CHAPTER FOUR

RESULTS AND DISCUSSIONS

In Chapter Four, the in-depth analyses are presented on the basis of the statistical data and subjects' response to the feedback questionnaire.

The first section provides the results of data analyses by One-Way ANOVA Repeated-Measures for Dependent Samples to examine the comparison significance for the performance on the GEPT posttest and pretest of all the subjects and of the different proficiency levels.

The second section discusses the comparison of mean scores in different texts also by One-Way ANOVA Repeated–Measures. Moreover, it is also shown in the section the correlation between text types and the performance on the GEPT posttest among all the subjects as well as the different proficiency levels

The analysis of the correlation between different task types, task types and the performance on the GEPT posttest of all the students as well as the different proficiency levels are presented respectively in the subsequent section.

The final one pertains to the percentile analyses for the results of students' responses to the feedback questionnaire, including the opinions on application of authentic broadcast as supplementary materials, designing of different task types and suggestions on listening curriculum.

Findings drawn from the analysis and the possible reasons behind them along with the correspondence to the research questions will be further discussed in the next chapter.

<u>**Results of Statistical Analyses**</u>

4.1 Comparison of the Performance on the GEPT Pretest and Posttest

In order to examine whether there is any improvement in the GEPT posttest, a One-Way ANOVA Repeated-Measures for Dependent Samples is performed among all the students and different proficiency levels as well. Table 4-1 and 4-2, the One-Way ANOVA results, present the difference between the GEPT pretest and posttest.

Table 4-1

Means and Standard Deviations for the Performance on the GEPT Pretest and Posttest

	Mean Scores of GEPT		Std Deviation		N
	Pretest	Posttest	Pretest	Posttest	
H	96.91	99.89	8.8435	8.6548	17
M	71.77	76.33	3.7591	9.7674	17
L	51.11	55.69	5.2113	9.6796	7
All	78.67	82.57	18.3156	18.7793	41

Note: 1.H: high-proficiency level

2.M: middle-proficiency level

3.L: low-proficiency level

4.N: numbers of subjects

5.The total raw scores of the GEPTs are 120.

As shown in Table 4-1, the subjects' overall performance on the GEPT posttest (total mean score: 82.57), compared with that of the pretest (total mean score:

78.67), does show progress on the average scores. Likewise, the mean scores of the posttest all prevail those of the pretest in the three proficiency levels. The summary of One-Way ANOVA for the significance of the difference of means is presented in Table 4-2.

Table 4-2

Summary of One-Way ANOVA for the Difference of Means between GEPT Pretest and Posttest

Proficiency Level	Sources of Variables	SS	df	MS	F	Sig.
H	Post-Pre	75.692	1	75.692	4.260	.056
	Error(Post-Pre)	284.317	16	17.770		
M	Post-Pre	176.609	1	176.609	2.901	.108
	Error(Post-Pre)	973.909	16	60.869		
L	Post-Pre	73.326	1	73.326	.837	.395
	Error(Post-Pre)	525.502	6	87.584		
All	Post-Pre	313.211	1	313.211	6.975	**.012***
	Error(Post-Pre)	1796.145	40	44.904		

Results obtained from the One-Way ANOVA summary show that for all the subjects the difference between GEPT Pretest and Posttest presents significant difference ($F(1,40)=6.975$, $p=.012<.05$). Needless to say, the results demonstrate a significant improvement in students' performance on the GEPT posttest. As for the individual levels, the difference between the GEPT pretest and posttest is not

significant (**H**: p= .056; **M**: p= .108; **L**: p= .395). However, the p value for the high-achievers is almost close to the significant value p=.05. As shown in Table 4-1, the middle- and low-achievers have higher average on gain-scores (76.33-71.77= **4.56**; 55.69-51.11= **4.58**) than that of the high-achievers (99.89-96.81= **2.98**), whereas the difference between the GEPT pretest and posttest still doesn't meet the criteria for significant difference. Maybe this is due to the small pool of subjects in the present study.

4.2 Comparison of Different Text Types

This section presents the overall means of production of the ten different texts chosen from authentic ICRT broadcast. Moreover, the test of within-subjects effects on different text types will also be presented to see if there is any significant difference among various text types. With respect to the correlation between individual text types and the GEPT posttest, Pearson Correlation for Dependent Samples is administered among overall as well as individual proficiency levels.

4.2.1 Comparison of the Mean Scores[1] of Different Text Types for All Subjects

As shown in Table 4-3, among the ten texts, the highest overall average is weather forecast item 3 (Appendix C), which is about world weather forecast by Richy Walker, whose speed is the fastest one. The second in rank is Special Topic item 3 (Appendix H), which is about the rising career of a singer, Mariah Carey. The third is Easy News item 1 (Appendix I), about a movie star, Russell Crowe. Nevertheless, compared with these ten texts, grouped according to their sub-topics, the highest is commercials (total mean = .7161); the second is Easy News (total mean

---

[1] As mentioned earlier in Chapter Three, the mean scores of each text are computed based on the ratio of correctness.

= .6876), while weathers (total mean = .6073) rank the lowest.

Table 4-3

Means and Standard Deviations of All Subjects' Performance on the Ten Different Texts

Text Types		Mean	Std. Deviation	N
Weather Forecast	**1**	.5844	.1929	41
	2	.4593	.2900	41
	3	.7756	.1772	41
Total Weathers		.6073	.1767	41
Commercials	**1**	.7061	.1761	41
	2	.7220	.2319	41
Total Commercials		.7161	.1718	41
Special Topics	**1**	.5100	.2283	41
	2	.7020	.1514	41
	3	.7707	.1419	41
Total Special Topics		.6612	.1408	41
Easy News	**1**	.7578	.1731	41
	2	.6134	.1523	41
Total Easy News		.6876	.1353	41

Table 4-4 presents the test of within-subjects effects on different text types to examine if there is any significant difference of the means of different texts.

Table 4-4

Summary of One-Way ANOVA on Different Text Types

Source	SS	df	MS	F
Text Types	4.675	9	.517	22.568*
Error (Text Types)	8.255	360	2.293E-02	

Note: 1. *The mean difference is significant at the .05 level

Results obtained from Table 4-4 show that there is a significant difference in the means of various texts (p=.000, <.05). The result implies that the ten chosen authentic materials are significantly different and are suitable for teaching listening comprehension.

As for the effects of using these texts as supplements, a Pearson Correlation is used to see the correlation between individual texts and the GEPT posttest in Table 4-5.

4.2.2 Pearson Correlation Between Text Types and GEPT Posttest for All Subjects

Table 4-5 shows the correlation between individual text items and the performance on the GEPT posttest.

Table 4-5: Correlation between Text Types and GEPT Posttest[2]

Variables	Correlation	Sig.	N
GEPT Posttest -Weather 1	.275	.082	41
GEPT Posttest -Weather 2	.522**	.000	41
GEPT Posttest -Weather 3	.697**	.000	41
GEPT Posttest -TW	.617**	.000	41
GEPT Posttest -Commercial 1	.669**	.000	41
GEPT Posttest -Commercial 2	.365*	.019	41
GEPT Posttest -TC	.586**	.000	41
GEPT Posttest -Special Topic 1	.683**	.000	41
GEPT Posttest -Special Topic 2	.528**	.000	41
GEPT Posttest -Special Topic 3	.415**	.007	41
GEPT Posttest -TS	.701**	.000	41
GEPT Posttest -Easy News 1	.558**	.000	41
GEPT Posttest -Easy News 2	.616**	.000	41
GEPT Posttest -TE	.703**	.000	41

Note: 1.*Correlation is significant at the .05 level (2-tailed)
2.** Correlation is significant at the .01 level (2-tailed)
3.TW: total weathers
6.TC: total commercials
7.TS: total special topics
8.TE: total Easy News

---

[2] The scores of GEPT posttest are the ratio between the raw score and the total score while the scores of the texts are computered on the basis of the ratio between the number of correct answerw and the number of the test items.

Overall, the Pearson Product-Moment correlation coefficients reported in Table 4-5 indicate a high degree of correlation between each of the texts and the GEPT posttest. Except for the text weather item 1 (p = .082 >.05, not significant), the validity coefficients of each text, which range from .365 to .697, are all significant at the .05 level or even .01 level of probability. According to the definition of Pearson's r correlation coefficients[3], though commercial item 2 (r = .365) has low correlation, the other text items (except weather item 1) all reach the middle correlation, and weather item 3 (r = .697) even reaches the high correlation.

As shown in the above Table, for these ten texts grouped according to their sub-topics, the validity coefficients of each total sub-topic (total weathers: p = .000; total commercials: p = .000; total special topics: p = .000; total Easy News: p = .000) are all significant at the .01 level. The results indicate that all the chosen text types are correlated with the GEPT posttest. Among them, total Special Topics (r = .701) and total Easy News (r = .703) are even highly correlated with the GEPT posttest. The result implies that the performance on GEPT could be highly predictable by applying these chosen authentic materials to train students' listening comprehension.

Since listeners' proficiency is a determining factor on EFL listening comprehension (Li, 1999), in the following two sections (4.2.3 and 4.2.4), the writer tries to further analyze the means of different text types and the correlation between individual text types and the GEPT posttest in terms of the distribution of proficiency.

---

[3] The definition and meaning of **Pearson's r correlation coefficient** in absolute value: 1.00 (with complete correlation); .79~ .99 (with high correlation); .40~ .69 (with middle correlation); .10~ .39 (with low correlation); below .10 (with little or no correlation). (Chiou, 2002, p.12-6)

4.2.3 Comparison of the Mean Scores of Different Text Types with Proficiency Distribution

Table 4-6

The Means and the Standard Deviations (SD) of the Three Proficiency Groups' Scores in the Different Text Types and GEPT Posttest

Proficiency Text Levels Types	High (n=17)		Middle (n=17)		Low (n=7)	
	Mean	SD	Mean	SD	Mean	SD
TW	.7247	.1054	.5553	.1616	.4486	.1811
TC	.7988	.1312	.7047	.1456	.5429	.1988
TS	.7629	.1128	.6124	.1133	.5329	9.268E-02
TE	.7929	8.327E-02	.6224	.1063	.5900	.1371
GEPT Posttest	.8329	7.235E-02	.6365	8.200E-02	.4643	8.162E-02

Note: 1.TW: total weathers

2.TC: total commercials

3.TS: total special topics

4.TE: total Easy News

As shown in Table 4-6, the high-proficiency group performs better in all the text types than the other two groups. Among them, the high-achievers perform the best in the text types of commercials and Easy News. Moreover, the mean of total commercials in the middle-proficiency group (mean= .7047) is almost close to that of the high-achievers. The result may probably be due to the nature of these two commercials about Chunghua Telecom and McDonald's, which are closely related to students' daily life, and students are more exposed to such kind of knowledge.

4.2.4 Pearson Correlation between Text Types and GEPT Posttest for the Three

Proficiency Groups

Compared with the correlation coefficients between different text types and the GEPT posttest, as shown in table 4-7, only one significant correlation is found in the high-proficiency level, which is Special Topics (Pearson's r =.585, p =.014<.05). Although the high-achievers score high in the means of total commercials and Easy News (see Table 4-6), they are not significantly associated with the performance on the GEPT posttest.

With respect to the middle-proficiency group, a significant correlation between the text types and the GEPT posttest is found in commercials, Special Topics and Easy News. The correlation coefficient between commercials and the GEPT posttest is .696, at the significant level of .01, which also indicates a very strong correlation with the posttest by the definition of Pearson's r correlation coefficient. In addition, the significant correlation coefficients between Easy News and the GEPT posttest (Pearson's r =.571, p =.017<.05) as well as between Special Topics and the GEPT posttest (Pearson's r =.490, p =.046<.05) indicate that middle-achievers' performance on the GEPT posttest could be predicted by the performance of such text types. In other words, the correlation results indicate that if middle-achievers get good scores in such text types as commercials, Special Topics and Easy News, they would most probably do well on GEPT.

As for the low-proficiency level, there is no significant correlation between the text types and the GEPT posttest.

Table 4-7

Pearson Correlation Coefficients (r) between the GEPT Posttest and Individual

Text Types in the Three Proficiency Groups

Proficiency Text Levels Types	High (n=17)	Middle (n=17)	Low (n=7)
	GEPT Posttest	**GEPT Posttest**	**GEPT Posttest**
Total Weathers	.169	.447	-.173
Total Commercials	.070	.696**	.031
Total Special Topics	.585*	.490*	-.546
Total Easy News	.302	.571*	.271

Note: 1.** Correlation is significant at the .01 level (2-tailed).

2.* Correlation is significant at the .05 level (2-tailed).

4.3 Comparison of Different Task Types

In this section the task types, which were regrouped[4] into four categories from all the text questions, are analyzed. First, the results of One-Way ANOVA Repeated-Measures for all subjects show the differences of individual task types in Tables 4-8 and 4-9. Then, Pearson correlation statistics is administered to see if there is any significant correlation between task types and the GEPT posttest in terms of the distribution of proficiency.

The following comparison provides us with a general knowledge on which design of task types is the most beneficial to the students' listening training.

4.3.1 Comparison of Mean Scores of Different Task Types for All Subjects

Table 4-8

Means and Standard Deviations of All Subjects' Performance on Different Task Types

Task Types	Mean	Std Deviation	N
Blank-Filling	.6054	.1708	41
True/False	.7354	.1264	41
Multiple Choices	.7044	.1202	41
Cloze	.6427	.1586	41

As seen in Table 4-8, the subjects do better on true or false statements (mean= .7354) and multiple choices (mean= .7044) than on cloze (mean= .6427) and blank-filling (mean= .6054). The results are in correspondence to some of Teng's

---

[4] Since the types of test questions vary according to the design of materials, and for the convenience of statistical interpretation, task types of the same kind, though in different texts, are grouped in the same category, including blank-filling, true/false, cloze and multiple choices. The scores of each task types are collected by the ratio of correctness.

(1998) and Chen's (2003) studies. It is found in Teng's study that the mean of the cloze task type is lower than that of multiple choices, and in Chen's, the means of blank-filling test types are lower than those of multiple choices but the same as that of true or false statements. The results indicate that task types requiring students to write their answers on the test paper (i.e. blank-filling and cloze test) are more challenging for the subjects and that "it indeed will take the subjects a period of time to get ready for this kind of test type"(Chen, 2003, p.8).

The results above reveal that familiarity of test questions and avoiding its answers in written forms are the key for developing self-confidence in testing listening comprehension. As shown in Table 4-8, task types, such as true or false statements and multiple choices provide the subjects with a better chance to get higher scores. The results correspond to Deng's (1998) assertion that English listening comprehension questions should simply aim at evaluating students' comprehension in listening. Therefore, the task types adopted should avoid answers in written forms.

Table 4-9

Summary of One-Way ANOVA on Different Task Types

Source	SS	df	MS	F
Task Types	.425	3	.142	16.274*
(Error) Task Types	1.044	120	8.704E-03	

Note: *The mean difference is significant at the .05 level

Table 4-9 reveals a significant difference between each task type. A closer inspection of subjects' performance on different task types at different proficiency levels and the correlation between task types and the GEPT posttest provides additional information (see Tables 4-10 and 4-11).

4.3.2 Comparison of Mean Sores of Different Task Types with Proficiency Distribution

Table 4-10

The Means and the Standard Deviations with Proficiency Distribution in Different Task Types and the GEPT Posttest

Proficiency Levels / Task Types	High (n=17)		Middle (n=17)		Low (n=7)	
	Mean	SD	Mean	SD	Mean	SD
Blank-filling	.7241	.1065	.5524	.1627	.4457	.1086
True/False	.8012	.1093	.7076	.1295	.6429	7.387E-02
Multiple Choices	.7776	8.059E-02	.6841	9.747E-02	.5757	.1343
Cloze	.7718	.1169	.5741	.1212	.4957	7.955E-02
GEPT Posttest	.8329	7.235E-02	.6365	8.200E-02	.4643	8.162E-02

Note:1.The mean of the GEPT posttest is the average ratio of correctness.

2.The means of individual task types are the average ratio of correctness.

The results in Table 4-10 show that all the proficiency groups perform the best in true or false task type (High: mean= .8012, Middle: mean= .7076, Low: mean= .6429). In addition, the high-proficiency group scores higher than the other two groups in all the four task types, and the low-proficiency group performs obviously the worst among the three groups.

With respect to the performance on task types among individual proficiency levels, the high-achievers perform the best in the task type of true or false statements

(mean= .8012) while they score the lowest in blank-filling (mean= .7241). This is the same with the middle-achievers (true/false: mean= .7076, blank-filling: mean= .5524) and the low-achievers (true/false: mean= .6429, blank-filling: mean= .4457).

It is evident from Table 4-10 that individual differences do exist among these groups, constituting the difference in their performance. Nevertheless, task types on multiple choices and true or false statements, which do not require any spelling, yield better results.

This phenomenon can serve as a reminder for any teacher that, when training students on listening competence, it is better to avoid testing students on spelling, especially in the initial stage.

4.3.3 Pearson Correlation with Proficiency Distribution

The following table further indicates the correlation between different task types and the GEPT posttest to see which task type can better evaluate students' listening performance.

Table 4-11

Pearson Correlation between Task Types and GEPT Posttest with Proficiency Distribution

Proficiency Task Levels Types	High (n=17)		Middle (n=17)		Low (n=7)	
	Pearson r	Sig. (2-tailed)	Pearson r	Sig. (2-tailed)	Pearson r	Sig. (2-tailed)
GEPT Posttest - **Blank-filling**	.181	.487	.439	.078	-.179	.700
GEPT Posttest - **True/False**	.285	.268	.377	.135	.031	.948
GEPT Posttest - **Choices**	.447	.072	.745**	.001	-.241	.602
GEPT Posttest - **Cloze**	.429	.085	.562*	.019	.180	.699

Note: 1. ** Correlation is significant at the .01 level (2-tailed)

2. * Correlation is significant at the .05 level (2-tailed)

3. The mean of the GEPT posttest is the average ratio of correctness.

4. The means of individual task types are the average ratio of correctness.

As Table 4-11 shows, no significant correlation between individual task types and the GEPT posttest is found in the high-proficiency group and the low-proficiency group. However, with respect to the middle-achievers, the values of the Pearson correlation coefficients turn out to be positively related. The statistics show that a significant correlation is found between the GEPT posttest and multiple-

choice task type (Pearson's r = .745, p< .01) as well as cloze task type (Pearson's r = .562, p< .05). This means if one middle-achiever scores high on the task types of multiple choices or cloze, he/she would most probably perform well on the GEPT posttest, however, in which there is only one task type, multiple choices, in the listening part.

Results of Questionnaire

4.4 Analysis of the Subjects' Responses to the Feedback Questionnaire

This section presents the results and the percentile analysis of the subjects' responses to the feedback questionnaire, including (1) views on applying authentic broadcast materials as supplements for listening comprehension, (2) views on the design of task types consisting of blank-filling, cloze, true or false statements and multiple choices, and (3) an open-ended question for comments on the listening comprehension curriculum or suggestions for the following semester.

4.4.1 Views on Applying ICRT as Supplements for Listening Comprehension

Table 4-12

The Result of Percentile Analysis with All Subjects (Item 1~ 5)

Questions	Responses									
	1		2		3		4		5	
	N	%	N	%	N	%	N	%	N	%
Item 1	0	0	1	2	6	15	14	34	20	49
Item 2	0	0	1	2	4	10	19	46	17	41
Item 3	2	5	1	2	11	27	17	41	10	24
Item 4	0	0	1	2	10	24	15	37	15	37
Item 5	0	0	2	5	6	15	20	49	13	32

Note: Response "1"=strongly disagree; "2"= disagree; "3"=neutral; "4"=agree; "5"=strongly agree

Question items 1 to 5 are questions about the contents of the selected authentic broadcast materials. For question item 1, the intention is to determine whether

test-takers feel the listening materials are appropriate supplements for listening materials. Question item 2, 3, 4 ask the subjects if the selected authentic materials are in consonance with their life experiences, very interesting, and practical respectively. Question item 5 focuses on appropriateness of the length of each text.

As shown in Table 4-12, most of the subjects (34%+ 49%= 83%) agree that the authentic listening texts chosen from ICRT are appropriate listening materials as supplements. In item 2, a large percentage of the subjects (46%+ 41%= 87%) agree that the contents of the supplementary materials are related to their life experiences. As for item 3, also more than half of the subjects (41%+ 24%= 65%) approve that the supplementary materials are interesting while those who don't think the materials interest them are only 7% (5%+2%). This might attribute to their lack of confidence and competence in listening to English, making them feel nervous when listening to English. In item 4 and item 5, when put together the percentage of " strongly agree" and "agree", 74% (37%+37%) of the respondents thought the materials are practical and 81% (49%+32%) also agree that the length of each text is appropriate.

The following are Figures with percentile analyses concerning question items 1~ 5 with the distribution of proficiencies.

Question 1: *Are the selected broadcast suitable as listening comprehension*

supplements?

Figure 4.1

**Breakdown Responses for Item 1 by Proficiency Levels**
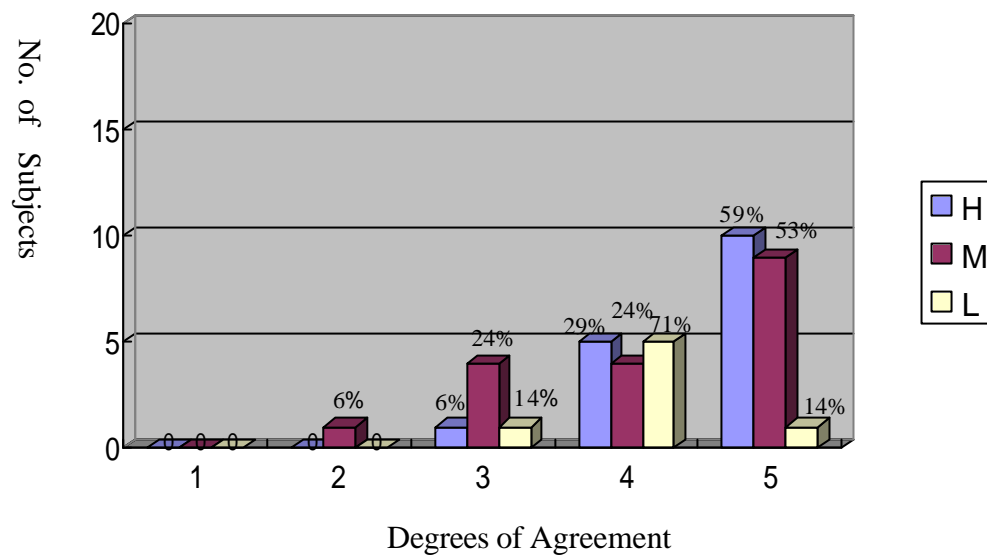
With regard to the distribution of proficiency, Figure 4.1 indicates that 88% (59%+29%) of the high-achievers and 77% (53%+24%) of the middle-achievers think the materials are suitable listening supplements, only 6% of the respondents in the middle-proficiency group do not think so. What's more, 85% (14%+71%) of the low-achievers respond that the authentic broadcast materials are suitable for them as listening comprehension supplements. In a word, from the subjects' responses, the selected authentic broadcast fits their current English language competence.

Question 2: *Are the supplementary materials related to subjects' life experiences?*

Figure 4.2

Breakdown Responses for Item 2 by Proficiency Levels
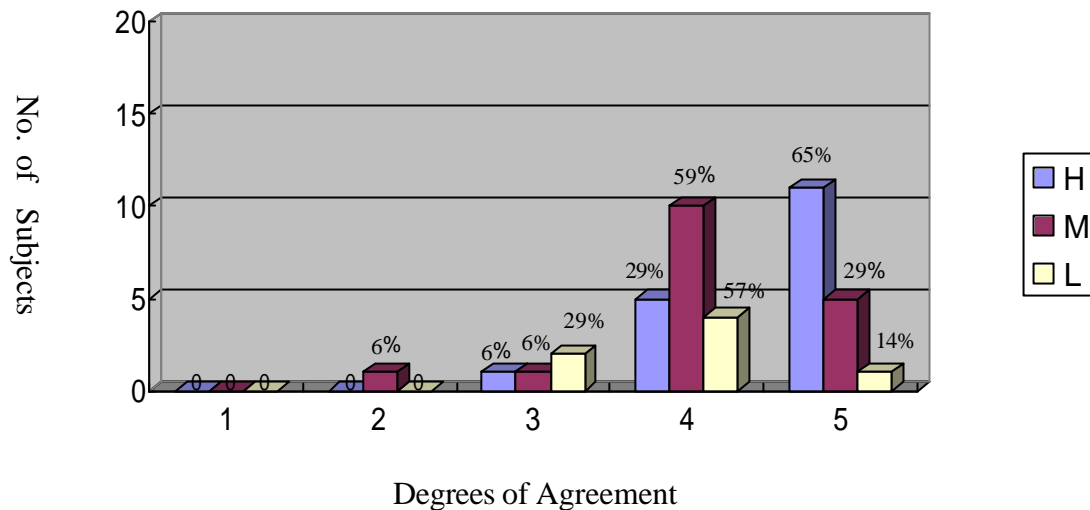
For question item 2, as seen in Figure 4.2, 87% of the respondents indicate that the authentic broadcast materials are related to their life experience (as mentioned earlier in Table 4-12), which is agreed strongly by the high-achievers (65%). In addition, more than half of the middle-achievers (29%+59%=88%) and low-achievers (14%+57%=71%) approve or strongly approve of the idea that the supplementary listening materials are in relation to their life experiences .

Question 3: *Are the topics interesting?*

Figure 4.3

Breakdown Responses for Item 3 by Proficiency Levels
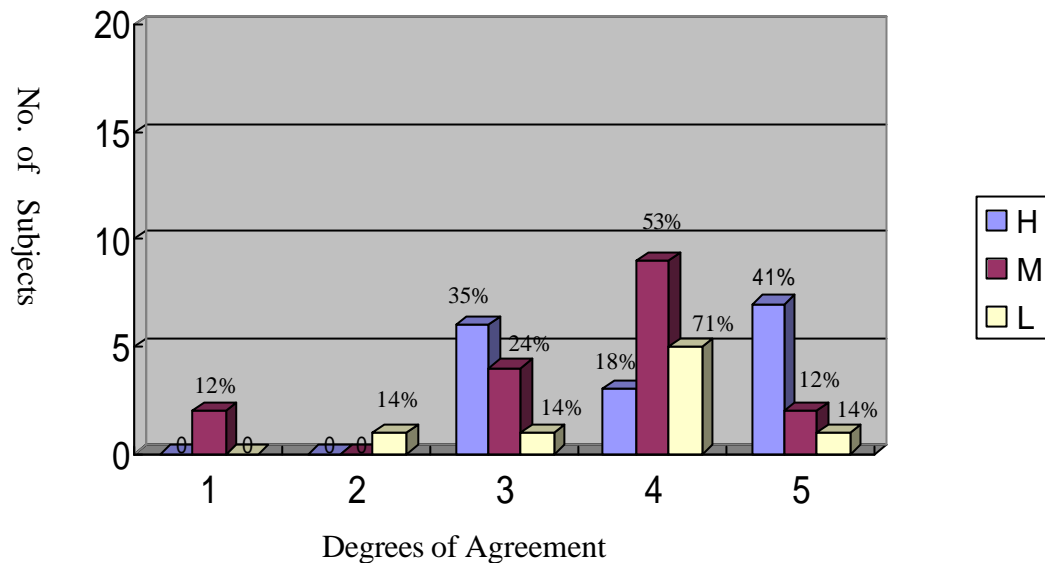
A breakdown of the responses to questions by proficiency levels in Figure 4.3 shows that 41% of the respondents in the high-achievers feel the materials are quite interesting, and 53% of the middle-achievers and 71% of the low-achievers also favor the materials. Those who don't think the materials interest them are among the respondents in the middle-proficiency level (12%) and in the low (14%).

Question 4: *Are the topics discussed very practical?*

Figure 4.4

Breakdown Responses for Item 4 by Proficiency Levels
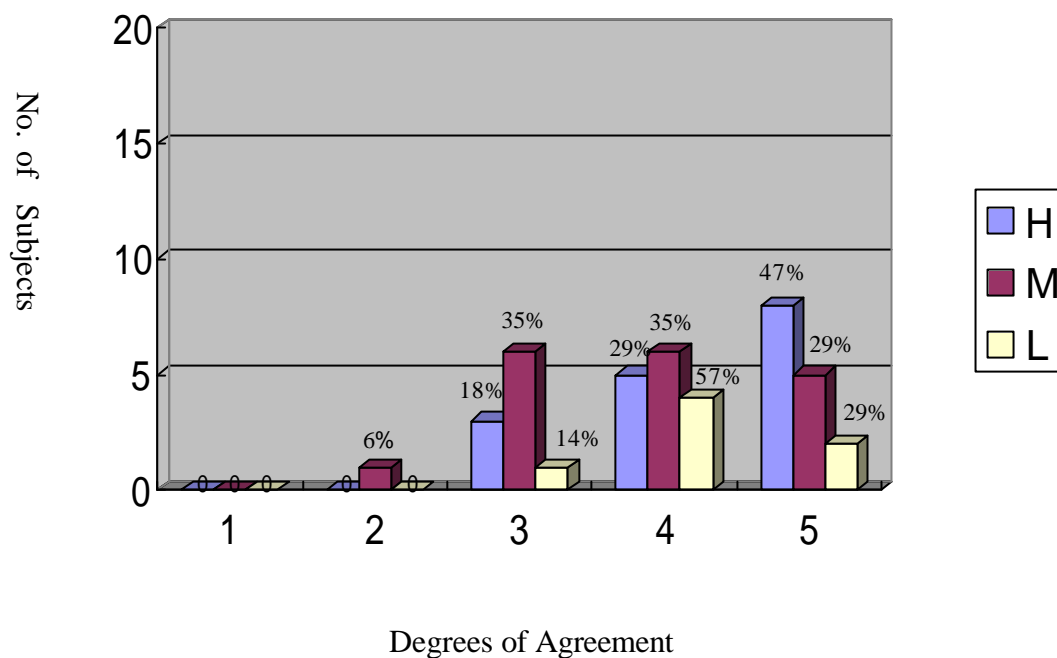
As to the practicability of the authentic broadcast materials, it is obvious to find in Figure 4.4 that most of the respondents in each level agree the materials are practical. Among those who strongly agree or agree, the ratio of the high-proficiency group is 76% (47%+29%), the middle-proficiency group 64% (29%+35%), and the highest ratio is in the low proficiency level (29%+57%=86%). Hence, from the responses, the materials extracted from broadcast are proved to be practical.

Question 5: *Is the length of passages and dialogues just appropriate for its intended purpose?*

Figure 4.5

Breakdown Responses for Item 5 by Proficiency Levels
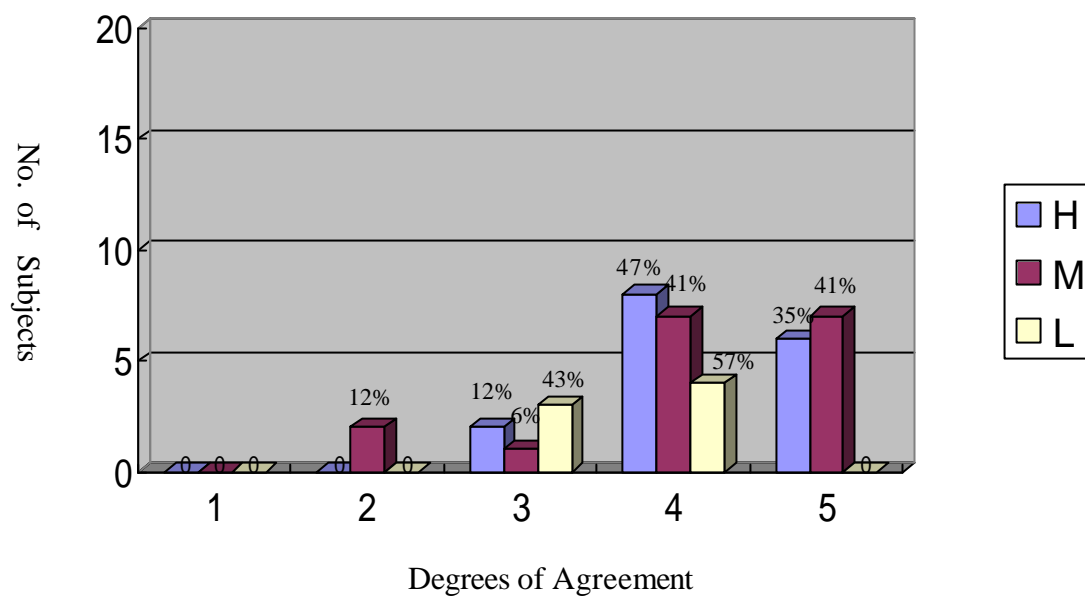
As shown in Figure 4.5, 41% of the middle-proficiency group, which is the highest ratio, and 35% of the high-proficiency group strongly agree that the length is appropriate but none of the low-achievers strongly support this idea. However, more than half of the low-achievers (57%) feel the length of the given texts is appropriate.

From Figures 4.1 to 4.5, the researcher have found that the authentic broadcast materials, which are carefully selected to cater to students' life experiences and interest and systematically designed by the teacher with appropriate length, are not only suitable but also practical as listening comprehension supplements for the subjects.

Table 4-13

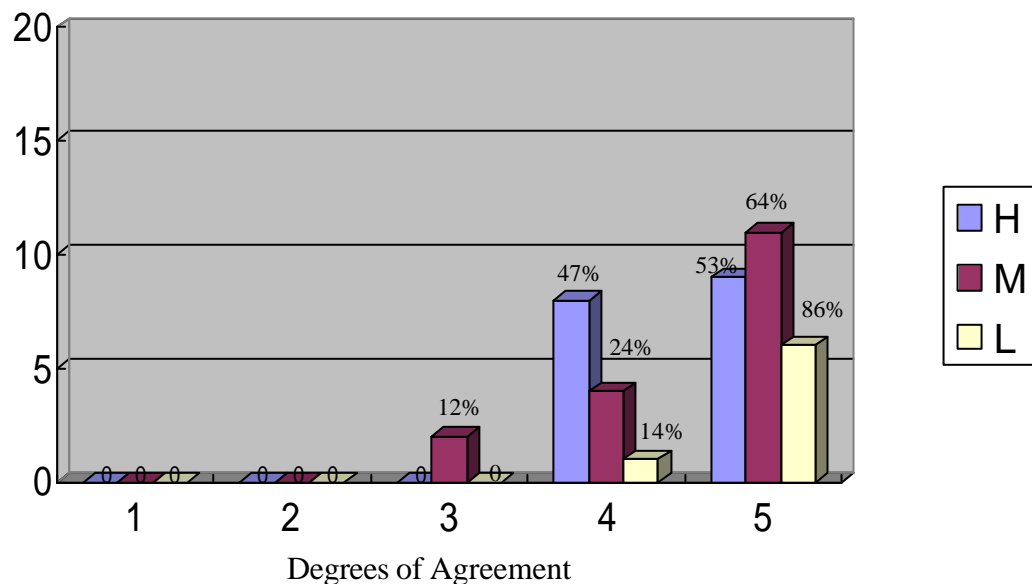**The Result of Percentile Analysis with All Subjects (Item 6~ 7)**

Questions	Responses									
	1		2		3		4		5	
	N	%	N	%	N	%	N	%	N	%
Item 6	0	0	0	0	3	7	13	32	25	61
Item 7	0	0	3	7	8	20	16	39	14	34

Note: Response "1"=strongly disagree; "2"= disagree; "3"=neutral; "4"=agree;

"5"=strongly agree

Question 6: *Are the authentic listening materials challenging?*

Figure 4.6

Breakdown Responses for Item 6 by Proficiency Levels

From table 4-13, a large percentage of the subjects (61%+32%=93) agree that the authentic listening materials are challenging. A breakdown of the responses of

proficiency levels (Figure 4.6) indicates that a high ratio (86%) of the low-achievers strongly feel the materials are challenging. This is very possible with their low competence in English. When the responses "strongly agree" and "agree" are added up, a marked ratio of the high-proficiency level (100%), and 88% (64%+24%) of the middle-proficiency level deem the materials are challenging. The results reveal that the supplementary authentic listening materials, interesting and related to students' life experiences, could also challenge subjects with all levels.

Further analyses of the assistance of the authentic broadcast supplements are presented in the following Tables and Figures.

Question 7: *Do the supplementary materials contribute to building up one's vocabulary?*

Figure 4.7

Breakdown Responses for Item 7 by Proficiency Levels
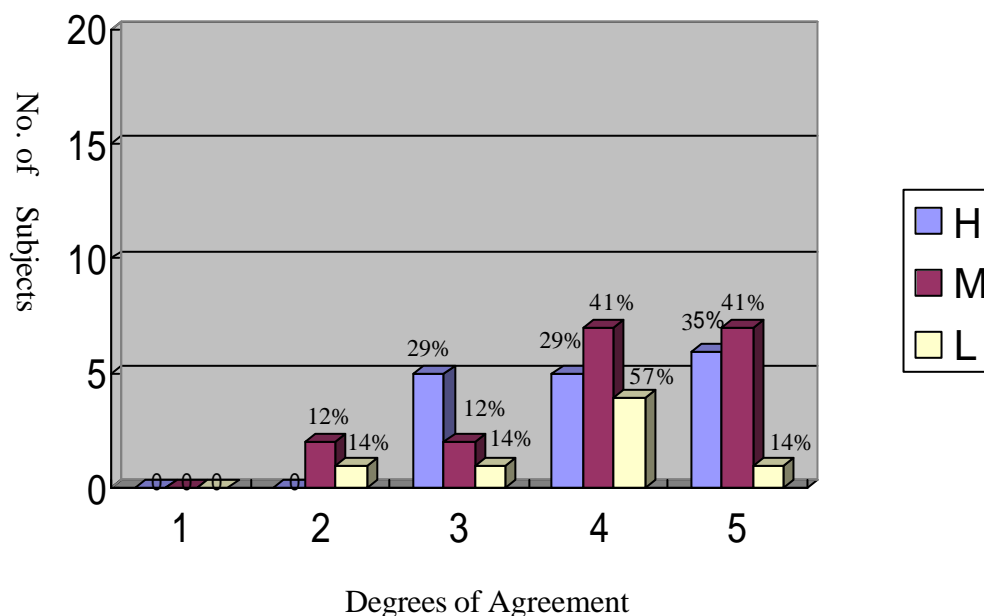
As shown earlier in Table 4-13, 73% (34%+39%) of the subjects support that

their vocabulary have increased during the ten weeks' systematic curriculum. Furthermore, when considering the proficiency levels, most of the middle-achievers (41%+41%=82%) agree that they have learned a lot of new words in the process of the authentic listening curriculum. Even for the low-achievers, there is a sign of increase in their vocabulary (14% with strong agreement ; 57% with agreement). As for the high-achievers, while more than half of them (64%) indicate that they have increased vocabulary, still 29% of the high-achievers hold the neutral opinion. Such attitude could be the result of the fact that they have higher competence of English and that they are familiar with most of the topics.

In a word, the finding indicates that most of the subjects in different proficiencies have enhanced their vocabulary in the process of the listening curriculum.

Table 4-14

The Result of Percentile Analysis with All Subjects (Item 8~ 9)

Questions	Responses									
	1		2		3		4		5	
	N	%	N	%	N	%	N	%	N	%
Item 8	0	0	1	2	6	15	14	34	20	49
Item 9	0	0	1	2	3	7	18	44	19	46

Note: Response "1"=strongly disagree; "2"= disagree; "3"=neutral; "4"=agree;

"5"=strongly agree

Questions item 8 and 9 aim to investigate subjects' opinions on the effects of the teacher's hints on vocabulary and phrases as well as the instruction on the features of

authentic natural speech, such as pitch, tone, intonation and linkage, which are some of the factors that affected the subjects' listening comprehension. (Chen, 2002, p.73)

As shown in Table 4-14, a large percentage of the subjects (34%+49%= 83%) indicate that the hints on vocabulary and phrases provided on the quiz sheet by the teacher help their better understanding of the context. Furthermore, the view of question 8 is repeated across high-, middle-, and low-proficiency levels, as shown in Figure 4.8.

Question 8: *Do hints by the teacher on vocabulary and phrases contribute to better understanding of the context?*

Figure 4.8

**Breakdown Responses for Item 8 by Proficiency Levels**
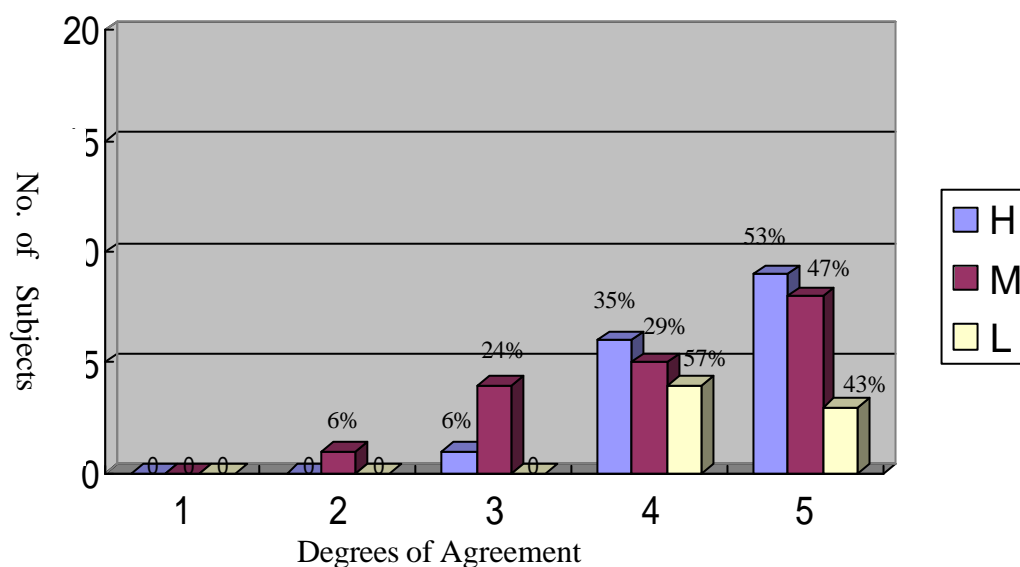
From Figure 4.8, more than half of the high-achievers (53%) vigorously support teachers' hints on vocabulary and useful phrases facilitate their listening comprehension. When put up the responses of "strongly agree" and "agree", 100% (43%+57%) of the respondents in the low-proficiency group point out that teachers'

hints on vocabulary or phrases either on the quiz sheet or by oral do help them better comprehend the authentic broadcast. It is obvious that students with low English competence need more instruction on vocabulary and phrases while listening to authentic broadcast since their vocabulary is usually limited.

Question 9: *Do language features such as pitch, tone, intonation, linkage given to students by the teachers contribute to better understanding of the context?*

Figure 4.9

Breakdown Responses for Item 9 by Proficiency Levels
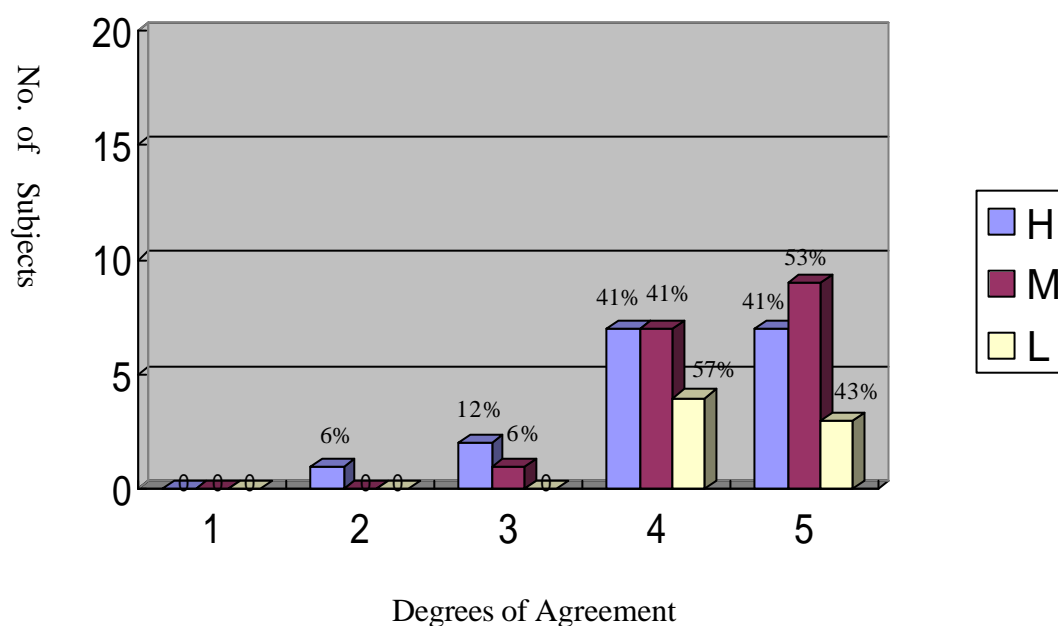
Some researchers (Li, 1999; Chien, 1999; Hsiung, 2002) have pointed out that the problems EFL/ESL students encountered in listening comprehension are the language features, such as stress, intonation, pitch, linking of sound, etc. Question item 9 aims to investigate the effect of the instruction on the language features when students listen to authentic broadcast. As shown in Table 4-14, a large percentage of the respondents (90%= 44%+46%) agree that the instructions of language features by

the teacher assist them in the process of listening to authentic speech. As seen in Figure 4.9, the instruction of language features is especially effective for the low-achievers (43%+57%=100%) and middle-achievers (53%+41%=94%).

Table 4-15

The Result of Percentile Analysis with All Subjects (Item 10~ 11)

Questions	Responses									
	1		2		3		4		5	
	N	%	N	%	N	%	N	%	N	%
Item 10	0	0	0	0	4	10	24	59	13	32
Item 11	0	0	0	0	13	32	23	56	5	12

Note: Response "1"=strongly disagree; "2"= disagree; "3"=neutral; "4"=agree;

"5"=strongly agree

Question item 10 is intended to investigate whether students pay attention to the authentic listening materials or not in the process of listening comprehension. Observed from Table 4-15, a large percentage of subjects (32%+59%=91%) do pay attention in the listening process. Figure 4.10 shows further indication.

Question 10: *Do students pay concentration when doing the series of listening*

comprehension from ICRT broadcasted programs?

Figure 4.10

**Breakdown Responses for Item 10 by Proficiency Levels**
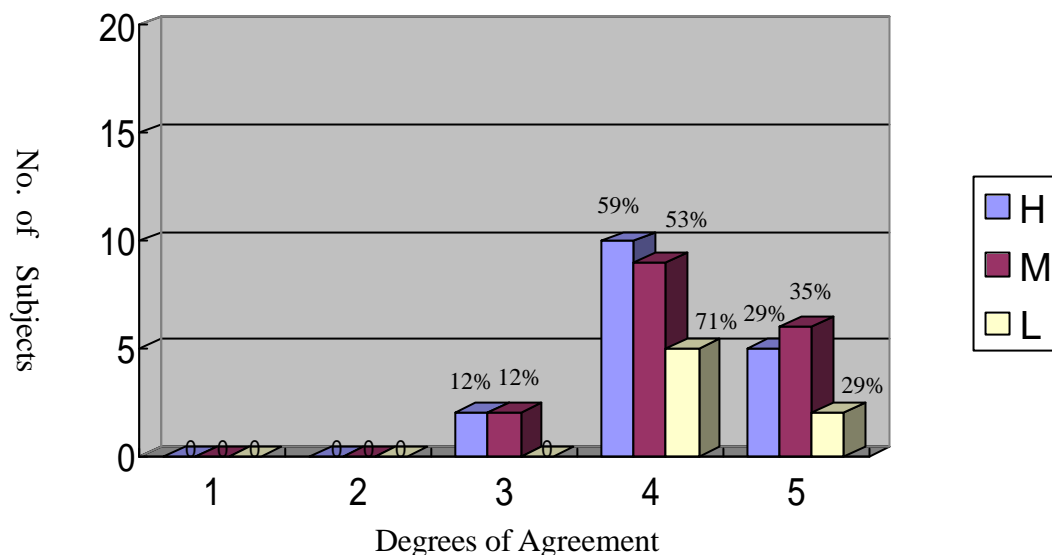
As shown in Figure 4.10, all the low-achievers pay more concentration in the process of listening to authentic broadcast than the other two levels. In addition, both the high-achievers and middle-achievers also concentrate on the listening with a marked ratio of 88%. Generally speaking, students would pay more concentration on authentic broadcast materials which interest them, otherwise they would miss a lot of the broadcast. Moreover, whether they concentrate or not during the listening process would affect their comprehension of the context.

Question 11: *Do subjects attain improvement in listening ability after listening to the series of listening materials extracted from ICRT programs?*

Figure 4.11

Breakdown Responses for Item 11 by Proficiency Levels
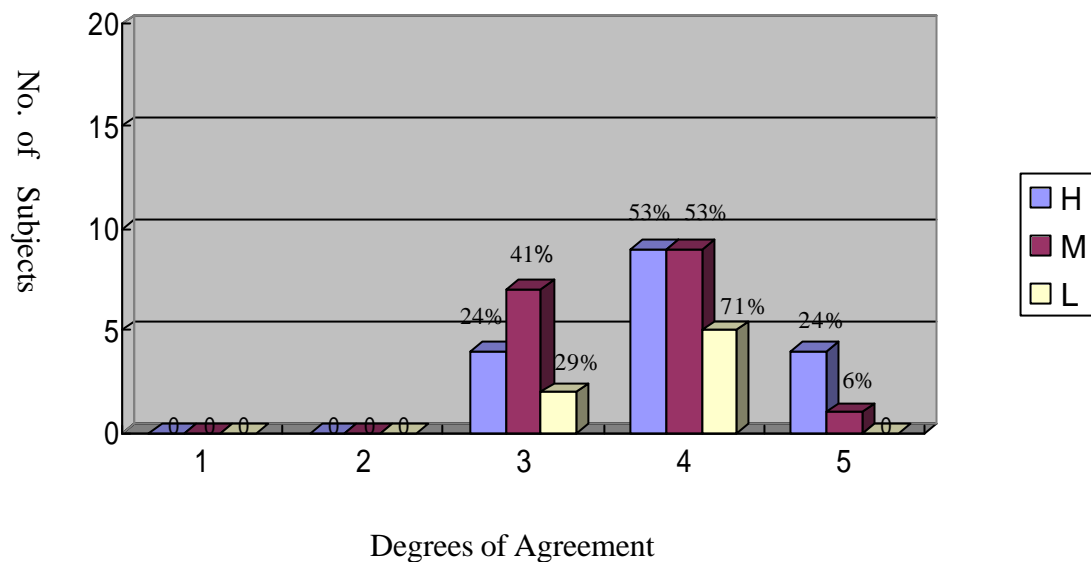

As Table 4-15 shows, 68% of the subjects indicate that their listening comprehension improves after being exposed to authentic broadcast environment for ten weeks. With respect to distribution of proficiency, the high-achievers (53%+24%=77%) improve more than the middle-achievers (53%+6%=59%) and low-achievers (71%). The result could be in correspondence to their performance on the GEPT posttest (see Table 4-2).

Question 12: *Which of the following topics give you most a sense of*

accomplishment?

Figure 4.12

Breakdown Responses for Item 12 by Proficiency Levels
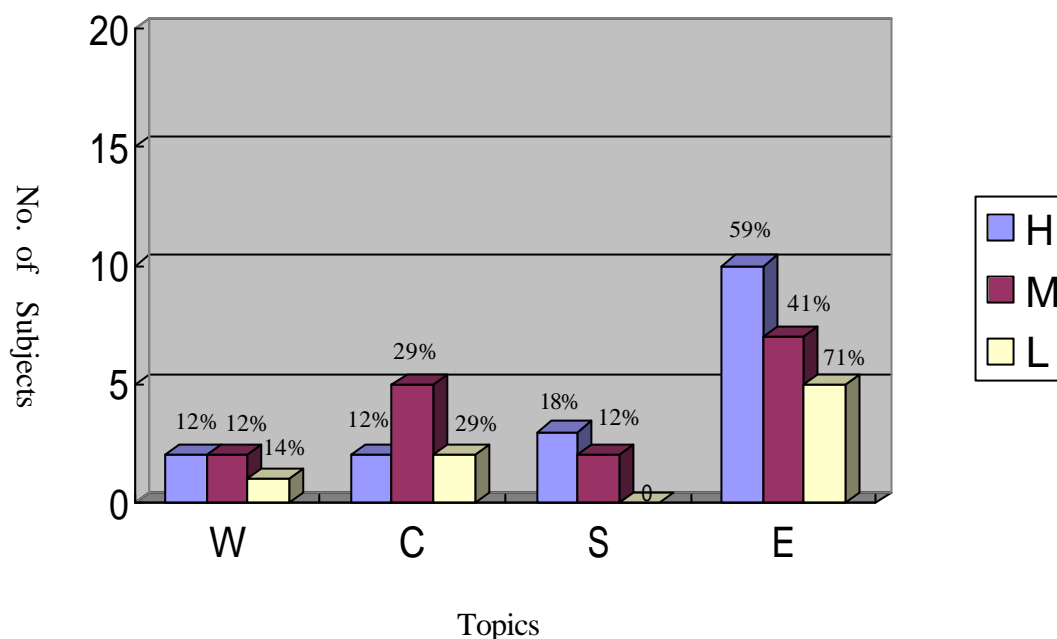
Note: Topics "W" =Weather Forecast; "C" = Commercials; "S"= Special Topics

"E"= Easy News

Questionnaire item 12 aims to investigate which authentic topic category the subjects feel confident in. The results in Figure 4.12 point out that a large percentage of the subjects in each proficiency group (H:59%, M:41%, L:71%) deem that they have more confidence in listening to the authentic texts extracted from Easy News. As for the other topic categories, Special Topics (18%) is the secondary ratio for the high-achievers, whereas the secondary ratio for the middle-achievers and the low-achievers is commercials (M: 29%, L: 29%). The findings are a little different from the results of subjects' performance on the quizzes. As indicated earlier in Table 4-6, high-achievers do best on commercials, followed by Easy News;

middle-achievers do best on commercials, followed by Easy News, while low-achievers do best on Easy News, followed by commercials. Since it is near the end of the present research when the materials of Easy News are applied, the possible reasons for students' feeling confident in listening to Easy News may probably because they are gradually adapted to listening to authentic broadcast.

4.4.2 Views on the Design of Task Types

The second part of the feedback questionnaire is intended to understand subjects' opinions on the design of different question types based on the same topic. They are asked to arrange their answers in the order of preference by numbers 1 to 4[5]. The results of the subjects' responses with different proficiency groups are demonstrated in the following Table 4-16.

---

[5] Task type "1" represented cloze test; "2", true or false statements; "3", multiple choices; "4", blank-filling.

Table 4-16

Percentile[6] Analysis of the Subjects' Responses

to the Design of Task Types with Proficiency Distribution

Questions	Task Types	Proficiencies					
		H		M		L	
		N	%	N	%	N	%
1.The type of test questions that truly helps me understand the article.	1	12	71	6	35	2	29
	2	1	12	3	18	2	29
	3	4	24	8	47	3	43
	4	0	0	0	0	0	0
2.The type of test questions that most gives me a sense of accomplishment.	1	10	59	4	24	2	29
	2	2	12	4	24	2	29
	3	4	24	8	47	3	43
	4	1	12	1	12	0	0
3.The type of test questions that most helps me concentrate on the content of text.	1	13	76	6	35	2	29
	2	2	12	2	12	1	14
	3	2	12	9	53	4	57
	4	0	0	0	0	0	0

Note: 1.Only the responses of students' first preference in each category are tabulated.

2. Task type "1" represented cloze test; "2", true or false statements; "3", multiple choices; "4", blank-filling.

---

[6] The ratio is obtained from the comparison within the same proficiency group.

As shown in Table 4-16, a large percentage of the respondents in the high proficiency group point out that cloze tests aid them in better understanding the authentic listening text (71%) and give them a sense of accomplishment (59%) because they could see the words while listening. The reason could be explained by Don & Jon's research (1981), which indicates that a transcription "is very helpful in aiding segmentation of the stream of sound and recognition of weak forms or other features of reduction"(p.186). However, when compared with Table 4-10 (means of different task types), the high-achievers score the highest in the task type of true or false statement, followed by multiple choices and then cloze.

On the other hand, most of the middle-achievers and low-achievers deem that they could concentrate more (M: 53%; L: 57%), understand better the listening text (M: 47%; L: 43%) and gain more confidence (M: 47%; L: 43%) when listening to authentic broadcast with the task type of multiple choices. The result, either, do not conform to those of Table 4-10, which shows both the middle-achievers and low-achievers perform the best on the task type of true or false statement, followed by multiple choices and then cloze. The reason why they differ from the high-achievers in their responses is probably due to their lack in confidence or competence in listening with writing. The results are in accordance to the investigation of Chen's (2003) that "once writing is involved in a listening task, the task itself becomes more difficult"(p.14).

4.4.3 Responses of the Open-ended Question

An open-ended question is provided at the end of the questionnaire, pertaining to the free opinions or suggestions from the subjects on the listening curriculum. Since there is such a wide range of their answers on this part, the responses are tabulated in Table 4-17 according to their frequencies.

Table 4-17

Descriptions of Subjects' Responses
1.The authentic listening curriculum is so wonderful that we'd like it to be continued next semester.
2.The authentic listening curriculum is interesting and down to earth.
3.I do enlarge my vocabulary, which is crucial in listening comprehension.
4.The listening curriculum provides me with a good chance to get used to the fast pace of authentic speech.
5.Listening to ICRT is difficult to me, but it does enhance my English proficiency.
6.By practicing listening to the authentic broadcast, I feel the exams based on text books are a lot easier because I don't feel that nervous as before.
7. I like the curriculum very much but the class time for the listening activity is not sufficient.
8. After the training, I make it a rule to tune into ICRT almost every day.
9.The curriculum is systematic and well-organized.
10.Most of the time I don't quite understand the content of the materials, but when I could catch what it means, I feel proud of myself because I could understand "real" English.

Note: Students were asked to answer this part in Chinese to ensure they can fully
      express their ideas.

It is obviously seen from Table 4-17 that most of the subjects hope to continue the authentic listening activity for the following semester. In addition, most of them deem that they have improved in their listening competence, including adapting themselves to the fast speed of authentic speech, increasing their vocabulary, and hence the authentic materials aid them to perform better on regular tests based on text-books. Furthermore, some would spontaneously tune into ICRT to polish up their listening skills, but regret for the limited time of conducting the curriculum. In a word, the above Table once again proves the effectiveness of applying authentic broadcast as a listening supplement for EFL senior high school students.

4.5 Conclusion

In this chapter, the result of the data analysis has been interpreted first by statistical analysis, focusing on (1) comparison of the GEPT pretest and posttest, (2) correlation between selected texts and the GEPT posttest, and (3) correlation between task types and the GEPT posttest. In addition, the analysis of the feedback questionnaire has been interpreted by percentile from three aspects: (1) opinions on selection and application of the authentic broadcast materials, (2) opinions on the design of different task types, and (3) open-ended questions on the listening curriculum. In conclusion, the results show that the selected broadcast materials and the designed task types place a high premium on the development of students' listening competence.

First of all, with the systematic introduction of the materials and designed activities provided to the students, it is evident from the above statistical results that the students as a whole have improved significantly.

This is done possible because all the selection of materials are properly designed

to meet the students' needs and expected results—that is the materials are closely related to their life experience and in conformity with their interest and competence in the English language. This is coupled with a very systematic presentation of selected texts, which are fitted as supplements for training senior high school students on listening comprehension. In addition, such listening supplements would not only help students improve their listening competences but also enlarge their vocabulary.

Although the correlation between the task types and the GEPT posttest doesn't show much significance, the feedback from different students shows that different task types provide them with different stimulation. Nevertheless, in general the subjects believe that familiarity with the test questions helps them obtain self-confidence, score better, and motivate their interest to keep on improving their listening abilities.