



Concordia Working Papers
in Applied Linguistics

Proceedings of the International Symposium on the Acquisition of Second Language Speech
Concordia Working Papers in Applied Linguistics, 5, 2014 © 2014 COPAL

The Effectiveness of Interactive Teaching Methods in EFL Classrooms: A Comparison with Bottom-Up and Top-Down Methods

Mamiko Orii-Akita

Waseda University

Abstract

This study evaluated the effectiveness of three teaching methods that employ three models of listening comprehension—bottom-up, top-down, and interactive—in the context of teaching English listening comprehension to Japanese university students.

The study involved twenty 90-minute sessions with three experimental groups (bottom-up, BG; top-down, TG; interactive, IG) that underwent customized listening training using a self-developed textbook/CD. The BG engaged in dictation activities. The TG received explicit instruction on listening strategies, and the IG, on listening strategies and prosody training. In addition, the informants engaged in speaking activities: the BG and TG produced and presented dialogues in pairs, while the IG engaged in basic prosody instruction/practice, followed by shadowing activity.

Listening tests were administered three times: before, immediately after, and two months after the 20 sessions. Results show that the IG outperformed the other two groups, indicating the interactive model's superiority in improving EFL learners' listening comprehension ability.

Previous research into how language users derive meaning from the sounds they hear has proposed three types of processing models: bottom-up, top-down, and interactive. Early research in this area proposed a bottom-up model in which listeners build understanding by starting with the smallest units of the acoustic message, such as phonemes, which are then combined into words, phrases, clauses, and sentences (Carrel, 1988; Flowerdew & Miller, 2005; Rost, 1990). This model is based on the traditional understanding of communication as a means of simply transmitting information. This model is also referred to as a text-based or perception-based model since it positions listening as a passive process.¹

On the other hand, Vandergrift (1997) argues that effective listeners use top-down processing. The top-down model differs from the bottom-up model in that listening is seen as an active process that involves seeking necessary information rather than simply analyzing speech as a series of sounds (Flowerdew & Miller, 2005, p. 26). The top-down model, also referred to as process-based approach, presupposes the use of previous background knowledge to predict content (Ross, 1975; Rubin, 1994). Previous studies have shown that effective listeners use top-down processing, more than the bottom-up model (Cross, 2009; Vandergrift, 1997, Vandergrift & Tafaghodtari, 2010). Furthermore, a number of researchers argue that training learners to use listening strategies are beneficial to their improvement in listening ability (Field, 2008; Lynch & Mendelsohn, 2002; Rost, 1990, 2011).

More recent studies have suggested that listening comprehension is the result of the interplay of information at different levels, and that listeners employ both bottom-up and top-down models in order to process this information (O'Malley & Chamot, 1990). This type of processing is known as the interactive model. Several researchers have argued that effective learners employ the interactive model more frequently than the top-down or bottom-up models (Buck, 2001; Cleary et al., 2007; Field, 2008; Higuchi, 1998) and listening comprehension teaching should be carried out including both top-down and bottom-up training (Hinkel, 2006, p. 110; Vandergrift, 2007; Vandergrift & Goh, 2012).

¹This work was supported by JSPS KAKENHI Grant-in-Aid for Scientific Research (C) [Grant Number 24520660]. I appreciate the audience at New sounds 2013 (Concordia University) for their feedback. In the presentation, the effectiveness of an interactive teaching approach was examined from the viewpoint of human cognitive processing in the discussion section; however, due to space limitations, I decided to focus on reporting the results of the experimental study (and exclude the processing part).

There are important pedagogical implications of these models of listening comprehension and a number of studies have provided quantitative data to evaluate the effectiveness of teaching methods that use bottom-up, top-down, and interactive models. However, simply putting these recommended methods into practice and providing a theoretical report on their educational benefits does not provide sufficient evidence of their effectiveness compared with other methods. By carefully studying these methods, we will be able to verify the most effective method for teaching listening comprehension in a foreign language. This study addresses this empirical gap by comparing and evaluating the effectiveness of three different teaching methods that employ three different models of listening comprehension—bottom-up, top-down, and interactive—in the context of teaching English listening comprehension to Japanese university students.

METHODOLOGY

Overview of Experimental Design

The data for this study were collected during experimental classes conducted in three intermediate English communication classes taught by the author in 2012. The 90-minute experimental classes were given 20 times in 2012 from April to November.

Participants' listening abilities were measured on three occasions using the Web-based Test for English Communication (WeTEC, provided by CASEC). It uses an IRT-based CAT system and adapts questions to the student's individual level. It is therefore shorter than traditional paper-based tests. The test last approximately one hour for reading, vocabulary and listening, with 45 listening questions, and provides an accurate measure of students' abilities. It is considered a reliable assessment since the results correlate with students' TOEIC scores. The WeTEC also provides a breakdown of the results for listening ability, showing individual scores for overall understanding (section 3) and specific information (section 4). This approach is in line with the top-down teaching model, which is another reason the test was judged to be a superior option compared to other online tests. In addition, the test is utilized for student placement and academic assessment by the university where the author is employed, and over 10,000 people take the WeTEC test each year (including those who take it multiple times).

The entire WeTEC test (section 1-4) was administered to the participants; however, since this study aims to investigate the relationship between teaching method and listening ability, I decided to analyze only the listening sections, i.e. section 3 and section 4. In this paper, “total score” will therefore refer to a combined total of scores from sections 3 and 4.

A pretest was conducted before the first session of the experimental classes and a posttest was administered immediately after the final session in November 2012. In addition, I implemented a delayed posttest two months after the end of the experimental classes to find out if the educational effect was sustained. This final test was conducted in February 2013.

Experimental Classes and Lesson Format

Overview. Each group underwent group-specific listening training while using the same textbook developed by the author: the bottom-up group (henceforth, BG) completed the dictation activity, and the top-down group (henceforth, TG) and interactive group (henceforth, IG) completed listening strategy exercises. The listening comprehension took approximately 60 minutes out of the 90-minute class. The participants also engaged in speaking activities that took approximately 30 minutes to complete. For these speaking activities, the BG and TG produced and presented dialogues in pairs on relevant topics, while the IG engaged in basic prosody instruction/practice, followed by a shadowing activity.

Lesson Format for Listening Training. A listening textbook (on travel) that the author developed was used for each group, and dialogues from the CD were played the same number of times (i.e., five times) for each group. Most previous studies have used different teaching materials for each experimental group and very few studies have provided uniform input across groups because of the associated practical difficulties. This study takes a different approach because I controlled the aural input received by the participants by using the same listening materials for each group. In this way, I aimed to measure the education effect of each teaching method with greater accuracy. This type of methodology has not yet been implemented in long-term studies.

The bottom-up group (BG, 24 participants) engaged in dictation activities that were designed using the bottom-up model. These activities *implicitly* directed students' attention to sound changes. The blanks in the

dictation activity represented clear examples of the prosodic point being covered as well as key travel expressions. The instructor then distributed the entire script and provided explanations about the vocabulary and grammar.

The top-down group (TG, 25 participants) received *explicit* instructions on listening strategies (e.g., skimming and scanning) that were designed using the top-down model. Based on Vandergrift (1997), following eleven strategies were chosen for instruction: (1) directed attention; (2) selective attention; (3) verify predictions; (4) linguistic inferencing; (5) voice inferencing; (6) inferencing between parts; (7) Personal elaboration; (8) world elaboration; (9) questioning elaboration; (10) note taking; (11) Cooperation.

The following activities were carried out: (1) the pre-listening activity to activate learners' schemata on the topic, (2) the first listening with exercises on listening for overall understanding, and (3) the second listening with exercises on listening for specific information (Ellsworth, 2006; Wilson 2008). When necessary, a listening strategy handout (distributed at the start of the semester) was used to explicitly teach students how to scan for essential information, predict the content and intended meaning of the story, and utilize logical and visual information. The interactive group (IG, 25 participants) also received instructions on listening strategies and completed the same listening exercises as TG.

Lesson Format for the Speaking Activity. As for BG and TG, students were tasked with creating dialogues and sentences about the unit topic, either in pairs or individually, which they presented afterwards in small groups or to the rest of the class. For example, unit 3 covered the topics of getting around and paying for transport. Students therefore practiced buying railway tickets. The entire production activity lasted for about 30 minutes.

As for IG, participants were given prosody instruction along with pronunciation practice for 30 minutes. The first 10 minutes were devoted to outlining the main features of the new pronunciation point in simple, jargon-free terms on the blackboard. For example, units 1 and 2 focused on the rhythm of English and the use of weak forms. Subsequent units (with a new pronunciation point generally being introduced every other unit) covered thought groups, assimilation, linking, and elisions.

The last 20 minutes were spent on practical pronunciation exercises. First, the class used very simple sentences to practice pronunciation (e.g., the phrase "I miss you" was used to practice assimilation). Students then

engaged in pronunciation exercises using the listening script that had been distributed earlier in the lesson. The relevant prosodic points were either pointed out by the instructor or identified by the students in pairs. The instructor then demonstrated and asked students to repeat individually. Finally, students worked on pronunciation exercises corresponding to one paragraph of the spoken material. The dialogues in the listening textbook were designed as much as possible to incorporate the types of prosodic issues covered in the unit. Specifically, I tried to include at least 10, but if possible up to 20, instances of the relevant prosodic points in each dialogue (each dialogue lasted 30-45 seconds in total). From units 14 to 20, participants were instructed to carry out a shadowing activity instead of basic pronunciation training through the oral reading of the manuscript. Shadowing is a technique which students immediately repeat the aural input as they listen to the CD (Kadota, 2006, 2012).

RESULTS

Overview of Data Analysis

We conducted a two-way repeated measures ANOVA of students' scores on section 3 (measuring listening for overall understanding), section 4 (measuring listening for specific information), and combined total scores (section 3 + section 4). For each set of scores, the within-subjects factor that compared tests (3 levels: pretest, posttest, and delayed) as well as the between-subjects factor that compared groups (3 levels: BG, TG, and IG) were tested at the significance level of 5%. Mauchly's sphericity test for homogeneity of variance, a prerequisite for carrying out the repeated measures ANOVA was carried out for each section; indicated that the assumption of sphericity had not been violated for section 3 ($p = 0.469$ (n.s.)), section 4 ($p = 0.389$ (n.s.)) and total score ($p = 0.992$ (n.s.)). The general linear model (GLM) in SPSS (SPSS Statistics ver.17.0 for Windows. SPSS Japan Inc., Tokyo, Japan) was used to calculate the repeated measures ANOVA.

In cases where there was a significant interaction effect, a multiple comparison of the three groups by test and the three tests by group were carried out. Tukey's HSD test to compare the three groups by test (BG vs. TG, BG vs. IG, and TG vs. IG) were employed. For the comparison of the three tests by group, we used Bonferroni-adjusted paired t-tests (pretest

vs. posttest, pretest vs. delayed, and posttest vs. delayed). Table 1. and Table 2. show the results of all the analyses.

Summary of ANOVA

Table 1 summarizes the results of the ANOVA and shows the basic statistics of test scores. The data is broken down by *test* and section as indicated by the two left columns and by experimental group across the row.

Table 1. Summary of ANOVA

Section	Test	Group						Main effect				Interaction (group x test)	
		BG (n=17)		TG (n=20)		IG (n=18)		Group		Test			
		Mean	SD	Mean	SD	Mean	SD	F value	p value	F value	p value	F value	p value
Section3	Pretest	165.65	22.366	159.40	24.659	163.22	12.959	2.088	0.134	13.477	0.000***	5.830	0.000***
	Posttest	171.71	14.225	166.50	21.309	173.17	16.078						
	Delayed	166.29	21.502	169.15	23.703	190.72	13.123						
Section4	Pretest	148.00	17.346	144.80	19.606	134.33	14.000	0.702	0.500	9.065	0.000***	3.183	0.016*
	Posttest	153.94	19.986	146.60	15.823	148.61	19.135						
	Delayed	151.47	20.100	147.90	18.046	153.17	16.104						
Total score	Pretest	313.65	33.386	304.20	38.164	297.56	20.171	0.622	0.541	28.308	0.000***	10.288	0.000***
	Posttest	325.65	26.907	313.10	31.616	321.78	29.883						
	Delayed	317.76	32.128	317.05	35.685	343.89	21.884						

Finding 3

Finding 1

Finding 2

First, as shown in Finding (1), The ANOVA on the within-subjects factor *test* shows that *test* as main effect is significant in all sections, including Section 3, Section 4, and Total Scores, indicating that the mean scores of three tests differed from each other.

Furthermore, as you can see in Finding (2), the interaction between *test* and *group* is also significant in all sections, indicating that the test scores changed differently by *group*, namely, BG, TG, and IG, because of a synergetic effect.

Next, as shown in Finding (3), the results of the ANOVA on the between-subjects factor *group* showed that *group* was not significant as the main effect. Therefore, there was no difference in the mean score among the three groups.

Summary of Post-hoc Analyses

Table 2. summarizes the results of all post-hoc analyses. Only the key findings are presented here. The far left column shows the test scores analyzed. The left half of the table shows the results of paired comparison of groups, or the comparison of all combinations of BG, TG, and IG, using Tukey's HSD test. The right half of the table shows the results of paired comparison of tests, or the comparison of all combinations of Pretest, Posttest, and Delayed, using Bonferroni-adjusted paired t-test.

Table 2. Summary of Post-hoc Analysis (Multiple Comparison)

Section	Post-hoc analysis (comparison of groups by test): Multiple comparison (Tukey's HSD test)					Post-hoc analysis (comparison of tests by group): Bonferroni-adjusted paired t-test				
	Test	Group (I)	Group (J)	Difference in means (I - J)	p value	Group	Test (I)	Test (J)	Difference in means (I - J)	p value
Section3	Pretest	BG	TG	6.25	0.635	BG	Pretest	Posttest	-6.06	0.578
		BG	IG	2.42	0.936		Pretest	Delayed	-0.65	1.000
		TG	IG	-3.82	0.838		Posttest	Delayed	5.41	0.436
	Posttest	BG	TG	5.21	0.647	TG	Pretest	Posttest	-7.10	0.447
		BG	IG	-1.46	0.968		Pretest	Delayed	-9.75	0.139
		TG	IG	-6.67	0.482		Posttest	Delayed	-2.65	1.000
	Delayed	BG	TG	-2.86	0.903	IG	Pretest	Posttest	-9.94	0.123
		BG	IG	-24.43	0.002**		Pretest	Delayed	-27.50	0.000***
		TG	IG	-21.57	0.005**		Posttest	Delayed	-17.56	0.000***
Section4	Pretest	BG	TG	3.20	0.840	BG	Pretest	Posttest	-5.94	0.307
		BG	IG	13.67	0.058		Pretest	Delayed	-3.47	0.820
		TG	IG	10.47	0.158		Posttest	Delayed	2.47	1.000
	Posttest	BG	TG	7.34	0.448	TG	Pretest	Posttest	-1.80	1.000
		BG	IG	5.33	0.666		Pretest	Delayed	-3.10	1.000
		TG	IG	-2.01	0.939		Posttest	Delayed	-1.30	1.000
	Delayed	BG	TG	3.57	0.822	IG	Pretest	Posttest	-14.28	0.010*
		BG	IG	-1.70	0.959		Pretest	Delayed	-18.83	0.000***
		TG	IG	-5.27	0.646		Posttest	Delayed	-4.56	0.787
Total score	Pretest	BG	TG	9.45	0.642	BG	Pretest	Posttest	-12.00	0.004**
		BG	IG	16.09	0.300		Pretest	Delayed	-4.12	1.000
		TG	IG	6.64	0.797		Posttest	Delayed	7.88	0.250
	Posttest	BG	TG	12.55	0.411	TG	Pretest	Posttest	-8.90	0.456
		BG	IG	3.87	0.921		Pretest	Delayed	-12.85	0.051
		TG	IG	-8.68	0.642		Posttest	Delayed	-3.95	1.000
	Delayed	BG	TG	0.71	0.997	IG	Pretest	Posttest	-24.22	0.000***
		BG	IG	-26.12	0.039*		Pretest	Delayed	-46.33	0.000***
		TG	IG	-26.84	0.025*		Posttest	Delayed	-22.11	0.000***

(2)

(4)

(6)

(7)

Section 3. The results of Tukey's HSD test shows a difference between the mean values of BG and IG ($p = 0.002$) and of TG and IG ($p = 0.005$) on the delayed test, with IG showing significantly higher mean scores in Section 3 when compared with both BG and TG (as shown by (1) in Table 2). Then, we carried out a multiple comparison (Bonferroni-adjusted paired t-test) of tests by group. As shown by (2), there is a difference between the mean values of pretest and delayed ($p = 0.000$) and of posttest and delayed ($p = 0.000$) for IG, with Section 3 mean scores shown to be significantly higher on the delayed test than on both the pretest and the posttest. See also Figure 1.

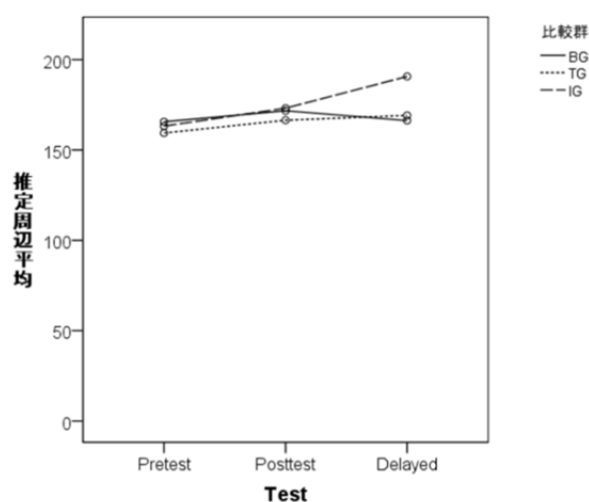


Figure 1. Mean values for section 3

Section 4. Regarding factors related to score changes in Section 4, the main finding was that there was a significant difference between the three tests (pretest, posttest, and delayed) but not between the groups (BG, TG, and IG). As you can see from (3) in Table 2, none of the pairs had mean scores that differed from each other. However, as shown in (4), results of the multiple comparison tests for each group show that the comparison of scores for the IG group's pretest and posttest scores ($p = 0.010$), and the pretest and delayed scores ($p = 0.000$), we observed that there was a significant difference in mean values between the two tests. In other words, as can be seen in Figure 2, IG improved their mean score in the posttest, and the mean score in the delayed test remained at the same level as the posttest.

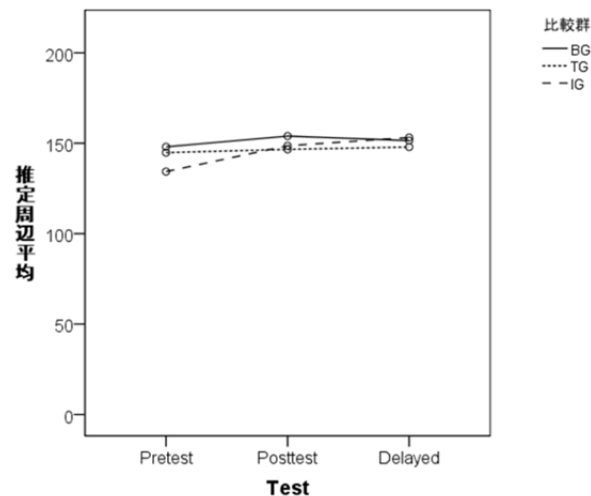


Figure 2. Mean values for Section 4

Results of Total Score. Results of the multiple comparison by *test* showed that differences in the delayed test between the BG and IG groups and between the TG and IG groups were significant ($p = 0.039$ and $p = 0.025$, respectively) as shown by (5) in Table 2, indicating that there were mean differences among the groups for the delayed test. Scores were significantly higher for the IG compared to the BG group and the TG group.

Next, results of the multiple comparisons for each *test* by *group* indicated that there was a significant difference between the pretest and posttest within the BG group ($p = 0.004$): the posttest average was significantly higher when compared to the pretest average (as shown by (6) in Table 2, see also Figure 3). Meanwhile, mean differences between the following pairs of tests were observed within the IG group (see (7) in Table 2): pretest vs. posttest ($p = 0.000$), pretest vs. delayed ($p = 0.000$), and posttest vs. delayed ($p = 0.000$). For the IG group, scores improved over time from the pretest to the posttest and then from the posttest to the delayed test.

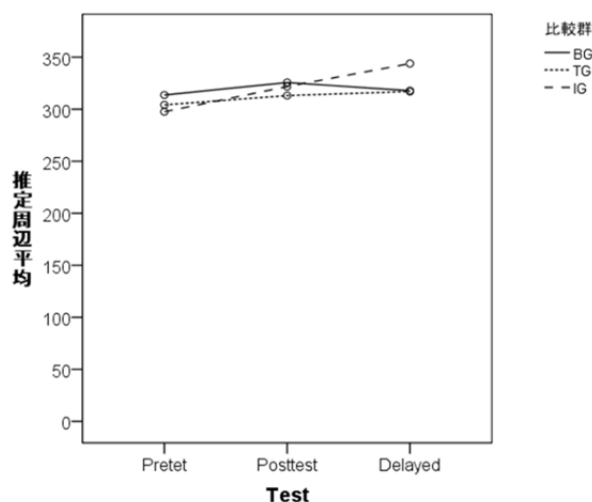


Figure 3. Mean values for total score

DISCUSSION

Summary and Examination of Results

The objective of this research was to examine effective methods for the teaching of listening skills in English as a foreign language (EFL) at the university level. Specifically, it was designed to investigate the effectiveness of an interactive teaching approach by comparing and verifying multiple teaching methods.

Here, results obtained will be summarized by experimental groups, supplementing the detailed presentation by section above.

For BG, no improvement in scores was observed on either Section 3 or Section 4. However, their posttest average for total score was higher than the pretest value. Even though BG's scores for Section 3 and Section 4 did not increase significantly, when the scores for both sections were added together, the increase did become significant. However, despite the fact that its total score improved on the posttest, its delayed posttest score returned to the level of the pretest. In other words, the effects of the classes were only temporary in this group.

Unfortunately, the scores for TG did not improve. No change in mean values was observed on Section 3, on Section 4, or for total score. From this, it can be concluded that when implemented as a single method, the teaching of listening strategies did not contribute to any improvement in listening abilities for TG. In particular, the instruction provided by the classes was specifically intended to train the students in listening for overall understanding by teaching them about various listening strategies;

this goal does not seem to have been achieved, as the measurement of listening for overall understanding in Section 3 showed no improvement in scores for TG.

In contrast, a significant improvement was observed in Section 3, Section 4 and total score for IG, which received both top-down (listening strategies) and bottom-up (prosody) teaching.

In Section 3, which measured listening for overall understanding, despite the lack of significant improvement on the immediate posttest, significant improvement was observed on the delayed posttest. From this finding, it seems that interactive teaching had educational effects that contributed to improving students' scores even after the teaching itself was completed (i.e., after the classes were finished). A comparison between groups for each test shows significant differences in delayed posttest scores between BG and IG and between TG and IG.

In Section 4, which measured listening for specific information, an improvement in scores was observed in the immediate posttest and maintained in the delayed posttest. However, the score of IG in Section 4 was not significantly higher than those obtained by TG and BG.

In total for Sections 3 and 4 (total score), IG's score improved significantly in the posttest and again in the delayed posttest, showing that the results not only were maintained but also improved over time. A comparison between groups for each test shows significant differences in delayed posttest scores between BG and IG and between TG and IG.

CONCLUSION

The objective of this research was to examine effective methods for the teaching of listening skills in English as a foreign language (EFL) at the university level. Specifically, it was designed to investigate the effectiveness of an interactive teaching approach by comparing its effectiveness to that of methods in which instruction was implemented as a single method (either bottom-up or top-down).

The data show that interactive teaching was an effective teaching method in conjunction with the classes used in this experiment. In other words, learners assessed in the present study gained the ability to more smoothly use bottom-up and top-down processing in parallel as a result of practice in bottom-up processing and the teaching of listening strategies to encourage top-down processing, provided in tandem.

REFERENCES

- Buck, G. (2001). *Assessing Listening*. Cambridge: Cambridge University Press (CUP).
- Carrel, P. (1988). Some causes of text-boundness and schema interference in ESL reading. In P. Carrel, J. Devine, & D. Eskey (Eds.) *Interactive Approaches to Second Language Reading*. New York: CUP, 101-13.
- Cleary, C., Holden, B. & Cooney, T. (2007). *Top-Up Listening*. Tokyo: Abax ELT Publishers.
- Cross, J. (2009). Effects of listening strategy instruction on news videotext comprehension. *Language Teaching Research*, 13(2), 151-176.
- Ellsworth (2006). Developing an English listening course for low proficiency Japanese university students. *The Journal of Chiba University of Commerce*, 43(4), 77-91.
- Field, J. (2008). *Listening in the Language Classroom*. Cambridge: CUP.
- Flowerdew, J. & Miller, L. (2005). *Second language listening*. Cambridge: CUP.
- Higuchi, A. (1998). Suki-ma riron ni motozuita listening strategy- oral communication (B) wo chusin to shite [Listening Strategy based on schma theory - focus on oral communication (B)-]. *The bulletin of Faculty of Education Kagoshima University, Part of Education Science* 49, 145-158.
- Hinkel, E. (2006). Current perspectives on teaching the four skills. *TESOL Quarterly*, 40, 1, 109-131.
- Kadota, S. (2006). *Reading and phonologocal processes in English as a Second Language*. Tokyo: Kuroshiro Shuppan.
- Kadota, S. (2012). *Shadowing, reading aloud, and science of English acquisition*. Tokyo: Cosmopier.
- Lynch, T., & Mendelsohn, D. (2002). Listening. *An Introduction to Applied Linguistics*. London : Arnold, 193-210.
- O'Malley, J. M. & Chamot, A. U. (1990). *Learning strategies in second language acquisition*. Cambridge: CUP.
- Ross, R. N. (1975). Ellipsis and the structure of expectation. *San Jose State Occasional Papers in Linguistics* 1, 183-91.
- Rost, M. (1990). *Listening in language learning*. Essex: Longman.
- Rost, M. (2011). *Teaching and researching listening. Second edition*. Edinburgh: Pearson Educational Limited.
- Rubin, J. (1994). A review of second language listening comprehension research. *The Modern Language Journal*, 78, 2, 199-221.
- Vandergrift, L. (1997). The comprehension strategies of second language (French) listeners: a descriptive study. *Foreign Language Annals*, 30, 3, 387-409.
- Vandergrift, L. (2007). Recent developments in second and foreign language listening comprehension research. *Language Teaching*, 40, 191-210.
- Vandergrift, L. & Tafaghodtari, M. H. (2010). Teaching L2 learners how to listen does make a difference: an empirical study. *Language Learning* 60, 2, 470-497.
- Vandergrift, L. & Goh, C. (2012). *Teaching and learning second language listening: Metacognition in action*. New York: Routledge.
- Wilson, J. J. (2008). *How to teach listening*. Edinburgh: Pearson Educational Limited.