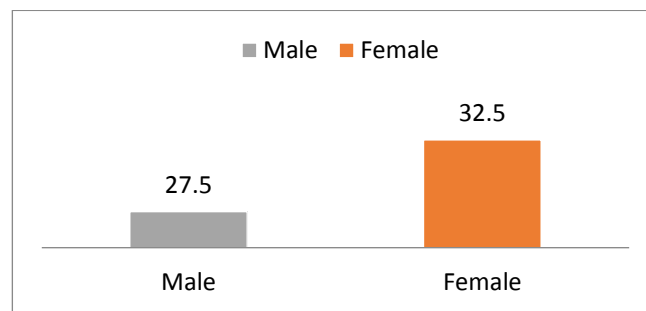
This section underlines some interesting findings related to the second research question in the current study i.e. “What are the percentages of students’ ability in answering inference-type questions of the TOEFL-like Listening Comprehension test based on gender?”. Table below shows the percentages of students’ correct answers on the types of inference questions in the TOEFL-like Listening Comprehension test based on gender.

**Table 4.** Percentages of Students’ Correct Answers on the Types of Inference Questions of the TOEFL-like Listening Comprehension Test based on Gender.

No	Type of Inference	Freq. (#)		Perc. (%)		Category	
		Male	Female	Male	Female	Male	Female
1	Functions: Suggestion	11	13	27.5	32.5	very poor	poor
2	Idiomatic language	9	8	45.0	40.0	poor	poor
3	Almost negative	7	8	35.0	40.0	poor	poor
4	Contrary meaning	7	7	35.0	35.0	poor	poor

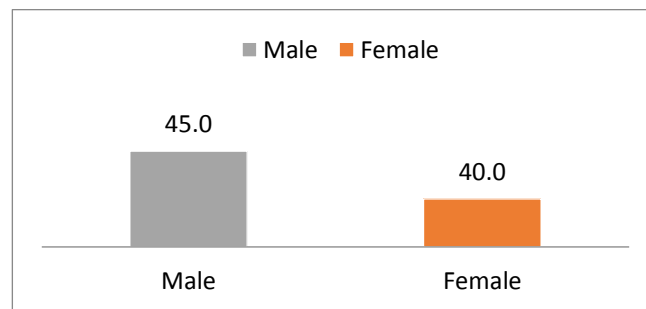
Data in Table 4 show that:

1. In answering **Inference Question Type 1: Function of Suggestion**, 32.5% of female and 27.5% of male students managed to answer this type of inference questions. It means that 68% of female and 73% of male students failed to answer this type of inference questions. Using Arikunto's classification, the percentages can be converted into "very poor" to "poor" category.



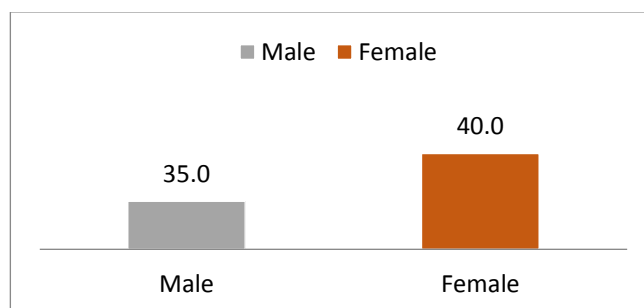
**Figure 2.** Inference Question Type 1: Function of Suggestion.

2. In answering **Inference Question Type 2: Idiomatic language/expressions**, 45% of male and 40% of female students managed to score. In other words, 55% of female and 60% of male students did not succeed to score well in the test.



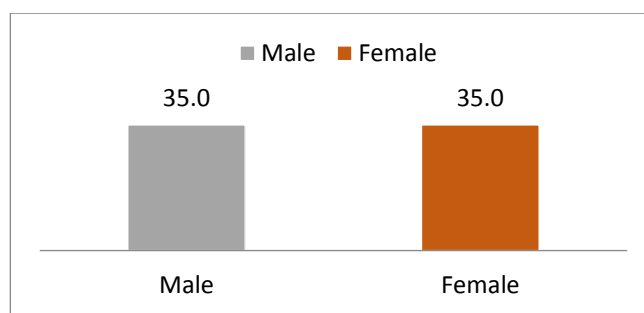
**Figure 3.** Inference Question Type 2: Idiomatic Language.

3. In answering **Inference Question Type 3: Almost Negative**, 35% of male and 40% of female students succeeded to answer correctly in the test. In other words, 60% of female to 65% of male students failed to score correctly in this type of inference questions.



**Figure 4.** Inference Question Type 3: Almost Negative.

4. In answering **Inference Question Type 4: Contrary Meaning**, both genders shared equal percentage with 35% each. This means that 65% of male and female students failed to score well in this type of inference question test.



**Figure 5.** Inference Question Type 4: Contrary Meaning.

Based on the presented percentages of correct answers in the four types of inference questions (function of suggestion, idiomatic language, almost negative, and contrary meanings) that range between 35%-65%, it can be categorized that both male and female students level of ability in answering all four types of inference questions is in “Very poor” to “Poor” category.

## CONCLUSION

As the title implies, this present study aims to analyze and describe the gender and scores in answering inference-type questions of the TOEFL-like Listening Comprehension test among English Education department students. It also aims to find out the percentages of students’ correct answers of the types of inference questions in the test. There are 4 types of inference questions found in the test that included (1) functions of suggestion, (2) idiomatic language, (3) almost negative, and (4) contrary meaning. Based on the data presented in the findings and discussion section, this study concludes that: (1) the

level of students' ability in answering inference-type questions of the TOEFL-like listening comprehension test based on gender groups are in the "Poor" category with the mean scores are 34.0 for male group and 36.0 for female group. Despite having a 2.0 mean score gap, both genders remain at the same level of ability i.e. "poor" category. The highest score in answering inference-type questions in the female group is 100 points ("Excellent" category) while in the male group is 80 points ("Good" category). 20% of female group did not score any points whatsoever while 15% of male group did not either. The percentages of students who correctly answered inference-type question are: (1) in function of suggestion, male group scored 27.5% while female group scored 32.5%; (2) in idiomatic language, male group scored 45%, female group scored 40%; (3) in almost negative, male group scored 35%, female group scored 40%; and (4) in contrary meaning, both male and female gender groups share the same amount of percentage that is 35%. The percentages of students' correctly answering to these four types of inference questions range between 27%-45% which means around 55%-73% of them failed to answer these types of inference questions. By using Arikunto's classification, it can be interpreted that the students' level of ability in answering inference questions containing function of suggestion expressions, idiomatic expressions, almost negative expressions, and contrary meaning expressions is in "poor" to "very poor" categories.

## REFERENCES

- Abusaeedi, R.A.A., & Boroomand, F. (2015). A quantitative analysis of Iranian EFL learners' sources of written errors. *International Journal of Research Studies in Language Learning*, 4(1), 31-42.
- Ahmad, T. (2019). *Metodologi penelitian praktis*. Yogyakarta: Teras.
- Alavinia, P. & Sameei, A. (2012). *Ikatan potensial antara ekstroversi/introversi dan kemampuan pemahaman mendengarkan pelajar EFL Iran*. *Pengajaran Bahasa Inggris*, 5(5), 19-30. <https://doi.org/10.5539/elt.v5n5p19>.
- Arifuddin, A. (2014). Ranking of causes of failure to infer implicature in TOEFL-like based on gender. *Theory and Practice in Language Studies*, 4(7), 1334.
- Arifuddin, A. (2015). Pragmatic-based listening: A Solution for gender-based Listening. *Pragmatic-based listening: a solution for boosting the English proficiency of the Indonesian learners of English*, 282-286.
- Arikunto, S. (2013). *Prosedur penelitian suatu pendekatan praktik*. Edisi revisi. Jakarta: Rineka Cipta.

- Destiyanti, C., Amin, M., Putera, L.J. (2021). Gender-Based Analysis of Students' Ability in Answering Factual and Vocabulary-In-Context Questions of the TOEFL-like Reading Comprehension Test. *PALAPA*, 9, no. 1, 1-17.
- Educational Testing Service (ETS). (2003). *TOEFL Institutional Testing Program (ITP) and TOEIC Institutional Program (IP): Two On-Site Testing Tools from ETS at a Glance*. Handout Berlin Conference. Princeton: Educational Testing Service.
- Gilakjani, A. P., & Sabouri, N. B. (2016). Learners' Listening Comprehension Difficulties in English Language Learning: A Literature Review. *English language teaching*, 9(6), 123-133.
- Giuffre, M. (1997). *Designing research: Ex post facto design*. *Journal of Peri Anesthesia Nursing*, 12(3), 191-195 : Published by Elsevier Inc. Cited in Edmonds, W. A., & Kennedy, T. D. (2016). *An applied guide to research designs: Quantitative, qualitative, and mixed methods*. Sage Publications.
- Graham, S., & Santos, D. (2015). *Strategies for second language listening: Current scenarios and improved pedagogy*. Springer.
- Halawah, I. (2005). The relationship between effective communication of high school principal and school climate. *Education*, 126(2), 334-346.
- Hien, T. V. (2015). Difficulties and strategies in listening comprehension. *Truong Dai Hoc Lac hong: Lac Hong university*.
- Kerlinger, F. N., Lee, H. B., & Bhanthumnavin, D. (2000). Foundations of behavioral research: The most sustainable popular textbook by Kerlinger & Lee (2000). *Journal of Social Development*, 13, 131-144.
- Mahmud, M. (2014). *The EFL Students' Problems in Answering the Test of English as a Foreign Language (TOEFL) A study in Indonesian Context*. Findland: Academy Publisher.
- Nasri, M., Biria, R., & Karimi, M. (2018). Projecting gender identity in argumentative written discourse. *International Journal of Applied Linguistics and English Literature*, 7(3), 201-205.
- Peterson, J. (2018). *Middle school students' perceptions of caring teacher behaviors* (Doctoral dissertation, Eastern Michigan University).
- Phillips, D. (2001). *Longman complete course for the TOEFL test: Preparation for the computer and paper tests*. London: Longman.
- Putera, L. J., & Sugianto, R. (2020). Perception and optimism about two-semester off-campus internship program of the kampus merdeka-merdeka belajar (freedom campus-freedom to learn) policy among university students. *Journal of Languages and Language Teaching*, 8(3), 264-275.
- Remache, A. (2016). Developing Students' Communicative Competence in University English Language Programs. *International Journal of Arts & Sciences*, 9(1), 183.
- Sharpe, P. J. (2005). *Barron's Practice Exercises for the TOEFL 5th Edition*. Jakarta: Bina Rupa Aksara.
- Simon, M.K. & Goes, J. (2013). *Dissertation and Scholarly Research: Recipe for Success*. Seattle, W.A: Dissertations Success LLC.
- Sugiyono. (2015). *Metode Penelitian Kuantitatif Kualitatif dan R&D*. Cetakan ke-21. Bandung: Alfabeta.

- Suter, E., Arndt, J., Arthur, N., Parboosingh, J., Taylor, E., & Deutschlander, S. (2009). Role understanding and effective communication as core competencies for collaborative practice. *Journal of interprofessional care*, 23(1), 41-51.
- Tarigan, H.G. (2015). *Menyimak Sebagai Suatu Keterampilan Berbahasa*. Bandung: Angkasa.
- Walker, N. (2014). Listening: The most difficult skill to teach. *Encuentro*, 23(1), 167-175.