The Effects of Inference-Training and Text Repetition on Chinese Learners' Incidental Vocabulary Acquisition While Listening

Rod ELLIS
University of Auckland, New Zealand

CHANG Le
Bohai University

Abstract

This article reports a study of the effects of inference training and text repetition on Chinese university students' performance of two listening information-transfer tasks that provided built-in measures of their comprehension and opportunities for the acquisition of ten unknown target words embedded in the listening texts. One group just listened to the text once, a second group three times, while the third listened three times and received inference-training support. The results showed that text repetition had a positive effect on both comprehension and vocabulary acquisition. However, the inference-training had no additional effect. The study lends support to the findings of previous studies which have shown that repeated opportunities to process oral input have a positive effect on listening comprehension and extends these studies by showing that it also facilitates incidental vocabulary acquisition.

Key words: text-repetition; inference-training; listening comprehension; vocabulary acquisition

Introduction

Listening activities potentially serve a dual purpose – to provide practice in helping second language (L2) learners comprehend oral texts and thereby developing their listening comprehension ability and to provide exposure to L2 input that can serve as data for linguistic development. Thus there are two potential outcomes from such activities

– the development of listening comprehension skills and language acquisition. By and large, however, these outcomes have been discussed and researched separately. The study reported in this article examines both.

Listening to comprehend and listening to learn

The processes involved in listening to comprehend and listening to learn are different, although, clearly, when learners listen to input they can potentially both comprehend it and acquire new linguistic forms and strengthen their control over partially acquired forms.

Central to any model of listening-to-comprehend is the distinction between bottom-up and top-down processing, which Field (2004) defined in this way:

> In a bottom-up process, small ('lower level') units are progressively reshaped into larger ones; in a top-down process, larger units exercise an influence over the way in which smaller units are perceived. (p. 364)

As a result of the difficulty that L2 learners, especially those of lower proficiency, experience with bottom-up processing (i.e., phoneme and word recognition in the input stream) – see Goh (2008) – L2 learners are likely to rely heavily on top-down processing by drawing on their content knowledge of a topic and on available contextual clues.

In a context where learners are primarily focused on comprehending input and where this input comes at them in a continuous stream (as in a typical listening comprehension activity) there will be limited opportunities to perceive the form and understand the meanings of unfamiliar words. If learners rely primarily on top-down processing, the chances are further reduced. Nevertheless, given that some bottom-up processing must also take place, some opportunities arise for noticing unfamiliar words and inferring their meanings with the help of the co-text and context will arise. Thus, while incidental acquisition of new words from oral input is likely to be limited, it can take place and indeed there is clear evidence that it does (see below).

Inferencing-training

Researchers have investigated a number of different types of listening support – topic preparation (Alavi & Jambaz, 2014; Sarandi, 2010), visual support (Chang & Reid, 2007), pre-viewing sentences from the listening text (Jafari & Hashim, 2012), pre-viewing the comprehension questions (Alavi & Jambaz, 2014; Berne , 1995; Chang & Reid, 2008; Elkhafai, 2005), vocabulary preparation (Berne, 1995; Chang, 2007; Elkhafai, 2005; Jafari & Hashim, 2012). Of these different types of support, some are aimed at facilitating top-down processing (e.g., topic preparation and visual support), some selective listening (e.g., pre-viewing the comprehension questions) and some bottom-up processing (e.g., pre-viewing sentences from the listening text and vocabulary preparation). In all of the studies

the support was provided prior to listening.

Interestingly, none of these studies investigated the effect of strategy training on listening comprehension although there is clear support for this in the pedagogic literature. Evidence showing that strategy training improves listening comes from studies that show that more proficient listeners make use of a broader selection of strategies especially meta-cognitive strategies such as selective listening (Vandergrift, 2005). Experimental studies have typically been longitudinal comparing a group that received strategy training with one that did not, with some reporting a positive effect for strategy training (e.g., Coskun, 2010). However, the overall effectiveness of strategy-training for listening remains doubtful. Plonsky (2011), in a general meta-analysis of strategy-instruction studies, reported that the effect on listening was almost non-existent (i.e., the effect size was very small: d = .06).

The study reported below investigated a particular type of strategy training of likely benefit not just to comprehension but also vocabulary acquisition – inference-training. Nation (2001) noted that "inferring vocabulary meaning from context … is an essential strategy for promoting reading comprehension and promoting lexical acquisition" (p. 240) and went on to argue that training in making inferences was desirable. However, the results of studies that have investigated inference-training for reading are not encouraging. Fraser (1999) for example, provided direct instruction and contextualized practice in the use of a range of "lexical processing strategies" (e.g., using cognates, word structure, grammatical function and structural redundancy) but found it had no direct effect on vocabulary learning. If it is not very effective for reading, it is unlikely to be so for listening, although there is evidence that L2 learners do engage in it. Vandergrift (2003), for example, reported that the learners he investigated used a variety of knowledge sources to infer the meanings of unfamiliar words while listening. Wei and Wu (undated) found that Chinese English majors differed in how they went about inferencing depending on their proficiency, with low-proficiency students resorting to their world knowledge and high-proficiency students drawing more on their linguistic knowledge.

Text-Repetition

Giving learners the opportunity to listen to a text several times can potentially aid both comprehension and vocabulary acquisition. Learners have the opportunity to apply a variety of strategies to fill in gaps in their comprehension when they listen a second or third time. Also, text repetition increases exposure to unknown words which may help build form recognition and form-meaning mapping. Several of the studies that investigated the effect of pre-listening support also included text repetition as a listening condition (Berne, 1995; Chang & Reid, 2007, 2008), reporting that it proved more effective than pre-listening support. Berne for example concluded "the most effective means of improving listening comprehension performance is through additional exposure to the passage" (p. 326). O'Bryan (2010) also found that text repetition assists listening comprehension.

Incidental vocabulary learning while listening

As Van Zeeland and Schmitt (2013) pointed out, relatively few studies have investigated incidental vocabulary acquisition through listening. In a number of task-based studies (e.g., Ellis & He, 1999; Ellis, Tanaka & Yamazaki, 1994; Loschky, 1994), learners listened to input containing unknown words and demonstrated their understanding of the input non-verbally. These studies reported that the learner-participants were able to acquire some new words when the input was pre-modified (i.e., simplified). For example, in Ellis, Tanaka and Yamazaki, a group of high school Japanese students demonstrated receptive knowledge of 14% of the target words and maintained this level over time. In Ellis and He, a group of adult ESL learners in the US demonstrated receptive knowledge of 62% of the target words and productive knowledge of 56% and again maintained these levels in delayed post-tests. The difference in the results of these studies can be explained by the difference in the learners' L2 proficiency. In both cases, however, the learners' proficiency was notably higher than in Brown, Waring and Donkaewbua's (2008) study of extensive listening where their participants (Japanese university students) acquired only 2% of the target words in an immediate test and virtually none in a 3 month delayed test.

Other studies have sought to identify the properties of an oral text that predict the incidental acquisition of vocabulary. These properties relate to both the general characteristics of the listening text (e.g., speed; linguistic complexity) and the characteristics of the words targeted for acquisition (e.g., their frequency, form and type; and also the degree and type of elaboration in the listening texts). Vidal (2003) reported that "explicitness" (i.e., whether the text includes explicit clues about the meanings of the target words) assisted vocabulary acquisition, but Revesz and Brunfaut (2013) found no such effect. These mixed findings probably can be explained by the fact various properties of a text interact to determine difficulty so the effect of a single factor will necessarily depend on other properties and also the differences in the learners' proficiency. Given that input frequency has been shown to be a major determinant of acquisition (N. Ellis, 2002), we might expect that the frequency of the target words in a listening text will be related to learning. However, surprisingly several studies (Revesz & Brunfaut, 2013; Van Zeeland & Schmitt, 2013; Vidal, 2003, 2011) found no or only a weak relationship between target word frequency and acquisition.

Research questions

To maximize the chances of inference-training having an effect on incidental vocabulary acquisition, the study investigated it in combination with text repetition. The assumption was made that learners would be better equipped to make use of the training if they had the opportunity to apply it in multiple listenings.

The following research questions were formulated:

1. Does text repetition help the incidental acquisition of vocabulary embedded in a listening

comprehension text?

2. Does inferencing-training combined with text repetition enhance the incidental acquisition of vocabulary embedded in a listening comprehension text?

Method

Participants

The participants were first year, second-semester students in a Chinese university in north-eastern China. They were aged 17 to 19 and 70% were male. These students had received at least six years of formal English instruction at high school and were pre-intermediate to intermediate in level. They were placed in four intact first-year listening classes. An examination of the listening scores from the previous semester's final examination showed there were no differences in the listening ability of the students in the three classes. The students were all enrolled in a scheduled listening course which followed the same syllabus and used the same materials.

Design

A pilot study was conducted on a different but comparable set of students to identify a set of target words that the students were unlikely to know. This was then checked with the participants in the main study by means of a self-report vocabulary test (i.e., Wesche & Paribakht's (1996) Vocabulary Knowledge Scale). The target words that were new to the participants were then embedded into two information-transfer listening tasks. Three intact classes were used. One class listened one time and the other two classes groups three times. In addition, the third class received 10 minutes inference-training before each task. Listening comprehension scores were obtained by inspecting the completed chart/table that were part of the listening tasks. Both tasks were completed in the same lesson. Immediately after the students had finished the second task, they completed three vocabulary tests (a Form Test, a Reception Test and a Production Test). One week later they took the same tests.

Listening tasks

The information-transfer tasks consisted of a text (specially written) and a table or chart to be completed as they listened. In one task – *Drawing a Sales Line on a Company's Yearly Sales Chart* – students completed a chart. The text was 455 words long. In the other task – *Completing a Conference Registration Form* – students completed a registration form. The text was 475 words long. There were five information items to fill in for each task. The students' listening comprehension was measured by scoring whether they had inserted the correct pieces of information in the chart in the first activity and in the table in the second. The maximum possible comprehension score was 10 for the two tasks.

Five target words (all nouns or adjectives), all not reported as known by the learners in a pre-test, were embedded in the text for each task but they did not appear in the table/chart the students completed:

Text 1: abrupt, competitive, decline, domestic, optimistic
Text 2: accommodation, architecture, available, linguistics, permanent

Each word occurred twice in a context that would provide an implicit clue to its meaning. For example, one of the contexts for *optimistic* was:

At the start of the year I felt very *optimistic* for our company. I felt we had a good chance of having an excellent year.

Listening conditions

Each group performed the two listening activities under different listening conditions:

Group A: The students heard the text once only and were asked to complete the chart/table while listening. They received no listening support.

Group B: The students heard the text three times. The first time they just listened. The second time they completed the chart/table. The third time they made any changes they wanted to the chart/table. No pre-listening support was provided.

Group C: This group also listened to each text three times and also had an inference-training session. First, they were given a simple definition of inferencing (i.e., *inferencing means guessing the meaning of a word you do not know.*). Then two common ways of how to infer the meanings of words from context were explained with the help of example sentences. "Using your general knowledge" was explained using sentences such as "Smoking is the major cause of bronchitis". "Using the context" was explained using sentences such as "There are usually two kinds of natural disasters in Northern Australia: in the wet season, the heavy rains cause flooding, whereas in the dry season farmers face the problem of drought". Students' attention was drawn to key lexical items (e.g., "wet season" and "dry season") that would help them infer the meaning of a word (in this case "drought"). The students then practiced inferring the meanings of words in a set of sentences. The training took approximately 10 minutes before each task.

Vocabulary tests

There were three tests that students completed after they had finished the listening comprehension tasks in this order:

1. The Form Test
Each target word was placed in a list along with four other words some of which were similar in spelling. The students were asked to circle the word in each list that they

remembered hearing when listening to the texts. They were also asked to indicate how certain they were about their answers by entering a percentage in a box after each word list. Here is an example:

optional – statistic – optimistic – mysterious – opportunity ☐

One point was awarded for each correctly circled word giving a total possible score of 10.

2. The Reception Test

This was based on Reid's (2000) Matching Items Test. It tested whether the students could recognize the meanings of the target words by matching the word in the left hand column with its definition in the right hand column as in this example:

1. complicated
2. chemical _____ exceptional; higher
3. optimistic _____ difficult and complex
4. advanced _____ expecting good things
5. stable

Students scored a point if they wrote in the number of a target word against the correct definition (i.e., number 3 in the above example). The maximum possible score was 10.

3. The Production Test

This was a cued recall test. Students were shown sentences from the listening texts with the target words blanked out and were asked to write in the missing words as in this example:

But we ended with stronger sales than we expected and I am very _____ for next year.

A point was scored when a student filled in the exact word from the listening text. The total possible score was 10.

Each completed test was collected before the learners took the next test.

Analysis

Descriptive statistics for both the listening comprehension and the vocabulary tests' scores were presented. As the vocabulary scores were not normally distributed, the Kruksil Wallis test was used to establish whether differences in the groups' vocabulary scores were statistically significant and, where appropriate the Mann-Whitney test with Bonferroni correction was used as a post-hoc test to compare pairs of groups. Within group differences between the immediate and delayed tests were examined by means of Wilcoxon Signed Ranks Tests again with Bonferroni correction. Effect sizes were also calculated for both between group and within-group differences using Cohen's *r*. The interpretation was made using Cohen's (1988) benchmarks (i.e., small = .10, medium = .30, and large = .50).

Results

Listening comprehension

Table 1 shows the mean, standard deviation and range of each group's listening comprehension scores. The mean scores for Group A (one time listening) were notably lower than those for the other two groups, both of which listened three times.

Table 1. Descriptive Statistics for Listening Comprehension Scores

Group	Mean	SD	Range
A (N =45)	4.87	2.05	2-10
B (N = 44)	8.09	1.60	3-10
C (N = 41)	7.68	1.65	3-10
Total (N = 130)	6.88	1.77	2-10

Note: A = listening one time; B = listening three times; C = inference training + listening three times

Vocabulary acquisition

Table 2 provides the descriptive statistics for vocabulary scores on the immediate and delayed Form Test. The means of the two groups that listened three times were both greater than the mean of the group that listened only once in both posttests. The immediate and delayed Form Test scores were analyzed separately using the Kruskal-Wallis test. In the immediate test there was a significant group effect ($X^2(3) = 9.409$, $p = .024$). The post-hoc *Mann*-Whitney U test with Bonferroni correction ($p = .013$) showed that Groups B and C outperformed Group A with a small and medium effect size respectively (Group B: $U(1) = 680.500$, $Z = -2.591$, $p = .01$, $r = .29$; Group C: $U(1) = 630.500$, $Z = -2.576$, $p = .01$, $r = .31$). There was no significant difference between Groups B and C and the effect size was small ($U(1) = 891.000$, $Z = -.099$, $p = .92$, $r = .05$). In the delayed test there were no significant group effects ($X^2(3) = 4.855$, $p = .183$), indicating that the differences between the groups had disappeared one week later.

Differences between the immediate and delayed test scores were examined using Wilcoxon Signed Ranks Tests with a Bonferroni correction, resulting in the significance level set at $p < .0125$. None of the groups showed any statistically significant change and the effect sizes were all small. (Group A: $Z = -2.279$, $p = .02$, $r = .18$; Group B: $Z = -.799$, $p = .42$, $r = .06$; Group C: $Z = -.730$, $p = .47$, $r = .06$).

Table 2. Descriptive Statistics for Form Test

Group	Immediate			Delayed		
	Mean	SD	Range	Mean	SD	Range
A (45)	2.11	1.37	0-5	2.60	1.37	0-5
B (44)	2.96	1.45	0-5	3.14	1.62	0-8
C (41)	3.12	1.78	0-8	3.33	1.95	0-7
Total (130)	2.73	1.53	0-8	3.02	1.65	0-8

Note: A = listening one time; B = listening three times; C = inference training + listening three times

Table 3 gives the descriptive statistics for the Reception Test. As for the Form Test, all the groups that listened three times outscored the group that listened just once in both the immediate and delayed test. Scores were analyzed separately using Kruskal-Wallis tests. There were no significant group effects for either the immediate test ($X^2(3) = 7.613$, $p = .055$) or the delayed test ($X^2(3) = 1.658$, $p = .646$). The effect sizes for the group comparisons in the immediate tests were small for A vs. B ($r = .25$) and negligible for B vs. C ($r = .06$). In the delayed tests all the effect sizes were negligible except for A vs. B which was small ($r = .13$). In contrast to the results of the Form Test, scores decreased over time but none of the groups showed any statistically significant change with small effect sizes (Group A: Z = -.371, $p = .71$, $r = .01$; Group B: Z = -1.235, $p = .22$, $r = .14$; Group C: Z = -1.415, $p = .16$, $r = .10$).

Table 3. Descriptive Statistics for Reception Test

Group	Immediate			Delayed		
	Mean	SD	Range	Mean	SD	Range
A (45)	3.47	1.80	0-7	3.42	1.95	0-5
B (44)	4.32	1.54	2-8	3.89	1.60	0-7
C (41)	4.12	1.78	1-8	3.75	1.77	0-8
Total (130)	3.97	1.71	0-8	3.69	1.77	0-8

Note: A = listening one time; B = listening three times; C = inference training + listening three times

Table 4 shows the descriptive statistics for the Production Test. There was very little evidence of productive knowledge of the target items in any of the groups in either the immediate or the delayed test. There was also only limited variance in the individual groups' scores (i.e., the range was at most 0 to 3). For this reason no inferential analyses were conducted. Little change was evident from the immediate to the delayed test.

Table 4. Descriptive Statistics for Production Test

Group	Immediate			Delayed		
	Mean	**SD**	**Range**	**Mean**	**SD**	**Range**
A (45)	0.00	0.00	0-0	0.04	0.08	0-1
B (44)	0.11	0.32	0-1	0.11	0.49	0-3
C (41)	0.27	0.67	0-3	0.28	0.60	0-2
Total (130)	0.13	0.33	0-3	0.14	0.39	0-3

Note: A = listening one time; B = listening three times; C = inferencing training + listening three times

Summary of results

1. The students demonstrated reasonable comprehension of the listening texts with scores above 50% in all groups except Group A (one-time listening). There were, however, marked differences in listening scores within each group and for the whole sample (see scores for range in Table 1). Group A performed poorly compared to the other two groups. There was no evidence that comprehension improved as a result of the inference-training. In other words the single factor that assisted comprehension was the opportunity to listen to the text three times as opposed to once.

 2. Scores on the Production Test were very low indicating that overall the learners did not develop productive control of the target words. Scores were higher on the Form and Receptive Tests with the three-time listening groups (i.e., B and C) obtaining higher scores than the one-time listening group (i.e., A). However, the group differences reached statistical significance only for the immediate Form Test with Groups B and C both outperforming Group A. No statistically significant differences were found between Groups B and C on either the Form or Reception Tests.

 3. By and large there was no decrease in scores from the immediate to the delayed test (one week later).

Discussion

Overall the learners demonstrated reasonable success in comprehending the listening texts. The whole sample's mean group score for comprehension was 6.95 (maximum 10). In general, then, the learners were successful in extracting the relevant information from the texts and transferring it to the chart in the first task and to the application form in the second. The general level of comprehension compares favorably with those reported in other task-based studies (e.g., Ellis et al, 1994; Loschky, 1994). However, as Table 1 shows, the learners differed markedly in their ability to comprehend, with some scoring zero.

 Vocabulary scores (out of 10) varied according to test. The learners largely failed to develop productive control over the target words. Clearly exposure to the target words even three times was insufficient to ensure productive knowledge. However, they were able to demonstrate receptive knowledge and also, to a lesser extent, they could recognize the

form of the target words.

It is clear, then, that many of the learners were able to both listen-to-comprehend and listen-to-learn when they performed the tasks. While they listened they completed the chart/application form by identifying the information required. At the same time they were able to attend to at least some of the target words, take note of their form, and use the context to help them establish a form-meaning mapping. However, the ability of these learners to dual task in this way varied considerably and knowledge of the target words was quite shallow as reflected in the failure to develop productive knowledge. This was not surprising given that the tasks themselves did not directly attract attention to the target words.

The results showed that those learners who had the opportunity to listen to the text three times (i.e., in Groups B and C) demonstrated better comprehension than those learners who listened only once (i.e., Group A). The effect sizes for the comparisons between Group A and the other two groups were all large. In other words, as in previous studies (e.g., Cervantes & Gainer, 1992; Chang & Reid, 2007; O'Bryan, 2010), text repetition aided comprehension. It should be noted, however, that some of the students in Group A were able to complete the tasks successfully (see Table 1). It would seem, therefore, that the advantage that text repetition conferred was mainly for those learners with limited listening abilities. Chang and Reid (2008) found that text repetition was less effective for low proficiency than for high-proficiency learners and this may have been the case in this study too as some learners in Groups B and C did not appear to benefit much from the text repetition.

The second research question asked whether the listening conditions had a differential effect on the learners' vocabulary acquisition. The two groups (Groups B and C) that listened to the texts three times achieved higher scores on all three vocabulary tests (see Table 2) than the group that only listened once. In the immediate Form Test, the differences in scores between Group A and B and between A and C reached statistical significance with a small to medium effect sizes for A vs. B ($r = .29$) and for A vs. C ($r = .31$). In the immediate Reception Test, Groups B and C again outperformed A but none of the pairwise comparisons reached statistical significance and effect sizes were small. Although none of the groups acquired much productive knowledge, Groups B and C again outscored Group A.

Overall, then, just as listening three times aided comprehension, so too it helped vocabulary acquisition. In fact, there were positive correlations – some of them statistically significant between the learners' comprehension scores in Groups B and C and their vocabulary acquisition scores, whereas the correlations between comprehension and vocabulary scores for the learners in Group A were all negative [1].

Target word frequency might account for the three-time listening groups' advantage in vocabulary acquisition; Group A heard the target words twice only whereas the other two groups all heard them 6 times. However, previous studies (e.g., Revesz & Brunfaut, 2013; Vidal, 2003) did not find a relationship between word frequency and vocabulary acquisition from listening. More important than sheer frequency, perhaps, is whether learners are able to make use of the context to infer the word meaning. The contextual

clues that were embedded in the texts were designed to help the learners work out the meanings of the target words. However, these clues were implicit in nature. Vidal (2003) reported that implicit word elaboration was much less effective than explicit (i.e., providing an explicit definition). This may be one reason why those learners who listened to the text just once found difficulty in establishing a link between the forms of the target words and their contextual meanings. In contrast, those learners who listened three times were better able to pick up the contextual clues and make connections between form and meaning. Text repetition, then, may be especially important when the clues to the meanings of new words are only implicit in nature.

It might be expected that inference-training would help. Group C did achieve statistically higher vocabulary scores than Group A but so did Group B, which received no such training. Overall, then, it was the text-repetition that mattered and the inference-training had little effect. The scores on the delayed vocabulary tests were very similar to those for the immediate tests suggesting that whatever vocabulary learning had taken place was durable at least in the short term (one week).

Conclusion

The clear lesson to be learned from the study is that repetition of a listening text assisted both comprehension and vocabulary acquisition (especially of form) while inference-training provided no additional benefit for the L2 learners we investigated. As Goh (2008) pointed out the difficulty that such learners have when listening is with phoneme and word recognition. When they have the opportunity to listen a second and third time they can build on their initial conceptualization of the text by processing those parts that they initially skipped, by clarifying propositions that were only partially understood first time round, and by adding details to stored propositions. Text-repetition facilitates those processes that learners experience difficulty with – namely, acoustic-phonetic and syntactic parsing. It helps both comprehension and also the noticing of new words needed for acquisition to take place (Schmidt, 2001).

Text repetition may work in the same way as interactionally-modified input, which task-based studies (e.g., Ellis & He, 1999; Ellis, Yamazaki & Tanaka, 1994) have shown can assist both comprehension and vocabulary acquisition. However, there is a difference. Interactionally-modified input allows learners some control over the input, whereas text-repetition does not. However, O'Bryan (2010) found no significant difference in comprehension between a group that experienced text repetition and a group that was allowed to control access to the input. Therefore where comprehension and incidental vocabulary acquisition is concerned what really seems to count is repeated exposure to the input.

It might be argued, however, that text repetition is "pedagogic" and does not occur in real-life listening. In other words, repeating a text might help learners comprehend a specific text but will not help them become better listeners. In fact, though, there are real-life opportunities for repeated listening (e.g., re-watching the same TV show) and

motivated learners are likely to take advantage of such opportunities. Also, repeated listening can help build listening proficiency by contributing to the development of linguistic proficiency. For example, as learners' lexicon grows, so too will their ability to process text in real time. It would be unwise to dismiss text repetition in either teaching or testing listening ability.

It is of course premature to conclude that inference training has no role to play in listening instruction. The fact that it failed to enhance the effect of text-repetition may have been because the time devoted to it was relatively short (i.e., 10 minutes before each listening task) although the time we allocated was in line with that in other studies that have investigated the effects of pre-listening support. It is also possible that inference-training would have proved beneficial for those students who listened just once although given the problem the learners in Group A had in comprehending the texts this seems doubtful. However, a case might also be made out for the long-term effect of inference-training but that remains to be investigated.

The tasks we used in the study were designed to provide both listening practice and input for vocabulary learning. As noted in the introduction, listening-to-learn and listening-to-comprehend involve different processes but they can occur in parallel when learners have time to process the input. Thus, we can see no reason why listening activities should not have the dual goal of comprehension and vocabulary development. However, given that there were very low scores on the Production Test, production practice in the post-task phase of listening lesson may be needed to develop productive knowledge of new words acquired incidentally through listening.

Notes

1. Correlations between comprehension scores and four test scores (Immediate and delayed Form test, and Immediate and delayed Reception test).

Group	Correlation Coefficient	Immediate Form Test	Delayed Form Test	Immediate reception	Delayed reception
A (45)	r	-.120	-.191	-.132	-.043
	p	.431	.209	.387	.777
B (44)	r	.366[*]	.261	.121	.308[*]
	p	.015	.087	.436	.042
C (41)	r	.219	.238	.240	.219
	p	.169	.139	.131	.174

Acknowledgement

References

Alavi, S., & Jambaz, F. (2014). Comparing two pre-listening supports with Iranian EFL learners: Opportunity or obstacle. *RELC Journal*, *45*(3), 253-267.

Berne, J. (1995). How does varying pre-listening activities affect second language listening comprehension? *Hispania, 78*(2), 316-329.

Brown, R., Waring, R., & Donkaewbua, S. (2008). Incidental vocabulary acquisition from reading, reading while listening, and listening to stories. *Reading in a Foreign Language*, *20*, 136-163.

Cervantes, R, & Gainer, G (1992). The effects of syntactic simplification and repetition on listening comprehension. *TESOL Quarterly, 26*(4), 767-770.

Chang, A. (2007). The impact of vocabulary preparation on L2 listening comprehension, confidence and strategy use. *System, 35*, 534-550

Chang, A., & Reid, J. (2007). Support for foreign language listeners. *RELC Journal*, *38*(3), 375-395.

Chang, A., & Reid, J. (2008). The effects of listening support on the listening performance of EFL learners. *TESOL Quarterly, 40*(2), 375-397.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Coskun, A. (2010). The effect of metacognitive strategy training on the listening performance of beginner learners. *Novitas-ROYAL (Research on Youth and Language), 4*(1), 25-50.

Elkhafaifi, H. (2005). The effect of pre-listening activities in Arabic learners. *Foreign Language Annals, 38*(4), 505-513.

Ellis, N. (2002). Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition, 24*, 143-188.

Ellis, R. (1995). Modified input and the acquisition of word meanings. *Applied Linguistics*, *16*(4), 409-441.

Ellis, R., & He, X. (1999). The roles of modified input and output in the incidental acquisition of word meanings. *Studies in Second Language Acquisition, 21*, 319-333.

Ellis, R., Tanaka, Y., & Yamazaki, A. (1994). Classroom interaction, comprehension and the acquisition of word meanings. *Language Learning, 44*, 449-91.

Field, J. (2004). An insight into listeners' problems: Too much bottom-up or too much top-down. *System*, *32*, 363-377.

Fraser, C. A. (1999). Lexical processing strategy use and vocabulary learning through reading. *Studies in Second Language Acquisition, 21*, 225-241.

Goh, C. (2008). Metacognitive instruction for second language listening development. *RELC Journal, 39*, 188-213.

Jafari, K., & Hashim, F. (2012). The effects of using advance organizers on improving EFL learners' listening comprehension: A mixed method study. *System, 40*, 270-282.

Loschky, L. (1994). Comprehensible input and second language acquisition: What is the relationship? *Studies in Second Language Acquisition, 16*, 303-323.

Nation, P. (2001). *Learning vocabulary in another language*. Cambridge, England: Cambridge University Press.

O'Bryan, A. 2010. Investigating the effect of segmented and whole-text repetition in listening

comprehension, listening process and listening problems. Unpublished doctoral dissertation, Ohio State University.

Plonsky, L. (2011). The effectiveness of second language strategy instruction: A meta-analysis. *Language Learning, 61*, 993-1038.

Read, J. (2000). *Assessing vocabulary.* Cambridge: Cambridge University Press.

Revesz, A., & Brunfaut, T. (2013). Text characteristics of task input and difficulty in second language listening comprehension. *Studies in Second Language Acquisition, 35*, 31-65.

Sarandi, H. (2010). Context-related support and listening comprehension: Some limitations. *Procedia, Social and Behavioural Sciences, 3*, 5606-5611.

Schmidt, R. (2001). Attention. In P. Robinson (Ed.): *Cognition and second language instruction* (pp. 3-32). Cambridge: Cambridge University Press.

Vandergrift, L. (2003). From prediction through reflection: Guiding students through the process of L2 listening. *Canadian Modern Language Review, 59*, 425-440.

Vandergrift, L. (2005). Relationships between motivation orientations, metacognitive awareness and proficiency in L2 listening. *Applied Linguistics*, *26*, 70-89.

Van Zeeland, H., & Schmitt, N. (2013). Incidental vocabulary acquisition through L2 listening: A dimensions approach. *System, 41*, 609-624.

Vidal, K. (2003). Academic listening: A source of vocabulary acquisition. *Applied Linguistics, 24*, 56-89.

Vidal, K., 2011. A comparison of the effects of reading and listening on incidental vocabulary acquisition. *Language Learning*, 61, 219-258.

Wei, C., & Wu, Y. (undated). Lexical inferencing in second language listening comprehension. Downloaded from tp://www.sinoss.net/qikan/uploadfile/2010/1130/5039.pdf. 19th January 2015.

Wesche, M., & Paribakht, T. S. (1996). Assessing second language vocabulary knowledge: Depth vs. breadth. *Canadian Modern Language Review, 53*, 13-39.

(Copyedited by Duncan SIDWELL & DING Yanren)

About the authors

Rod ELLIS is Distinguished Professor of Applied Language Studies in the University of Auckland, and also Cheung Kong Scholar Chair Professor at Shanghai International Studies University. He was recently elected a fellow of the Royal Society of New Zealand. His published works include numerous articles and books on second language acquisition, language teaching and teacher education, including the *Study of Second Language Acquisition* (OUP). His latest books published in 2013 (with Natsuko Shintani) is *Exploring Language Pedagogy through Second Language Acquisition Research* (Routledge) and in 2015 *Understanding Second Language Acquisition 2*[nd] *Edition* (Oxford University Press). Email: r.ellis@auckland.ac.nz

Chang Le, male, PhD in Language Teaching and Learning from the University of Auckland, has been teaching EFL since 1994. He's currently a professor in Bohai University. His research interests include L2 vocabulary acquisition and corpus for special purposes. Email: gmwa@163.com