

CHAPTER IV

RESEARCH FINDINGS

This chapter presents the findings of the study. It is divided into two main parts. The first part concerns analysis of variance to test hypotheses one, two, and three for research questions one, two and three. The second part presents qualitative analysis of interview data to answer the fourth research question.

SECTION ONE: FINDINGS FROM THE MIXED FACTORIAL ANOVA

I. The comparison of the test format effects on the test takers' listening comprehension scores

Hypothesis 1: There is a significant difference in the average scores obtained from the multiple choice format and the short answer format.

$$(H_1: \mu_{MC} \neq \mu_{SA}, \alpha = 0.05)$$

To test Hypothesis 1, a mixed within-between ANOVA was performed to examine the differences between the scores. When an ANOVA with a repeated measures factor is conducted, SPSS automatically conducts a test of sphericity - the Mauchly's test. Sphericity requires that the variances of each set of different scores were equal. If the Mauchly's test statistic was less than .05, it meant that the assumption of sphericity was violated. However, in the present study the significance level was greater than .05 so sphericity can be assumed. The summary tables of the repeated measures effects in the ANOVA with sphericity assumed values, corrected F-values and the effect size value are below.

Table 4.1
Mauchly's Test of Sphericity

Within Subjects Effect	Mauchly's W	Approx. Chi-Square	df	Sig.	Epsilon		
					Greenhouse-Geisser	Huynh-Feldt	Lower-bound
format	1.000	.000	0	.	1.000	1.000	1.000

Table 4.2
Test of Within-Subjects Effects

Source		Type III Sum of Squares	df	Mean Square	F	Eta Squared	Cohen's <i>d</i>	Observed Power
Format	Sphericity Assumed	152.510	1	152.510	25.621*	.119	.735	.999
Error	Sphericity Assumed	1139.990	190	5.953				

* $p \leq .05$ (two-tailed)

Table 4.3
Means for Test Format Effects

Format	Mean	SD
1 (Multiple-Choice)	11.182	3.651
2 (Short Answer)	12.443	3.777

Table 4.2 illustrates that there was a significant difference between the subject main effect of the test formats, $F(1, 190) = 25.62, p < .001$, with an effect size of $\eta^2 = 0.119$. This suggests that the test-takers' listening ability was affected by the test formats used. From Table 4.3, the format 1, which is the multiple choice format, was more difficult than the short answer format for the test takers.

The effect size measurement was also performed to analyze the size of the experimental effect of the test formats. The Eta Squared was 0.119, which means that the factor test format by itself accounted for 11.9 % of the overall (effect + error) variance. Besides the Eta Squared, the magnitude of effect size value was also analyzed using the suitable formula for F test – the Cohen's d . A Microsoft Excel Spreadsheet (Thalheimer and Cook, 2002) was used to compute the value. From the excel calculation, the Cohen's d was reported to be 0.735 which is interpreted as having a medium effect size (Cohen, 1992). The two values of effect size that were $\eta^2 = 0.119$ and the Cohen's $d = 0.735$ confirm that the main effect of the test formats has a medium effect size. The relationship between Cohen's d and η^2 is described in Appendix F.

Further, the Cohen's d effect size can be thought of as the average percentile standing. An ES of 0.735 indicates that the mean of the short answer scores is at 76th percentile of the multiple choice scores. This also can be interpreted in terms of the amount of nonoverlap between the two groups. An ES of 0.735 indicates a nonoverlap of approximately 43% (see Appendix F). This means that the average score of the persons taking the short answer format exceeded the average score of those taking the multiple choice format approximately 43 % (Becker, 2000).

Consequently, Hypothesis 1 is accepted. The scores from the multiple choice format were significantly different from the short answer format scores at the 0.05 level. The Cohen's d effect size value of 0.735, which is close to a large effect size level, indicates that employing different formats in the listening comprehension test had quite a strong effect on the participants' scores.

II. The comparison of the English accent varieties effects on the test takers' listening comprehension scores

Hypothesis 2: There is a significant difference in the average scores obtained from the test using English native speakers' model as listening stimuli and the test using other accent varieties of English as listening stimuli.

$$(H_1: \mu_{NS} \neq \mu_{NNS}, \alpha = 0.05)$$

A mixed within-between ANOVA was employed to test the second hypothesis. Since the main effect of the test version was a between subject variable, the result was listed separately from the repeated measure effect for Hypothesis 1. Before employing an ANOVA with a between subject variable, the test of homogeneity of variances or the Levene test and the test of normality were conducted first to test whether the scores of test versions A and B were normally distributed and the variances of the two groups were indifferent. Table 4.4 and Table 4.5 are the results of the test.

Table 4.4
Test of Homogeneity of Variances

		Levene Statistic	df1	df2
Total Score	Based on Mean	3.488	1	190
	Based on Median	3.591	1	190
	Based on Median and with adjusted df	3.591	1	185.904
	Based on trimmed mean	3.513	1	190

Table 4.5
Tests of Normality

		Kolmogorov-Smirnov		Shapiro-Wilk	
	Test Version	Statistic	df	Statistic	df
Total Score	1	.058	96	.981	96
	2	.072	96	.985	96

From Table 4.5, all significance values were more than 0.05. This means that the scores from the two test versions were distributed indifferently. Furthermore, both tests of normality that are Kolmogorov-Smirnov and Shapiro-Wilk indicate that the scores of the two groups were also distributed in a normal manner. This made the use of a between-subject ANOVA possible because the assumption was not violated. Below is the result of a between-subject ANOVA employed to test the difference in scores of the two test versions.

Table 4.6
The Main Effects of Test Versions

Source	Type III Sum of Squares	df	Mean Square	F	Eta Squared	Cohen's <i>d</i>	Observed Power
Intercept	53581.500	1	53581.500	2534.199*	.930	7.300	1.000
Test Version	86.260	1	86.260	4.080*	.021	0.292	.520
Error	4017.240	190	21.143				

* $p \leq .05$ (two-tailed)

Table 4.7
Means for Test Version Effects

Test Version	N	Mean	Std. Deviation	Minimum	Maximum
A (NS)	96	24.5729	7.08407	9.00	38.00
B (NNS)	96	22.6771	5.86424	11.00	37.00

The result shown in Table 4.6 reveals a significant effect. There was a significant main effect of test version, $F(1, 190) = 4.080, p < 0.05$, with an effect size of $\eta^2 = .021$. From Table 4.7, the test version A which used native speakers' voice was easier than the test version B that used nonnative speakers' voice for the test takers in this present study. An Eta squared was just 0.021 and this means that the test version (accent varieties of English) factor by itself accounted for only 2.1% of the overall variance.

The magnitude of effect size value was analyzed using the suitable formula for F test – the Cohen's d . Microsoft Excel Spreadsheet (Thalheimer and Cook, 2002) was used to compute the value again. From the excel calculation, the Cohen's d was reported to be 0.292 which is interpreted as having a small effect size. The two values of effect size that are η^2 of 0.021 and the Cohen's d of 0.292 confirm that the main test version effect was small. Moreover, an effect size of 0.292 (Cohen's d) means that the average score of the persons in the test with native speaker voice stimuli exceeded the average score of those in the nonnative speaker voice group approximately 21% (Becker, 2000). As for the percentile standing, an ES of 0.29 indicates that the mean of the native speaker test version scores is at about 62 percentile of the nonnative speaker test version scores (see Appendix F).

Although the scores of both test versions were reported to be significantly different, the difference produced a very small effect size of $d = 0.29$. The following table might be the explanation of the little effect size.

Table 4.8
High Ability Test Takers' Scores

Sample Code	Mechanical Matching Scores*	Test B Scores**
0087	29	28
161	30	35
024	30	34
065	31	28
110	31	26
112	33	34
153	33	32
166	34	30
Mean = 31.375, SD = 1.767		Mean = 30.875, SD = 3.356

*Mechanical matching scores from the final Listening Comprehension examination in February 2006

**Test B scores from the nonnative speaker test version in March 2006

From Table 4.8, the high ability test takers, who scored more than 28 in their final examination test for mechanical matching, could also score high in the test version B, the nonnative speaker version. Some of them scored higher and some of them scored lower but they were still the same people who scored high in both tests. However, in order to be certain that the scores from their final examination and test version B are not different; the paired-samples *t*-test was conducted. The results are reported as follows:

Table 4.9
Paired Samples Test for High Ability Test Takers' Scores

Scores	Paired Differences			t	df	Sig.(2-tailed)
	Mean	Std. Deviation	Std. Error Mean			
Final exam and Test B	.500	3.625	1.281	.390	7	.708

From Table 4.9, the results from the *t*-test (t value = 0.390, $p = .708$) indicates that the scores from the two tests were not significantly different. This means that the ability of high test takers was similar in the final examination and the test version B. It implies that although unfamiliar accent varieties of nonnative speakers were reported to be more difficult to understand for the whole group of 96 subjects, the

nonnative accents did not obstruct comprehensibility of the message for the advanced test takers. If all the subjects were advanced test takers, the result might have been different – the ANOVA might have reported an insignificant effect of the test version. This might be an explanation for the small effect size of accent varieties of English variable.

Therefore, Hypothesis 2, which states the scores from the two test versions that use native speakers' voice and nonnative speakers' voice as test stimuli are different, was accepted. The effect size reveals that although there was a significant difference between the scores from the two test versions, the size was quite small. The Eta squared was just .021, which means that the test version factor by itself accounted for only 2% of the overall variance.

III. The comparison of the interaction effect between the test format variable and the English accent varieties variable

Hypothesis 3: There is a significant interaction effect between the test formats and the accent varieties of English on average listening comprehension scores.

($H_1: \mu_{MC NS} \neq \mu_{MC NNS} \neq \mu_{SA NS} \neq \mu_{SA NNS}$ at least one pair, $\alpha = 0.05$)

A two-way ANOVA also reports the interaction between the two main variables – the test formats and accent varieties of English. The advantage of factorial design over conducting multiple one-way ANOVAs lies precisely in this capacity of ANOVA to look at the interaction effect of the combination of variables. To test the interaction effect between the test formats and versions in this present study, a mixed within-between subjects 2 X 2 ANOVA was used to test Hypothesis 3. Table 4.10 reports the interaction effect of the combination of these two variables.

Table 4.10
The Interaction Effect of Test Format and Version

Source		Type III Sum of Squares	df	Mean Square	F	Eta Squared	Cohen's <i>d</i>	Observed Power
Format	Sphericity	37.500	1	37.500	6.300*	.032	.363	.704
*Version	Assumed							
Error	Sphericity	1130.990	190	5.953				
	Assumed							

* $p \leq 0.05$ (two-tailed)

There was a significant interaction between the test format and the test version, $F(1, 190) = 6.30, p < 0.05$, with an effect size of $\eta^2 = 0.032$. The effect size of Cohen's d was calculated with the Microsoft Excel Spreadsheet (Thalheimer and Cook, 2002) to investigate the magnitude of the effect. It was reported to be 0.363 which is in the level of small effect.

The interaction is significant. This means while the multiple choice format is more difficult than the short answer format, the difficulty condition might increase or decrease due to the combination of the test version factor. There are many possible patterns involved in interaction. The best way to interpret an interaction is to plot the means of the groups. This interaction can be shown graphically as in Figure 4.1 and Figure 4.2.

Figure 4.1
The Interaction Effect of Test Format and Versions

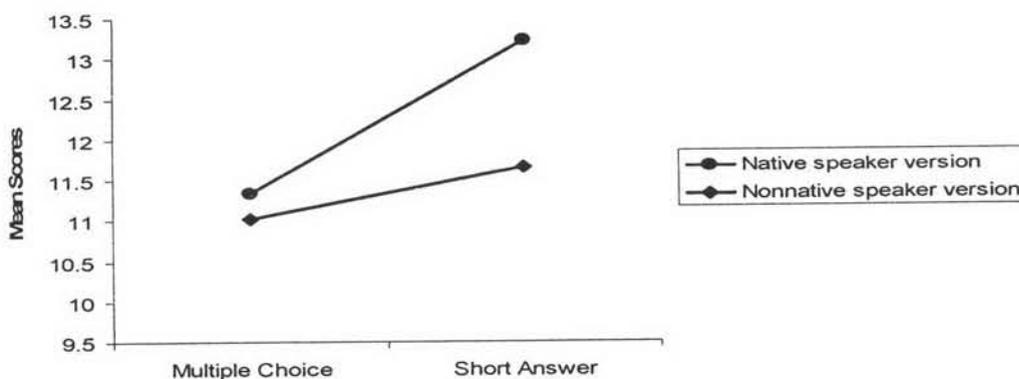
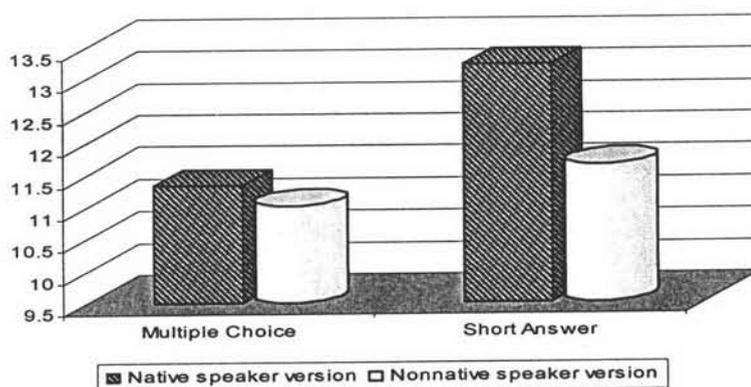


Figure 4.2
Means of Multiple Choice and Short Answer Format
In the Two Test Versions



Consider Figure 4.1, this pattern of interaction indicates that there was virtually not much difference between the scores in test versions A and B for the multiple choice format (version A mean =11.343, SD = 3.996; version B mean = 11.020, SD = 3.283). However, using the short answer format in test version A, the participants could perform to a greater extent than using the short answer format in test version B. Anyhow, the test takers also performed better with the short answer format in the test version B compared to the multiple choice format in the same version. The height of the bar graph in figure 4.2 clearly illustrates that the participants performed better in the short answer format in both test versions. They performed quite similarly in multiple choice format in test version A and B.

In analyzing an interaction, an ANOVA would look for the difference among the cell means. The present study produced the following cell means:

Table 4.11
The Cell Means for the Combination of Formats and Varieties

<i>Varieties of English</i>	<i>Test formats</i>	
	Multiple choice (MC)	Short answer (SA)
Native Speakers	<i>Native MC</i> 11.343	<i>Native SA</i> 13.229
Non-native Speakers	<i>Non-native MC</i> 11.020	<i>Non-native SA</i> 11.656

For the present study results, the differences among these cell means are large enough to say that they are statistically significant ($p = .013$). However, it cannot state precisely where the difference lies. By plotting the graph, we could look at the mean for each group and tell where the differences are the greatest or the smallest. To be certain precisely where the differences occur, a post hoc comparison of the means is usually conducted in an ANOVA. Anyhow, a Mixed within-between ANOVA used in the present study does not allow the post hoc comparison because there are only 2 levels in each case.

From Table 4.9, there was not much difference between the scores in the test version A and B for the multiple choice format. The mean scores of 11.343 for test version A and 11.020 for B are so close that we cannot be certain whether they are statistically different; consequently, the independent *t*-test was conducted to study the difference between the two means. The results from the *t*-test (*t* value = .612, *p* = .541) show that the difference between the two means was not great enough. It can be claimed that the multiple choice format produced similar scores in both test versions. Combining native or nonnative speaker varieties with the multiple choice format did not have a significant effect on the test takers' listening comprehension scores.

On the other hand, the mean scores for the short answer format with test versions A and B look quite different. The two means are 13.229 for version A and 11.656 for version B. The results from the *t*-test (*t* value = 2.942, *p* = .004) confirm their difference. The two means were significantly different which means that combining accent varieties with short answer format had an effect on the scores, apparently, the combination of short answer format with native speaker voice stimuli could produce an easier form of listening comprehension test.

On the other side, considering the aspect of accent varieties, using native speaker voice input with multiple choice and short answer format seems to make a big difference when looking at the mean scores of 11.3438 and 13.2292. The pair sample *t*-test was conducted since the participants of the two groups for the comparison now were the same people. The results, which were *t* value = 5.135, *p* < .001, demonstrate that combination of native speaker varieties with different formats gave different results and the short answer format provided higher scores. Another comparison to be considered is the combination of nonnative speaker varieties with the two formats. The mean scores of 11.0208 for MC format and 11.6563 for SA format are again quite close, and the *t* value of 1.889 and *p* value of .062 reveal that, unlike native speaker voice input, using nonnative speaker voice as test stimuli gave similar scores in both test formats. Thus, for the test takers, nonnative speaker varieties were still difficult to comprehend and changing the test format did not give much assistance for them.

To sum up, the interaction effect result shows that the combination of the two variables made a difference in the test takers' scores. The test takers did not perform differently in the multiple choice test format in both test versions. On the contrary,

they performed differently in the short answer format when it was combined with native speaker test version.

Therefore Hypothesis 3, which states there is a significant interaction effect between the test format and accent varieties of English, is accepted. The effect size value for the combination of the two variables was small to modest. This is due to the fact that the multiple choice format variable did not produce much difference in scores in both test versions because it was difficult for the test takers; whereas, the combination of the short answer format, which is less difficult, with native speakers' accent varieties has a greater impact on the participants' scores.

PART TWO: FINDINGS FROM THE SEMI-STRUCTURED INTERVIEW

In order to gain more in-depth information from the test takers concerning their attitudes and preferences towards accent varieties of English, semi-structured interviews were conducted a week after the test takers had participated in the listening comprehension test. Thirty participants who took the test were invited for the interview. These 30 participants were randomly selected from both groups who took version A and B test. The demographic data concerning the present study i.e. gender, the test version they took, and the scores they obtained from the listening comprehension test of the present study are presented in Table 4.12.

Table 4.12 reveals that majority of the sampled interviewees were females which accounts for 80% from the total number of the interviewees. Only 20% were males. The number corresponds with the character of the participants who took test versions A and B in that the female subjects were a lot more numerous than male subjects. From 192 participants, there were 146 females and 46 males which account for 76% and 24% respectively.

In terms of the interviewees' scores obtained from the listening comprehension test in both versions, the samples from all score range were successfully selected in a good proportion. However, the number of the interviewees who took test version B is more than the number of test version A (56.7% and 43.3%). This is because there were two selected students from test version A that did not turn up for the interview. To fix for these two students, another two participants from test version B who shared the same scores were called for the interview instead.

Table 4.12
Demographic Data of Thirty Interviewees

	Samples	Percent
Test Version they took		
Test A. NS	13	43.3%
Test B. NNS	17	56.7%
Total	30	100%
Gender		
Female	24	80%
Male	6	20%
Total	30	100%
Scores obtained		
11 – 15	3	10%
16 – 20	6	20%
21 – 25	9	30%
26 – 30	7	23.3%
31 – 35	2	6.7%
36 - 40	3	10%
Total	30	100%

I. The investigation of the test takers' attitudes towards English accent varieties

The sampled interviewees were asked to listen to the recorded speech and gave instructions without knowing which varieties were included. The 14 speech samples - 2 voices for 1 accent variety- lasted about 30 seconds each. The 7 varieties, which were from the United Kingdom, the United States of America, Australia, Japan, China, Malaysia, and Singapore, were arranged to be heard at random. While listening to each speech sample, the interviewees were asked to give their opinions on a four-point scale rating according to the answer sheet with the ten adjectives (Appendix J). All interview parts and instructions given were conducted in their native Thai language so that the participants had a clear explanation of the meaning of

adjective words used. The results of the mean score are shown in Table 4.13. The higher the score, the more positive the interviewees felt for that adjective word.

Table 4.13
Mean Scores of Varieties for Ten Adjectives

ADJECTIVES	ENGLISH VARIETIES							
	UK	AME	AUS	MAL	JAP	CHI	SING	
Solidarity	Sociable	2.91	3.10	2.53	2.50	2.70	2.23	2.51
	Sincere	2.75	3.08	2.56	2.46	2.70	2.75	2.55
	Comforting	2.63	2.83	2.63	2.33	2.65	2.48	2.46
	Friendly	2.46	2.96	2.51	2.30	2.91	2.46	2.53
	Reliable	2.90	3.23	2.70	2.50	2.63	2.33	2.41
Status	Elegant	3.03	2.45	2.51	2.53	2.36	2.10	2.21
	Educated	3.35	3.38	2.98	3.00	2.80	2.45	2.70
	Intelligent	3.31	3.35	2.91	2.70	2.60	2.36	2.66
	Wealthy	2.95	2.90	2.58	2.40	2.35	2.08	2.33
	Successful	3.11	3.31	2.78	2.61	2.70	2.33	2.53
Total mean	2.94	3.10	2.66	2.53	2.64	2.35	2.48	
NS	2.90		NNS		2.50			

The next step was to analyze the data. First, attitude differences among seven varieties for the solidarity and status dimensions were investigated. The mean scores for each dimension were ranked. The mean scores of the five adjectives for solidarity dimension and the other five adjectives for the status aspect are demonstrated in Figures 4.3 – 4.4 and Tables 4.14 – 4.15.

Figure 4.3
The Solidarity Means

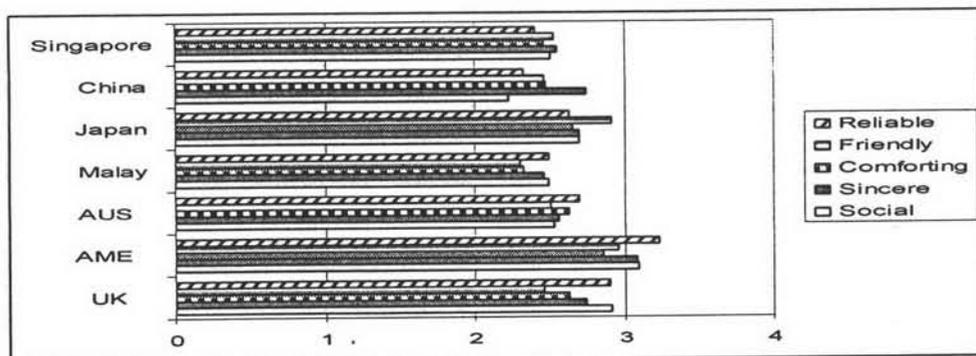


Figure 4.4
The Status Means

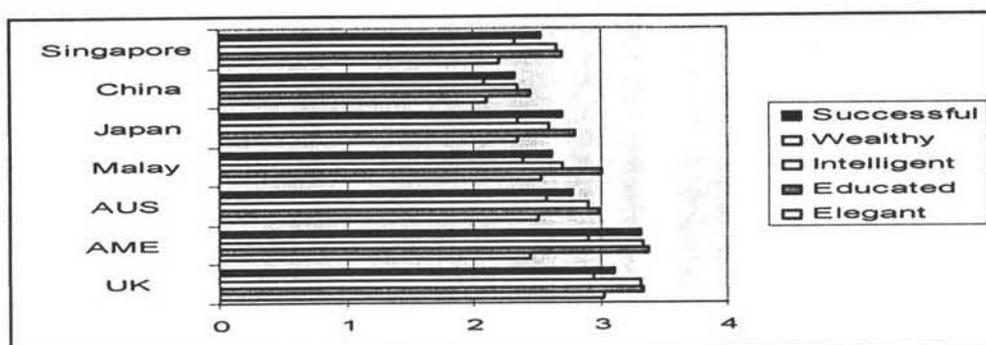


Table 4.14
Mean Scores Ranking for Solidarity Dimension

Varieties	Mean Scores	Mean Rank
America	3.04	1
United Kingdom	2.74	2
Japan	2.73	3
Australia	2.58	4
Singapore	2.49	5
China	2.45	6
Malaysia	2.41	7

Table 4.15
Mean Scores Ranking for Status Dimension

Varieties	Mean Scores	Mean Rank
America	3.16	1
United Kingdom	3.15	2
Australia	2.75	3
Malaysia	2.66	4
Japan	2.56	5
Singapore	2.48	6
China	2.26	7

In terms of solidarity dimension, the American variety was rated the highest and the Malaysian variety was rated the lowest. Also, the American variety was ranked the highest in the status dimension. The responses given by the 30 test takers showed that the American variety is the most acceptable accent in both solidarity and status aspects. Further, a hierarchy on the basis of attitude ranking and the total mean scores (NS mean = 2.90, SD = .222; NNS mean = 2.50, SD = .120) reveal that, on average, native speakers' varieties were rated higher than nonnative speakers' varieties, except on the aspect of solidarity that the Japanese variety received a higher score than the Australian's. Among native accent varieties, the American variety is the most preferred accent while the Japanese accent received outstanding attitude scores among nonnative accents. *It can be concluded from the data that the test takers have more positive attitude towards the native speakers' accent varieties than the nonnative speakers' accent varieties of English.*

II. The investigation of the test takers' preferences towards using accent varieties of English in a listening comprehension test

After listening to 14 speech samples and expressing their attitudes on the English varieties, the 30 participants were interviewed individually to find out their preferences towards using varieties of English accents as stimuli for a listening comprehension test. Seven questions were asked. The information gathered is presented in Table 4.16.

Table 4.16

Interview Data of Preferences towards Accent Varieties

Questions and Answers	Frequency	Valid Percent
1. Are you able to recognize different accent varieties of English?		
Yes	19	63.3%
No	11	36.7%
total	30	100%
2. Which accent do you find it difficult to comprehend?		
Native speakers	6	20%
Nonnative speakers	24	80%
total	30	100%

Questions and Answers	Frequency	Valid Percent
3. Do you find different accents equally pleasing?		
Yes	4	13.3%
No	26	86.7%
total	30	100%
4. Do your judgments depend on the voice of the speakers or the contents of the utterance?		
Voice	19	63.3%
Content	10	33.3%
Both	1	3.3%
total	30	100%
5. Do you like the listening comprehension test to incorporate accent varieties of English?		
Yes	10	33.3%
No	20	66.7%
total	30	100%
6. Does the inclusion of English accent varieties make you uncomfortable?		
Yes	14	46.7%
No	16	53.3%
total	30	100%
7. Does the inclusion of English accent varieties make the listening comprehension test more difficult?		
Yes	26	86.7%
No	4	13.3%
total	30	100%

From Table 4.16, the percentage of sampled interviewees who indicated that they were able to recognize different accent varieties of English is 63.3%. About 36.7% of the test takers admitted that they could not recognize different English accents used in the test. Most of the test takers (80 %) agreed that nonnative varieties were more difficult to comprehend, and more interviewees which are about 86.7%

thought there were some varieties that were more pleasing to hear for them. Further, a lot of interviewees (19 out of 30 participants) revealed that they made their judgments on the voice rather than the content of speech they heard.

When asked if they wanted the listening comprehension test to include different varieties of English, about two third of the interviewees, which was 66.7 %, disagreed with the idea to include different varieties of English as listening stimuli in the test. The reason was disclosed by their answer to question seven, when 86.7% of the test takers thought that inclusion of English varieties made the test more difficult. *It can be concluded from the data that the test takers expressed stronger preference to use the native speakers' accent varieties as the voice stimuli in the listening comprehension test.* However, in the answer to question six, more than half of the interviewees did not feel uncomfortable when listening to different English accents.

The interviewees were asked to express their preferences openly concerning their likes and dislikes on inclusion of English varieties in the listening comprehension test and their answers were varied. The findings are described in detail for each question.

Interview Question 1: “Are you able to recognize different accent varieties of English?”

Most of the test takers were confident that they were able to recognize English accent varieties. Approximately 63.3 % of the interview participants were certain that they could differentiate between native and nonnative varieties. Examples of some of their answers are:

- a. Yes, I can recognize the accents. Are they from India?
- b. Yes, I am quite sure that there are some Indonesia, and the Philippines.
- c. Of course, they are Japanese, French and Spanish.
- d. I know, I could hear the Indian English quite clear. Many varieties are from Asia, China, Malay and the Philippines.
- e. There are a variety of nationalities such as Korea, Japan, China, India, and Vietnamese. There are also some Europeans too.
- f. I'm sure they are native speakers of English, I can tell.
- g. Oh, I can tell that they are not native speakers. Why are these voices included in the test? I didn't pay attention to their accents, I want to understand and I don't think they are appropriate for listening lessons.

There were a lot of the interviewees who thought that they were able to recognize English accent varieties but not all of them could give correct varieties. A crosstabulation was conducted to investigate the results of correctness as shown in the table below.

Table 4.17
Varieties Recognition and Correctness Crosstabulation

			Correctness		Total
			Correct	Not correct	
recognize	Yes	Count	7	12	19
		% within right	77.8%	57.1%	63.3%
		% of Total	23.3%	40.0%	63.3%
	No	Count	2	9	11
		% within right	22.2%	42.9%	36.7%
		% of Total	6.7%	30.0%	36.7%
Total		Count	9	21	30
		% within right	100.0%	100.0%	100.0%
		% of Total	30.0%	70.0%	100.0%

There were 19 people who mentioned that they were able to recognize the varieties of English. However, only 7 of them could correctly identify the speakers' place of origin. The number 7 out of 30 accounts for 23.3% of the total. *This means that only 23.3% of the respondents could tell the origin of the speakers.* Moreover, there were also 2 people who successfully told the correct varieties but they indicated that they were not able to identify the varieties.

Question 2: "Which accent do you find it difficult to comprehend?"

Nonnative speakers' English varieties were pointed out by majority of the participants to be more difficult to comprehend when compared to those native speakers'. 80% of the interviewees agreed that nonnative varieties were hard to listen for them. Among these varieties, the interviewees have a preference for the American variety. Second to the American's is the British accent which also was often mentioned by the test takers. It is interesting to note that, though some interviewees said that they preferred one accent to others, they did not really know how the accent said to be preferred sounded like. For instance, one female respondent said in the interview that she very much liked the British accent. However, when asked to give

the name of the British teacher whose accent she liked, she gave the name of an American teacher.

The information on the accent varieties that are perceived as easy or difficult to understand by the respondents is demonstrated in Figures 4.5 – 4.6.

Figure 4.5

The English Accent Varieties Preferred by the Interviewees

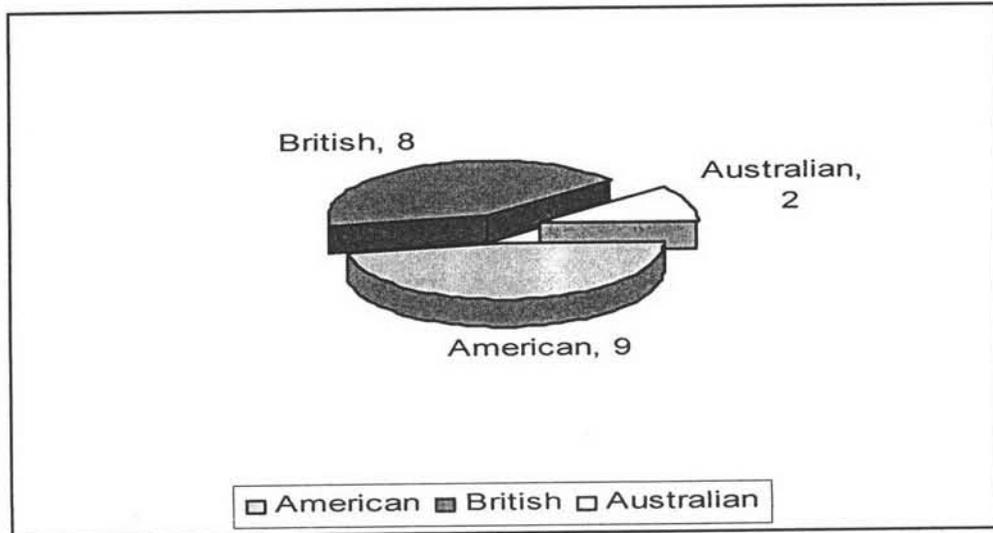
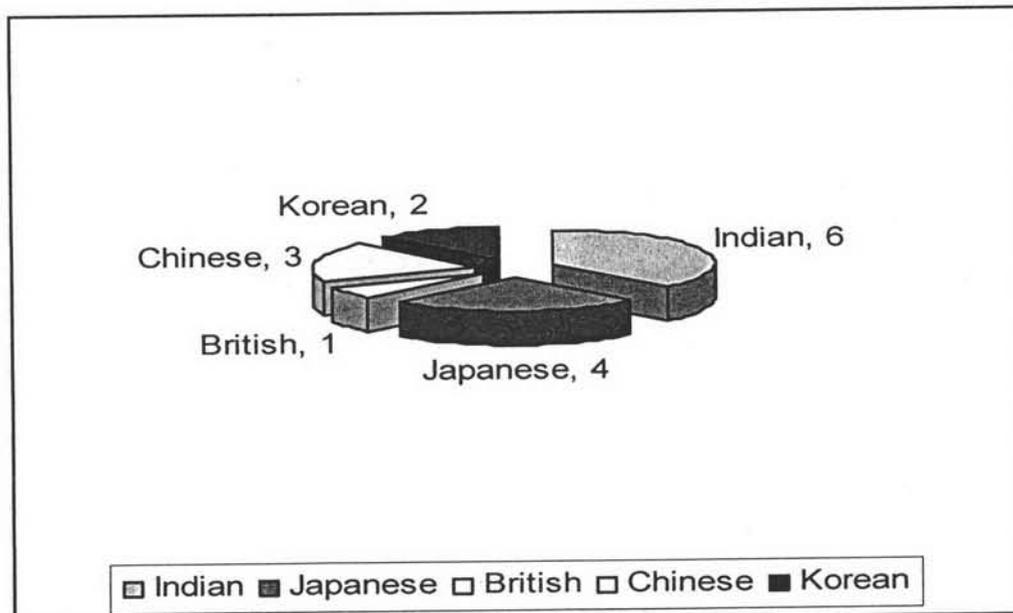


Figure 4.6

The English Accent Varieties Difficult to Comprehend



From Figure 4.5 and Figure 4.6, it is important to note that, although 20% of the respondents did not say that nonnative varieties were more difficult to understand, none of them mentioned that they preferred any nonnative speaker accent. They just only mentioned that all English accent varieties were hard to comprehend for them. Furthermore, when asked about their opinion on the most difficult accent variety they have heard, a lot of respondents mentioned some varieties such as Indian and Korean that were not included in the test from their own experience or belief before taking the test. Not all of the respondents named specific varieties that they preferred or not preferred. These respondents only named these varieties as native or nonnative English. *In conclusion, it is obvious that nonnative accent varieties were perceived to be more difficult to comprehend than those native accent varieties for the majority of the respondents, and some of the respondents had fixed stereotype on a particular accent.*

Question 3: “Do you find different accents equally pleasing?”

Most of the test takers (86.7%) did not find different accents equally pleasing and they said they liked native speakers’ varieties more. The answers for question 3 confirm the result of question 2 that the majority of the interviewees had a preference for native speakers’ varieties. Examples of their answers for question 3 are as follows:

- a. Not equal, I don’t want to listen to things that I can’t understand or are difficult to understand. If I can catch what they mean, their accent is pleasing for me.
- b. Not equal but I don’t know which one is more pleasing, I can’t tell.
- c. No, some of them sound very boring. It’s like they are being forced to speak (nonnative speaker’ voices).
- d. Oh, it’s terrible. I don’t want to hear them. I’m not happy (nonnative speakers’ voices).
- e. Oh, I don’t like the Indian accent.
- f. Oh, I don’t care but if it is like the native speaker’ voice that would be good.
- g. Some nonnative speakers’ accents sound unfriendly to me.
- h. Some accents (nonnative speakers’ accents) seem to be aggressive. I like those accents that are easy to comprehend.
- i. If they sound like native speakers, they are fine for me.
- j. I want to sound like native speakers.

- k. I want to sound British.
- l. I like the British accent.
- m. I like the American accent.

It can be concluded from these responses that the respondents did not find different accents equally pleasing and none of them showed favor to any nonnative varieties.

Question 4: “Do your judgments depend on the voice of the speakers or the contents of the utterance?”

More than half of the interviewees (63.3%) disclosed that their judgment on the English utterance they heard depended very much on the voice or accent rather than the content of the speech. Only one respondent said that she considered both factors equally. Some of them mentioned the emotion of the speakers as well. On top of that, it is worth to mention that there were 13 respondents who said, *“If a person sounds like a native speaker, it makes that person look smarter, more educated, and have a better personality than a person who sounds different.”* This reflects the rigid stereotype of native varieties of English among the participants. *To conclude, accents play a major role on the respondents’ judgments of the utterance.* To the respondents in this situation, native English accents could make the utterance being heard sound superior.

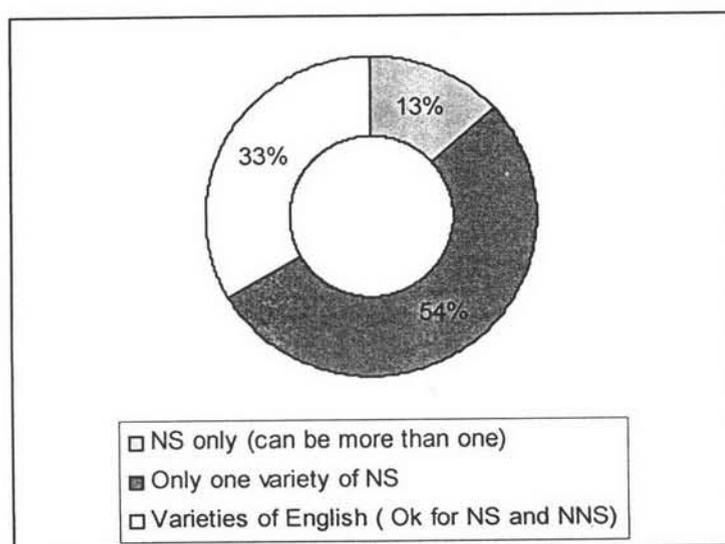
Question 5: “Do you like the listening comprehension test to incorporate accent varieties of English?”

The majority of the interviewees (66.7%) did not want the test to use nonnative varieties of English for listening stimuli, while the minority of the participants (33.3%) thought it was more authentic to incorporate these nonnative varieties. Their preferences on this matter are varied and can be categorized into 3 types as described in the following table and figure.

Table 4.18
Varieties Preferences from the Interview

Preferences	Frequency	Percent
Native speakers varieties only (can be more than one)	4	13%
Only one variety of NS	16	54%
Varieties of English (Ok for NS and NNS)	10	33%

Figure 4.7
Varieties Preferences from the Interview



Question 5 findings confirm the previous interview questions in that the respondents had a preference for native speakers' varieties. However, 13% of the interviewees mentioned that they wanted to listen to only one native accent variety. Their reasons share similarity in that they did not want to get confused and be frustrated when taking an important test. The following are the reasons from those who prefer native speaker varieties:

- a. I understand native speakers better.
- b. I have studied with native speakers and am more familiar with native speakers' accents
- c. If it is an important test, I prefer to listen to only one native speaker variety as I don't have to adjust to new accents again and again.
- d. I don't want different accent varieties, only one native speaker is the best for me because I can have enough time to get familiar.

- e. I don't want varieties because I want to score high in a test even though I know it is useful to listen to them.
- f. I was confused when listening to these nonnative speakers. I want only one native accent.
- g. I want to hear only the variety that I like. It is easy when taking a test.
- h. Even I know that there are more nonnative speakers of English in the world, I don't want these varieties to be included in an important examination.

For the interview participants who agreed that a listening comprehension test should incorporate varieties from both native and nonnative speakers based their reasons mainly on authenticity. They want to practice listening to varieties and also want to evaluate their own ability. Examples of their reasons are:

- a. I am fine with varieties if speakers speak clearly.
- b. Varieties are good to hear. I like to listen to things that are different.
- c. I want to practice listening to different accents so my listening ability will be improved.
- d. I think it is beneficial if we are provided with different English varieties, so we can practice and we can evaluate ourselves.
- e. It is a must to incorporate varieties both in listening lessons and in listening tests because in reality there are a lot more nonnative speakers. To have good listening comprehension ability in English should mean a person can understand different varieties too.
- f. It is interesting to hear different varieties. I want to try.
- g. Yes, I think it is good to have varieties in a test. I think having varieties in a test helps increase my chance to understand the speech, for example, if I can't catch a particular accent, I still have a chance with another variety used in the same test.
- h. It's kind of fun and new. I think it is challenging though I know it will be more difficult, I still want to try.
- i. It is exciting and amusing to hear varieties.

It can be concluded from the interview data that using accent varieties of English in the test was new to most of the test takers. English varieties especially of nonnative speakers were accepted among a minority of the test takers.

Question 6: “Does the inclusion of English accent varieties make you uncomfortable?”

More than half of the respondents (53.3%) said they were not uncomfortable to hear varieties even though they did not have much comprehension. Most of the reasons for those who said they were uncomfortable concerned their anxiety for not being able to answer the test questions. A lot of interviewees admitted they were stressed when listening to nonnative speakers’ varieties. Their answers to these questions can be summarized by using the words as seen in Table 4.17.

Table 4.17
Interviewees’ Feelings When Hearing Varieties

Feelings	Frequency
• Indifferent	16
• Uncomfortable	9
• Stressful	5
• Confused	2
• Nervous	1
• Worried	1
• Fun	1
• Awkward	1

It can be summarized that the majority of test takers did not have negative feeling towards varieties of English if they were not used as listening test voice input. In general situation, accent varieties of English can be interesting for a few of respondents to test their own listening ability.

Question 7: “Does the inclusion of English accent varieties make the listening comprehension test more difficult?”

The majority of the sampled interviewees agreed that the inclusion of English varieties would make the test more difficult. The percentage of the respondents who supported this view was very high at 86.7%. Five of the respondents believe that it was because they were not familiar with these varieties. One respondent said, when hearing varieties with which were not familiar, that she just ‘shut her ears down’. She

realized later that the test was not too difficult but she was nervous. Moreover, some of them said they were 'having a headache' listening to these unfamiliar varieties. They could not adjust and concentrate.

There were two people who thought that it was not because of varieties but the test itself was difficult. One student said that accent varieties made the test more difficult but only to a small degree. A few interviewees mentioned that they were having fun listening to varieties and doing the test. This means that though, for most of the respondents, the inclusion of accent varieties of English in the test made the test more difficult; there were some test takers who found these accent varieties interesting and encouraging.

To conclude, the information from the preference questions supports those from attitude questions that the test takers had a more positive attitude towards native speakers' varieties than nonnative speakers' accents. The positive attitude led to their preference for the varieties; therefore, they preferred a listening comprehension test to include only native varieties of which they have had more exposure.

Summary

This chapter reports the results of the findings. The results are presented according to the research objectives set. The mixed within-between ANOVA was employed to answer the first three quantitative research questions. For Research Question 1, it was found out that the test takers' listening ability was affected by the test format used and the multiple choice format was reported to be more difficult than the short answer format for the test takers. For Research Question 2, the use of varieties of English accents as listening input had a significant effect on test takers' performance. The results reveal that the test version with nonnative speakers' voice was harder than the test version with native speakers' voice. For Research Question 3, the interaction effect between the use of test formats and the use of English varieties was significant. The combination of the multiple choice format and nonnative varieties in the listening comprehension test was the most difficult test version. On the contrary, the combination of the short answer format and native speakers' voice was the easiest test for the test takers in the present study situation.

For the interview part, to answer Research Question 4, mean scores of the attitude scale were calculated to indicate the test takers' attitude towards the varieties of English. Further, the interview data concerning the test takers' preference for the

English varieties used in the test was displayed in the forms of percentage and frequency with graphs and charts. It could be summarized from the attitude questions that the respondents had a more positive attitude towards the native varieties than the nonnative varieties. Further, their positive attitude led to their preference for native varieties of English in the listening comprehension test.