CHAPTER 4
FINDING AND DISCUSSIONS

This chapter consisted of three parts, they were research finding, hypothesis testing and discussion. The findings of the research presented the description of the data collected through pre-test, test, and post-test. Moreover, discussion consisted of the result of the research.

4.1. Finding of the Research

The data of this study was quantitative data. Quantitative data was taken by pre-test, test, and post-test. Pre-test conducted in first meeting by the researcher. Those tests were given to both experimental and controlled group. The result of pre-test provides information about both of groups' ability in speaking. The post-test was administered then to both of groups' after the experimental group got the treatments and the control group taught through conventional teaching.

Table 4.1
Schedule of the research

No		Experimental Class	Control Class
1	Test validation	Friday, 05 February 2021	
2	Pre-test	Wednesday, 10 February 2021	Thursday, 11 February 2021
3	Treatment using time token arend methods	Wednesday, 17 February 2021 Wednesday, 24 February 2021	
4	Treatment using pair and conventional method		Thursday, 18 February 2021 Thursday, 25 February 2021
5	Post-test	Wednesday, 03 March 2021	Thursday, 04 March 2021

Based on Sugiyono (2018:193), validity means that the instrument can be used to measure what should be measured. Validity is used to know the index validity of test. In this study, the researcher used content and construct validity, and the writer ensured face validity by consulting to English teacher of SMA Negeri 1 Tahunan.

Table 4.2
Result of validation test

No	Category of validity	Indicator	Score				Comment
			1	2	3	4	
1	Content validity	1) The material related to syllabus of tenth grade students in second semester.				√	
		2) The material relates to basic competencies (KD) of tenth grade students in second semester.				√	
		3) The content of the material and topic are appropriate with grade and school level.				√	
		4) The picture is appropriate with the material.				√	
		5) The material is appropriate with the students' target skill.			√		
2	Construct validity	1) The test instrument has covered the aspects of speaking				√	

		test.				√	
		2) The test instructions are easy for students to understand.				√	
		3) The instructions are accordance with the aspects to be measured.			√		
		4) Time allocation of the test is quite effective.			√		
		5) The assessment rubric has covered all aspects and indicator of the test.			√		

Score = $\frac{total\ score}{40}$ $x\ 100 = \frac{36}{40}$ $x\ 100 = 90$

Based on the table above, the result of the test validation by using construct and content validity showed the score is 90 and have a very good category. It could be concluded that the instrument of this research was valid and could to measure the knowledge of the students and could use to pre-test and post-test of students.

4.1.1. The Result of Pre-test

In this study, the experimental group was the students of X MIPA 4. The following table shows the result of pre-test for experimental class and control class.

Table 4.3

Score of Pre-test

No	Experiment Class		Control Class	
	Code	Score	Code	Score
1	A-1	60	C-1	56
2	A-2	56	C-2	64
3	A-3	68	C-3	72
4	A-4	72	C-4	52
5	A-5	64	C-5	80
6	A-6	56	C-6	64
7	A-7	64	C-7	80
8	A-8	72	C-8	76
9	A-9	76	C-9	64
10	A-10	68	C-10	60
11	A-11	60	C-11	64
12	A-12	64	C-12	52
13	A-13	68	C-13	60
14	A-14	60	C-14	64
15	A-15	56	C-15	68
16	A-16	64	C-16	72
17	A-17	68	C-17	60
18	A-18	60	C-18	64
19	A-19	44	C-19	76
20	A-20	56	C-20	64
21	A-21	52	C-21	72
22	A-22	64	C-22	80
23	A-23	80	C-23	72
24	A-24	60	C-24	68
25	A-25	64	C-25	72
26	A-26	56	C-26	80
27	A-27	76	C-27	76
28	A-28	60	C-28	64
29	A-29	56	C-29	60
30	A-30	80	C-30	48
31	A-31	76	C-31	68
32	A-32	72	C-32	60

33	A-33	56	C-33	84
34	A-34	60	C-34	52
35	A-35	64	C-35	56
	∑	2234	∑	2324
	Mean	63,77	Mean	66,40

Based on the table above, the mean scores of pre-test in experimental group was 63.77, meanwhile the mean of pre-test in control group was 66.40. The biggest score of the experimental class was 80 and the lowest score of experimental class was 44, while the biggest score of control class was 84 and the lowest score was 48. It can be seen that both of class have the difference result. From the result, it can be concluded that the students' mastery of speaking in experimental class is lower than control class, but there were same level.

4.1.2. The Result of Post-test

The following table shows the result of post-test between experimental class and control class.

Table 4.4

Score of Post-test

No	Experiment Class		Control Class	
	Code	Score	Code	Score
1	A-1	76	C-1	72
2	A-2	80	C-2	76
3	A-3	84	C-3	80
4	A-4	80	C-4	68
5	A-5	84	C-5	92
6	A-6	72	C-6	72
7	A-7	76	C-7	68

8	A-8	88	C-8	64
9	A-9	84	C-9	84
10	A-10	72	C-10	68
11	A-11	80	C-11	80
12	A-12	68	C-12	72
13	A-13	80	C-13	76
14	A-14	76	C-14	80
15	A-15	72	C-15	84
16	A-16	76	C-16	68
17	A-17	84	C-17	80
18	A-18	76	C-18	72
19	A-19	68	C-19	76
20	A-20	76	C-20	80
21	A-21	80	C-21	84
22	A-22	84	C-22	76
23	A-23	96	C-23	72
24	A-24	84	C-24	88
25	A-25	80	C-25	76
26	A-26	76	C-26	72
27	A-27	88	C-27	76
28	A-28	76	C-28	68
29	A-29	80	C-29	76
30	A-30	92	C-30	80
31	A-31	80	C-31	68
32	A-32	84	C-32	72
33	A-33	76	C-33	92
34	A-34	80	C-34	64
35	A-35	84	C-35	68
	∑	2792	∑	2644
	Mean	79,77	Mean	75,54

Based on the table above, the mean score of post-test in experimental

class and control class was 79,77 and 75,54. Meanwhile the biggest score

for experimental class was 96 and control class was 92. The lowest score

of experimental class was 68 and for control class was 64. From the result, we can be concluded that the level of speaking mastery both of class was higher than the result of pre-test by using time token arend method for experimental class and paired discussions for control class.

4.1.3. Normality and Homogeneity Testing

The preliminary analysis consisted of normality and homogeneity test. It was intended to know whether the data was distributed normally and homogeneous or not. In this research, the writer used IBM SPSS Statistic 22 for analyzing the normality and homogeneity test. The result can be seen as follow:

4.1.3.1. Normality of the Pre-Test

Table 4.5
Normality of Pre-Test

| | Kelas | Kolmogorov-Smirnov[a] | | | Shapiro-Wilk | |
		Statistic	df	Sig.	Statistic	Df
Hasil Speaking Siswa	Experimental Pre Test	.146	35	.057	.956	35
	Control Pre Test	.145	35	.060	.965	35

| | Kelas | Shapiro-Wilk[a] |
		Sig.
Hasil Speaking Siswa	Experimental Pre Test	.179
	Control Pre Test	.322

a. Lilliefors Significance Correction

The normality test used to know whether both groups had normal distribution data or not. The data was normal when the significant value is bigger than 0.05 (p>0.05), when the data <0.05 the distribution data was not normal. The normality test above used *Kolmogrov-Smirnov* critical points table in the determining the t_{table} because the sample was 35 students. The table above showed the significance of experimental class was 0.057 and for controlled class was 0.060. The result shown that the significance of both classes were 0.05. Therefore, the distribution of pre-test score was normal.

4.1.3.2. Normality of the Post-Test

Table 4.6
Normality of Post-Test

	Kelas	Kolmogorov-Smirnov[a]			Shapiro-Wilk	
		Statistic	Df	Sig.	Statistic	Df
Hasil Speaking Siswa	Experimental Post Test	.142	35	.070	.952	35
	Control Post Test	.145	35	.061	.942	35

Tests of Normality

	Kelas	Shapiro-Wilk[a]
		Sig.
Hasil Speaking Siswa	Experimental Post Test	.128
	Control Post Test	.063

a. Lilliefors Significance Correction

The normality test used to know whether both groups had normal distribution data or not. The data was normal when the significant value is bigger than 0.05 (p>0.05), when the data <0.05 the distribution data was not normal. The normality test above used *Kolmogrov-Smirnov* critical points table in the determining the t_{table} because the sample was 35 students. The table above showed the significance of experimental class was 0.070 and for controlled class was 0.061. The result shown that the significance of both classes were 0.05. Therefore, the distribution of pre-test score was normal.

4.1.3.3. Homogeneity of the Pre-Test

The homogeneity was use to know the data are homogeneous or not (Sudjana, 250:2003). The homogeneity test used to show two or more groups of sample data from the populations whom have same variations. Homogeneity test also used to determine whether several variations the data of populations were same or not. The value of significance (α) = 0.05. The writer used *Levene* statistic in SPSS 22 to analyze the homogeneity.

Table 4.7
Homogeneity of Pre-Test

Hasil Speaking Siswa

Levene Statistic	df1	df2	Sig.
1.010	1	68	.318

Based on the result above by using Levene's test, it was known that the significance value of variances score was 0.318 it means that the value

was higher than 0.05. It assumed that both of the variances were homogeneous (0.318>0.05).

4.1.3.4. Homogeneity of the Post-Test

Table 4.8
Homogeneity of Post-Test

Hasil Speaking Siswa

Levene Statistic	df1	df2	Sig.
1.327	1	68	.253

Based on the result above by using calculation Levene's test, it was known that the significance value of the variances scores was 0.253 and it means that the value was higher than 0.05, therefor the data of post-test was homogeneous.

4.1.3.5. Independent T-Test of the Pre-Test

The independent sample T-Test used to compare the means for two different populations (experimental and control class). It was need this test to see the variance that the both groups were equal. The hypothesis for the t-test can be seen below:

a. Ho: both the variances were the same or equal (experimental and control group).

b. Ha: both of the variances were different and not equal (experimental and control group).

This research used standard significance (α) = 0.05 to test the hypothesis. The interpretations to test the hypothesis were state below:

a. If sig.(2-tailed) > 0,05 = there is no significant

b. If sig.(2-tailed) < 0,05 = there is significant

To analyze the data, the writer used SPSS 22 version. The result can be seen below:

Table 4.9
Group Statistic

	Kelas	N	Mean	Std. Deviation	Std. Error Mean
Hasil Speaking Siswa	Experimental Pre Test	35	63.7714	8.22866	1.39090
	Control Pre Test	35	66.4000	9.26473	1.56603

Based on the table 4.9 (group statistic) above, it showed there were two class, they were experimental class and control class. Then, at N it showed the number of students in experimental and control class. In experimental class there were 35 students with the mean score was 63.77 the standard deviation was 8.22866 and standard error mean was 1.39090. While, in control class there were 35 students with the mean score was 66.40 and standard deviation was 9.26473 and standard error mean was 1.56603.

From the result of mean score both experimental and control class was the mean of experimental class was lower than control class.

Table 4.10

Independent T-Test of Pre-Test

		Levene's Test for Equality of Variances		t-test for Equality of Means
		F	Sig.	T
Hasil Speaking Siswa	Equal variances assumed	1.010	.318	-1.255
	Equal variances not assumed			-1.255

		t-test for Equality of Means		
		Df	Sig. (2-tailed)	Mean Difference
Hasil Speaking Siswa	Equal variances assumed	68	.214	-2.62857
	Equal variances not assumed	67.066	.214	-2.62857

		t-test for Equality of Means		
		Std. Error Difference	95% Confidence Interval of the Difference	
			Lower	Upper
Hasil Speaking Siswa	Equal variances assumed	2.09452	-6.80813	1.55099
	Equal variances not assumed	2.09452	-6.80919	1.55204

In this research, the writer using F test (Levene's test) to know the result of homogeneity testing. From the result, it would be seen whether to use

"equal variances assumed (if variance is same) or "use equal variances not assumed" if variances is different in *t-test*. Because the data was homogeneous, the writer used independent sample t-test.

Based on the table above, it can be known that the result of *t-test* for pre-test score. It can be seen that the significant p value test was 0.214. In other word, p>0.05, and it's bigger than 0.05. It means that the null hypothesis is not rejected and there was no statistically significant difference between the two groups. Hence, the condition of experimental and control group in the beginning was no difference.

4.1.3.6. Independent T-Test of the Post-Test

Table 4.11
Group Statistic

	Kelas	N	Mean	Std. Deviation	Std. Error Mean
Hasil Speaking Siswa	Experimental Post Test	35	79.7714	6.13134	1.03638
	Control Post Test	35	75.5429	7.24505	1.22464

In addition to the t-test analysis of pre-test score, the researcher also analyzed the result of post-test score. It was intended to know the condition of the classes after giving the treatment both of class. However, the results both of the classes were increasing. The experimental class had mean score improvement from 63.77 in pre-test become 79.77 in post-test, meanwhile the control class had mean score improvement from 66.40 become 75.54.

Then, at N it showed the number of students. In experimental class there were 35 students with the standard deviation was 6.13134 and standard error mean was 1.03638. Meanwhile, in control class there were 35 students with the standard deviation was 7.24505 and standard error mean was 1.22464. It was seen from the mean score both of class, the experimental class was higher than the control class.

Table 4.12
Independent T-Test of Post-Test

		Levene's Test for Equality of Variances		t-test for Equality of Means
		F	Sig.	T
Hasil Speaking Siswa	Equal variances assumed	1.327	.253	2.636
	Equal variances not assumed			2.636

		t-test for Equality of Means		
		Df	Sig. (2-tailed)	Mean Difference
Hasil Speaking Siswa	Equal variances assumed	68	.010	4.22857
	Equal variances not assumed	66.190	.010	4.22857

		t-test for Equality of Means		
		Std. Error Difference	95% Confidence Interval of the Difference	
			Lower	Upper
Hasil Speaking Siswa	Equal variances assumed	1.60432	1.02721	7.42993
	Equal variances not assumed	1.60432	1.02562	7.43152

In this result of post-test score in the experimental and control class, as presented in the table above shows that the significant value of the t (2-tailed) was 0.010. It was showed that sig. (2-tailed) was lower than 0.05 (0.010 < 0.05) and there was significant between experimental and control class in post-test. Therefore, that the alternative hypothesis (H_a) is accepted and null hypothesis (H_0) was rejected. In other words, it can be concluded that there is significant difference score between students taught by using Time Token Arends strategy and those taught without using Time Token Arends Mmethod in speaking ability of SMA Negeri 1 Tahunan Jepara.

4.2. Hypothesis Testing

This research was intended to investigate the effectiveness of using Time Token Arends method on students' speaking skill at tenth grade of SMA Negeri 1 Tahunan Jepara. In other, to get the result of this research, the researcher used hypothesis testing. The steps for testing the hypothesis can be described as follows:

1. Alternative hypothesis (H_a): there was significance effect of using Time Token Arends method on students' speaking skill.

2. Null hypothesis (H_0): there was no significance effect between using of Time Token Arends strategy on students' speaking ability.

To prove the hypothesis, the result of t-test calculation was tested with these situations:

1. If Sig. (2-tailed) > 0.05 in significance degree of 5%, the Null Hypothesis (H_0) was accepted and the alternative hypothesis (H_a) was rejected. It means that there was no significant effect of using Time Token Arends Method on students' speaking ability.

2. If Sig. (2-tailed) < 0.05 in significance degree of 5%, the Null Hypothesis (H_0) was rejected and alternative hypothesis (H_a) was accepted. It means that there was significant effect of using Time Token Arends Method on students' speaking ability.

Based on the result of t-test, there was difference significant between experimental and control group. According to the result by using SPSS showed that the Sig. (2-tailed) < 0.05 (0.010 < 0.05). In other words, it meant that null hypothesis was rejected and there was statistically difference significant between the two classes. Hence, the H_0 was rejected and H_a was accepted. Thus, using of Cooperative Learning with Time Token Arends Method to improve students' speaking ability for tenth grade was effective.

4.3. Discussion of the Research

This study is about the use of Time Token Arends method to improve students' speaking ability of tenth grade at SMA Negeri 1 Tahunan. This section is intended to analyse the result of research finding based on the related theory. All data collected from the instrument provides the information of the result and after that all the data was calculate using t-test.

This research was conducted of four meetings for each class. In the first meeting, the pre-test was administered in both of the experimental and control class. The aim of pre-test was to know the students' speaking ability before giving the treatment. After that, the second and third meeting was conduct of treatment. Meanwhile, experimental class was teaching using Time Token Arend method and control class was taught by using conventional method. In the last meeting, the students were given a post-test after they got the treatment, the purposed of post-test was to measure the effectiveness of the method after getting the treatment. The result of pre-test and post-test bellow:

Table 4.13
Students mean score

No	Class	Mean pre-test score	Mean post-test score
1	Experimental class	63.77	79.77
2	Control class	66.40	75.54

Based on the table above, it could be concluded that the mean of pre-test in experimental class was 63.77 and in control class was 66.40. The mean of control class was higher than experimental class. And the mean score of experimental group in post-test was 79.77 and control class was 75.54. From the result on post-test, the score of experimental class was higher than control class after taught by using Time Token Arends method than using conventional method. So, there was significant difference between the students' score in speaking ability without using the method and not at tenth grade of SMA Negeri 1 Tahunan Jepara.

Table 4.14
Statistical analysis in pre-test and post-test

T-test Sig. (2-tailed)	
Pre-test	Post-test
$0.214 > 0.05$	$0.010 < 0.05$
H_0 was accepted	H_0 was rejected
H_a was rejected	H_a was accepted

From the T-test explanation, the result of statistical analysis was used as the data for testing hypothesis and to get the final result. According to the data analysis, it showed that the significant (2-tailed) was bigger than the coefficient significant 0.05 ($0.214 > 0.05$). It means that H_0 was accepted and H_a was rejected (there was no significant difference in students' speaking ability between the experimental and control group).

Then, the result of t-test statistical analysis showed that there was a significant difference in the students' speaking ability of experimental and control group after they got a treatment. The result of post-test was increasing to moderate level. It concluded that the H_0 was rejected and H_a was accepted.

The previous researcher also had proved that Time Token Arends strategy could be effective and increase the students' speaking ability. It was supported by some previous studies done related to the implementation of Time Token Arends in teaching speaking. The first, research conducted by Siti Mardiah (2017) the research found out that Time Token Arends technique could increase students speaking ability with four cycles.

The second was conducted by Mira Nurjannah and Ahmad Hambali (2018) the objective of this research was to find out whether Time Token Arends method could improve students' speaking ability. The result of this research showed that Time Token Arends method can improve students' speaking ability, especially in encouraging passive students become more active students.

The next research conducted by Raudhatul Jannah (2019). It was about the implementation of Time Token Arends Model in improving students' speaking achievement at Islamic Senior High Schools MAN 3 Banda

Aceh. The finding of this research was to show that using Time Token Arends could increase students' speaking ability.

The next previous research was conducted by Sholihatul Hamidah Daulay, Maryati Salmiah (2018). The researcher got the best result after conduct two cycles which each cycle consisted of planning, acting, observing, and reflecting. The mean of post-test 1 was still low, and the second post-test was better than the first cycle. It indicated that the students' skill in speaking improved and became well in the first meeting to the next meeting.

The last was conducted by Nur Faizah (2015) it was about the implementation of Time Token Arends to improve students' speaking skill. The result of the research showed that there were some improvements in students speaking in teaching narrative text and the teacher suggested using Time Token Arends for teaching speaking and another aspect of English.

From the explanation above, it can be concluded that Time Token Arends method could become the effective strategy for teaching speaking. In this study, the researcher focused on the use of Time Token Arends to improve students' speaking ability. Arends in Slavin (2001:15) said that time token one of type in cooperative learning which can be used to teach the social skill, to avoid the students who dominate the discussion or student kept quiet at all.

With this strategy, students are required to be able to cooperate with each other and discuss with other to express their ideas. This learning process made students more enthusiastic during learning. The implementation of Time Token Arends in this research showed that the students involved in discussion well.

From the explanation above, it can be concluded that Time Token Arends method could the appropriate strategy for teaching speaking at Senior High School. It can increased students' speaking ability and it can be conclude that this research using Time Token Arends was effective to teach speaking students at tenth grade of SMA Negeri 1 Tahunan Jepara.