

# Text-to-Speech Application for Foreign Language Learner' Listening Comprehension in Indonesia

1<sup>st</sup> Tika Nur Andriani<sup>1</sup>, 2<sup>nd</sup> Yayuk Widyastuti Herawati<sup>2</sup>, 3<sup>rd</sup> Teguh Sulistyio<sup>1</sup>  
{tikanurandriani@gmail.com<sup>1</sup>, yayukwherawati@uin-malang.ac.id<sup>2</sup>, sulistyoteguh@unikama.ac.id<sup>1</sup>}

University of Kanjuruhan Malang<sup>1</sup>, UIN Maulana Malik Ibrahim Malang<sup>2</sup>

**Abstract.** Listening takes role to provide the main aural input which is really important. The use of difference speech origins for the listening material will bring different effect for students' listening comprehension. This study identifies whether the students who are taught by synthetic speech have significant difference on their listening comprehension than those who are taught by natural non-native speaker. It used quasi experimental research design. The experimental group consisted of 34 students who treated by using synthetic speech audio material which produced by Text-To-Speech application. The control group consist of 32 students treated by non-native speaker. The instrument used in this research was test. The result of independent sample t-test from the post-test showed significant value of the data was 0.016. From the result above, it can be concluded that teaching listening using different speech origins has effect for students' listening comprehension. It was proved by the result of the test. It is suggested that English teacher can use synthetic speech audio material which produced by Text-To-Speech application for teaching listening to make the teaching and learning process more enjoyable.

**Keywords:** listening comprehension; natural speech; synthetic speech.

## 1 Introduction

In the process of learning English, there are four language skills that should be learnt and practiced by the learners namely listening, speaking, reading, and writing. Among those skills, listening is considered as the important skills that foreign language (FL) learners should be learnt. Listening as "the activity of paying attention to and trying to get meaning from something we hear"[1], [2]. Listening can be considered as the first step in learning a language [3], [4], [5], [6]. It means that language learning, initially, highly depends on listening skill. In listening to spoken language, the ability to decipher the speaker's intention is required of a competent listener. Meanwhile, listening comprehension is the ability to understand the spoken language of native speakers [7].

In language acquisition, listening takes role as the thing that provide the main aural input which is really important [5], [6]. The auditory input is different, it differs according to the speech origin of the speaker. Speech origin is any kind of auditory input that can recognize by the hearer. Whether the input is from natural speech or synthetic speech. When the aural input is from human who directly speak or record it as the listening material, it called natural speech or human-generated speech. The natural speech can be divided into native speaker and also

non-native speaker. Natural native speaker is people who speak English as their mother tongue or people who comes from English speaking countries, but non-native speaker is people who speak English as their second or foreign language or people who comes from non-English speaking countries. In the other hand, synthetic speech or computer-generated speech is a kind of listening material which is generated from the Text-To-Speech application that can directly transferred any computer text materials into audio files.

The use of difference speech origins for the listening material will bring different effect for students' listening comprehension. Natural speech requires little effort when listening to the speech signal [8]. When a person listens to natural speech he or she do not need to devote more effort to identification the phoneme because natural speech provides higher rate of phoneme identification than synthetic speech. Although some people say that listening to natural speech is easier then listening to synthetic speech [9], the use of text-to-speech to produce synthetic speech has increase during past decade and became more popular in many education settings [10].

One of advantage of using text-to-speech application to produce synthetic speech audio material is schools do not need to hire any readers to record students' listening materials. Teachers or people in the school just need to scanned students' book and change it into type-text and the text can easily be converted to speech [11], [12]. There is no human speaker who must read and record the text, it means the teacher or people in the school can save time, energy and money. Another advantage from using synthetic speech is teachers also can manipulate the sound effect, speaker, language, pause, pronunciation correction, volume, pitch, and the speed through the application. So, the teacher can adjust the listening material that has been created with the ability of the students [2], [4], [13], [14].

Some researchers have investigated about the use of synthetic speech. The first previous study investigated the implementation of text-to-speech media software in teaching listening for eight grades of junior high school. The finding showed that the students could catch the meaning from the audio which was produced by text-to-speech software easily and more enthusiastic with the audio file [15], [16].

The second previous study investigated comprehension of synthetic speech produced by rule: word monitoring and sentence-by-sentence listening times, the finding of this study was contrast to the previous study mention above. It shows that the on-line task performance was slower and less accurate for passages of synthetic speech than for passages of natural speech [17].

Other previous study, syntactic speech can help integrating character education and local genius through *balabolka* in teaching listening [18]. The possibility of a foreign language teacher to modify the materials based on learner's needs.

The last previous study, that investigated the comparison of collage students' ability to comprehend passage material when input (natural and synthetic speech) was provided in different modalities: L-only or listening to the text, R-only or reading the text silently, and RWL or simultaneously reading and listening to the text. The finding showed there was no difference in comprehension for natural versus synthetic speech in any conditions [8].

The finding from previous studies are various, most of them used various kind of text for their listening materials. It is important for the researcher to investigate whether the natural and synthetic speech give any impact for foreign language learners' listening comprehension especially at first grade high school students when they listen to narrative text as their listening material. This research is expected to reveal further information related to the effect of speech origin on foreign language learners' listening comprehension. It is also hoped that this research will also reveal which treatment is better for listening comprehension.

The objective of this study is to find out whether students who are taught by synthetic speech have significant difference on their listening comprehension than those who are taught by natural nonnative speaker. Theoretically the finding of the research is intended to add new insights and knowledge of theoretical understanding related to of speech origins on foreign language learners' listening comprehension. For the English teacher the finding of the research is expected to help teacher to select and use appropriate listening material for teaching and learning process. For further researcher the result of this research is expected can become some referential to conduct next research.

## **2 Methods**

This research used quasi-experimental research. There are two kind of variable for this research. They are dependent and independent variable, in this case the dependent variable is foreign language learners' listening comprehension and for independent variable is speech origin. For the population of this research the researcher used X grade students of SMA Diponegoro Tumpang in East Java, Indonesia while for the sample of this research the researcher took X IPS 4 and X Bahasa for the experimental group and the control group. The researcher used pre-test and post-test as the instrument of the research to collect the data. The instrument consisted of fifteen multiple choices questions. For calculate the data the researcher used SPSS 22 software.

## **3 Finding**

### **3.1 The Result of the Pre-test**

In this research, the primary data were taken from the test. The test was divided into pre-test and post-test [19], [20]. This finding showed significant difference between experimental and control group. The population of this research was first grade students at SMA Diponegoro Tumpang. There were two samples, they were X IPS 4 (treated by synthetic speech), X Bahasa (treated by natural non-native speaker). The total number of students of the sample was 66 students.

The researcher conducted research for both groups. The first step is the researcher gave pre-test to them to know whether both groups are homogeneous or not. The next step the researcher started to applied treatment which is using synthetic speech to teaching listening. But, this treatment was only given for the experimental group. For the control group, the researcher taught by conventional method like dictation by natural non-native speaker. The last step, the researcher gave post-test to both groups to know the significant difference on students listening comprehension between the group which was taught by using synthetic speech and the group which was taught by using natural non-native speaker.

There were three kind of activities; pre-test, treatments, and post-test. The researcher gave the pre-test at the first meeting, the researcher needed to calculate the normality and the homogeneity of the test before implement the treatment. The result normality and homogeneity of the test showed that the significant value in  $>0.05$  it meant that the test had normal distribution and the data was homogeneous before the researcher implement the treatment. After giving the pre-test the researcher gave the treatments at the second and third

meeting. For the treatment, the researcher gave treatment by using synthetic speech audio material for the experimental group and for the control group the researcher teach the students using conventional method or in this case the researcher deliver the listening material by directly dictate it. After implementing the treatment, the researcher gave post-test. The result of the data showed; the mean score of both groups was 77.65 for the experimental group and 73.91 for the control group. To prove whether the treatment was success or not, the researcher used independent sample t-test for analyzing the data. The result showed that the significant value from the independent sample t-test was 0.016. From the result above it was clear that the null hypothesis is rejected because the probability of sampling error is lower than the level set by the researcher ( $0.016 < 0.05$ ). In short, it is clear that the implementation of the treatment was successful.

The researcher gave the pre-test for experimental group and control group. The pre-test was conducted in the first meeting for both groups on May 8<sup>th</sup> and 9<sup>th</sup>, 2018. The pre-test for control group (X Bahasa) was done on May 8<sup>th</sup>, 2018 at 09.40 am until 11.00, while for experimental group (X IPS 4) was done on May 9<sup>th</sup>, 2018 at 09.40 am until 11.00 am. The researcher gave pre-test in the first meeting before researcher gave treatment for experimental group. The pre-test consisted of 15 items of multiple choices. The result of the pre-test will be showed in table 1.

**Table 1.** Pre-Test Score

Group Statistic						
Class	Class	N	Min	Max	Mean	Std. Deviation
	Experimental Group	34	30	70	50.88	9.085
	Control Group	32	30	70	48.59	9.266

The result of pre-test experimental group and control group showed that the number of the students in experimental group was 34. Mean score of experimental group (X IPS 4) was 50.88. The minimum score of experimental group was 30 and the maximum score was 70. Then the standard deviation was 9.085.

On the other hand, the number of the students in control group was 32. Mean score of control group (X Bahasa) was 48.59. The minimum score of control group was 30 while the maximum score was 70. Then the standard deviation was 9.266. The table showed both group experimental and control group had the same score between maximum and minimum score.

From the result, the mean score of experiment group was 50.88 and the mean of control group was 48.59. So, the researcher concluded that both groups had same performance and the average was almost the same before conducting the treatment.

### 3.2 Normality & Homogeneity of the Test

Normality test is for testing whether the data is normally distributed. In this research, to test the normality, the researcher also used SPSS 22 with Sapiro-Wilk method. It shows the significant value of the normality test is 0.512 for the experimental group and 0.624 in control group showing that the test has normal distribution.

The aim of homogeneity test is to know whether the respondents on population are equal in term or ability or not. The homogeneity of the respondents has been counted from the pre-test score. It is found that the test of homogeneity between groups is 0.983. The significant value of homogeneity is 0.05 level. It shows no significant differences between experimental and control groups in their listening comprehension, it can be concluded that the data were homogeneous.

### 3.3 Treatment for Experimental Group

The researcher had two meetings for giving treatment for the students on May 11<sup>th</sup> and 16<sup>th</sup>, 2018. The first treatment was held in the second meeting on May 11<sup>th</sup>, 2018 for the experimental group (X IPS 4). The students of the experimental group were given treatment using synthetic speech audio material for listening. The researcher was started teaching the students about the material which was suitable with the syllabus. The material was about narrative text.

Group Statistic						
Class	Class	N	Min	Max	Mean	Std. Deviation
	Experimental Group	34	30	70	50.88	9.085
	Control Group	32	30	70	48.59	9.266

The first, the researcher try to introduce the material by stimulate the students through some questions about what they know about the common legend story that familiar for them and asked about what narrative text is as the pre-listening activities. Almost all the students answered the question. Some students have learnt and know about what narrative text is. Some of them also mention about the example of legend story. However, there were some students who could not answer the questions and forget about what narrative text is. The researcher explained more detail about narrative text, the kinds of narrative text and the generic structures of narrative text.

Second, the researcher asked the students to listen to the synthetic speech audio material about the example of legend story and pay attention about it. The researcher played the audio three times, while the audio was played the researcher asked the students to find out the difficult word from the text at the audio for the whilst listening activities. At the end of the meeting as the post listening activity, the researcher asked the students to discuss the difficult words and find the meaning in Bahasa together.

The second treatment was held in the third meeting on May 16<sup>th</sup>, 2018 for the experimental group (X IPS 4). At the beginning of the lesson the researcher remaining the students about the narrative text by giving some question related with the previous lesson. Then, the researcher gave the students another example of legend story and asked the students to listen and pay attention to the synthetic speech audio file. Same as the previous lesson, the researcher played the converted audio three times.

Then, the researcher gave a small task to the students, the task was five multiple choices questions about the legend story which have played by the researcher. The aim of the task was for train the students to listen to the synthetic speech audio file and to know the students understanding toward the story that they have heard. At the end of the lesson the researcher asked the students to discuss about the question together and played the synthetic speech audio material once more to find out the best answer for each question.

### 3.4 Treatment for Control Group

In order to compare the experimental group, the students in control group was treated by using conventional method in their listening activity. In control group, the material that was taught was the same but the aural input that use for teaching listening was different. In conventional method, the students did not listen to the synthetic speech audio file for their listening materials but the students were asked to listen and pay attention to the researcher as the non-native speaker when the researcher read and doing dictation of the legend story. Before read the story, the researcher explain about what narrative text is, the kinds of narrative text, and the generic structures of the narrative text. There was no different for the material; the different was only in the aural input for the listening material that the researcher uses for teaching.

### 3.5 The Result of Post-test

The researcher gave post-test for experimental group and control group after implement the treatment, the researcher conducted the post-test in order to know the effect of speech origin for students listening comprehension that has been used in experimental and control group. Besides, the researcher also wanted to know the difference mean between the experimental and control group after giving the different treatments. The post test was conducted in the fourth meeting on May 18<sup>th</sup>, 2018 for both groups. The post-test for experiment group (X IPS 4) was done on May 18<sup>th</sup>, 2018 at 07.00 am until 08.20 am, while control group (X Bahasa) was done on May 21<sup>st</sup>, 2018 at 07.00 am until 08.20 am. The pre-test consisted of 15 items of multiple choices.

**Table 2.** Post-Test Scores

Group Statistics						
	Class	N	Min	Max	Mean	Std.
Class	Experimental	34	65	90	77.65	Devia tion 6.424 5.783
	Group	32	60	85	73.91	
	Control Group	32				
	Valid N (listwise)					

Based on table 2, the result of post-test for experimental group and control group showed that the number of the students in experimental group was 34. Mean score of experimental group (X IPS 4) was 77.65. The minimum score of experimental group was 65 and maximum score was 90. Then the standard deviation was 6.424.

On the other hand, the number of the student in control group was 32. Mean score of control group (X Bahasa) was 73.91. The minimum score of control group was 60 while the maximum score was 85. Then the standard deviation was 5.783.

From the result, we knew that the result of the post-test between experimental group and control group was different. The mean score of experimental group was 77.65 which higher than the mean of control group was 73.91.

**Table 3.** Independent Sample T-Test

	t-test for Equality of Means				
	Sig. (2-tailed)	Mean Diff.	Std. Error Diff.	95%Confidence Interval of the Diff.	
				Lower	Upper
Listening_ Score	.016	3.741	1.508	.729	6.753
	.015	3.741	1.503	.738	6.743

The qualification of hypothesis is that if the significant value is less than the significant level (0.05). It meant that there is a significant difference between the two groups (experimental group and control group). If the significant value is more than the significant level (0.05). It meant that there is no significant difference between the two groups (experimental group and control group). The conclusion is if there is significant difference, the null hypothesis (Ho) is rejected and alternative hypothesis (Ha) is accepted. While, if there is no significant difference, the null hypothesis (Ho) is not rejected and alternative hypothesis (Ha) is rejected.

Based on the result of the data above, it showed significant value of the data was 0.016. From the result above it was clear that the null hypothesis is rejected because the probability of the sampling error is lower than the significant level set by the researcher ( $0.016 < 0.05$ ).

Based on the descriptive statistic, the mean of the experiment group was higher than the mean score of control group ( $77.65 > 73.91$ ). So, descriptively the experiment group got higher mean score than control group. The researcher concluded that students who were taught by synthetic speech had better comprehension in listening than students who were taught by natural non-native speaker. So, speech origins had significant effect on foreign language learners' listening comprehension at first-grade students of SMA Diponegoro Tumpang Kabupaten Malang, EastJava, Indonesia.

### 3.6 The Analysis of Test Validity & Reliability

The researcher established content validity of the listening test in this research the test items were constructed based on the basic course outline of the 2013 curriculum for tenth grade students. To get the content validity, the researcher matched the material with the curriculum, as it was developed based on the blueprint of the test and for measure the blueprint of the test matched with the curriculum the researcher used expert validity.

The researcher calculated the data using Kuder Richardson – 21 (KR-21) reliability by Microsoft Excel on Windows. The result of reliability was important to prove whether the instrument was good enough to give for the students or not. It is found that the mean of the data was 12.18, the standard deviation from the data was 2.69 and the variance of the data was 7.25. From the data above the researcher then calculate the data manually using the formula of KR-21. The result of is 0.757. It meant that the value is between 0.70–0.79. Based on the criteria of coefficient of reliability by Kuder Richardson, the data of test has fair reliability. In short, the test items were given to the students were reliable.

### 3.7 Research Discussion

The effect of speech origins on foreign language learners' listening comprehension can be seen from the result of the analysis of the mean score on students' test. The result of the testing revealed that the experimental group, which was given treatment by using synthetic

speech was significantly have better listening comprehension with the mean score of the experimental group was 77.65 and the mean score of the control group was 73.91.

Based on the finding of the research, it is clear that using synthetic speech as the audio file material for teaching listening can help the students at SMA Diponegoro Tumpang being more interested and pay more attention to the listening material. So, they can get the message from the audio file easily. This point related to previous study which were investigated the implementation of text-to-speech media software in teaching listening for eight grades of junior high school have found that the audio file which was produced by text-to-speech software could build a good atmosphere in the classroom. The students could catch the meaning from the audio easily and more enthusiastic with the audio file [7].

However, the finding of this research was different with the finding of the studies which investigated about the comprehension of synthetic speech produces by rule: word monitoring and sentence-by-sentence listening times. The finding of previous study showed that the on-line task performance was slower and less accurate for passages of synthetic speech than for passage of natural speech [8]. It was different with present research because this research used simple and short text for the listening material so it helped students to understand the message from the text easily.

The finding of this research was also different with the finding of previous study who investigated the comparison of collage students' ability to comprehend passage material when the auditory input (natural and synthetic speech) was provided in different modalities: L-only or listening to the text, R-only or reading the text silently, and RWL or simultaneously reading and listening to the text [4]. The finding showed there was no difference in comprehension for natural versus synthetic speech in any conditions. It was different with present research because Taake used college students as the object of the research while the present research used high school students as the object of the research. High school students pay more attention to the synthetic speech audio material because they think that this kind of listening material is interesting for them.

#### **4 Conclusion and Suggestion**

The conclusion of this research is drawn in accordance with the result of the data analysis in the previous chapter. There are some significant differences between experimental group and control group. In the result of the post-test of experimental group was 77.65 which was higher than the control group 73.91. It means that teaching listening using synthetic speech audio material was better than teaching listening using conventional method by natural non-native speaker.

The result of calculation using the t-test showed that significant value of the data was 0.016, from the result it was clear that the null hypothesis is rejected because the probability of the sampling error is lower than the significant level set by the researcher ( $0.016 < 0.05$ ). It can be concluded that there is a significant difference between students who were taught by synthetic speech and the students who were taught by natural non-native speaker in their listening comprehension at SMA Diponegoro Tumpang Kabupaten Malang. From the explanation above, it can be concluded that teaching listening using different speech origins has effect for students' listening comprehension at first-grades students of SMA Diponegoro Tumpang Kabupaten Malang.



The researcher gives suggestion to further researcher to develop this research by applying the application on other kinds of text, such as report, procedure, or recount text. Then, the researcher hopes that the research can be perfected in the content and result by further researchers which are related to the effect speech origins on foreign language learners' listening comprehension.

**Acknowledgements.** This paper in conjunction with the 1st International Conference on Islam, Science and Technology (ICONIST) 2018

## References

- [1] Underwood, M.. *Teaching Listening*. New York: Longman Inc. 1990
- [2] Brown, H. Douglas. *Teaching by principles: an interactive approach to language pedagogy* Second Edition. White Plains: Addison Wesley Longman, Inc. 2001.
- [3] Margareth, J. *The Teaching Listening and Speaking*. New York: The Macmillan Company. 1988.
- [4] Davies, Paul. *Success in English teaching*. Oxford: Oxford University Press. 2000.
- [5] Krashen, S. D. *Principles and practice in second language acquisition*. California: Pergamon Press Inc. 1966.
- [6] Tricomi E.T. "Krashen's Second Language Acquisition Theory and the Teaching of Edited American English" *Journal of Basic Writing* Vol 5 No 2 1986
- [7] Mendelsohn, D.. *Learning to listen: A Strategy-Based Approach for The Second Language Learner*. San Diego: Dominic Press. 1994
- [8] Taake, K. A Comparison of Natural and Synthetic Speech: With and Without Simultaneous Reading. Saint Louis, Missouri: All Theses and Dissertations (ETDs). 2009.
- [9] Roring, R. W., Hines, F. G., & Charness, N. "Age difference in identifying words in synthetic speech". *Human Factors*, vol 49, 2007, pp. 25–31.
- [10] Koul, R. 2003. Synthetic speech perception in individuals with and without disabilities. Texas: Augmentation and Alternative Communication, 19: 49–58.
- [11] Harmer, J. 2007. How to Teach English. England: Longman.
- [12] Richard, W. A & Renandya, A. W. *Methodology in Language Teaching: An Anthology of Current Practise*. New York: Cambridge University Press 2002.
- [13] Richards, Jack C. *Curriculum development in language teaching*. Cambridge: Cambridge University Press. 2001.
- [14] Sadtono, E. The development of TEFL in Indonesia. Malang: Penerbit IKIP Negeri Malang 1997.
- [15] Sitepu, R., Munir, A., and Suwono. The Implementation Of Text-T O-Speech Media Software In Teaching Listening For Eighth Graders Of Junior High School. Surabaya: *International Conference on Education and Science*. 2017.
- [16] Wenden, A. dan Rubin, J. *Learner strategies in language learning*. Englewood Cliffs, New Jersey: Prentice Hall International 1987.
- [17] Ralston., J., Pisoni, D., Lively, S., Greene, Beth., and Mullennix, J. "Comprehension of synthetic speech produced by rule: word monitoring and sentence-by-sentence listening times". Bloomington, Indiana. *Human Factors*, vol. 33 no.4, 1991, pp. 471 – 491.
- [18] Herawati, Yayuk W. Integrating character education and local genius through Balabolka in teaching listening. *Lingua Journal*. vol 11, No 1, 2016

- [19] Ary, D et al *Introduction to Research in Education*. Eight Edition. Canada: Nelson Education. 2010.
- [20] Creswell. *Educational research: planning, conducting, and evaluating quantitative and qualitative research*. Boston: Pearson. 2012.