

# **Dictation: Is it really effective for language teaching?**

**TAKEUCHI, Osamu**

# Dictation: Is it really effective for language teaching?\*

TAKEUCHI, Osamu

## Introduction

As a testing method, dictation has been confirmed to be reliable and valid by many researchers. Based on the summary of empirical studies, Oller (1979), for example, considers dictation to be a type of integrative test, a test which requires a learner to use several language skills at the same time. Heaton (1988: 17) says, "the integrated skills involved in tests of dictation include auditory discrimination, the auditory memory span, spelling, the recognition of sound segments, a familiarity with the grammatical and lexical patterning of the language, and overall textual comprehension," and claims, "dictation tests can prove good predictors of global language ability." Dictation is, in a sense, an established means of testing.

As a teaching method, however, dictation lacks empirical grounds. Harris (1969), for example, insists that dictation can be a useful pedagogical device but offers little empirical support. Davis and Rinvolucri (1988) also mention that dictation is an effective measure of teaching, and provides us with various methods of dictation. However, empirical support for their claim of effectiveness is slight.

One example of empirical data was provided by Yoshida (1978, 1981, 1984). He found statistically significant positive correlations between dictation and the results of listening tests, and, based on these findings, claims that dictation can be a good teaching device. His findings do indicate that dictation is a good predictor of learners' listening ability. However, a good predictor does not necessarily turn out to be a good teaching device. The literature as a whole, therefore, shows that dictation has been utilized by foreign language instructors for teaching with precarious empirical support.

## Experiment and Procedure

### Purposes

The purposes of this experiment are 1) to determine whether dictation is in fact effective for language teaching, and 2) to confirm the relative effects of the three different methods of dictation.<sup>1</sup>

### Subjects

Subjects were 207 first-year students of a women's junior college, Kyoto, Japan. Their major was English language. Students who had stayed in an English speaking country for more than one year were excluded from the subject group.

## Grouping

Subjects were divided into three treatment groups based on the results of the listening section of the CELT, Form B (Harris and Palmer, 1986: henceforth "pre-test"), conducted in April, the beginning of the school year. Homogeneity among the three groups at the beginning of the experiment was confirmed ( $F_{2,204} = .22, p = .80$ ; see Table 1 for descriptive statistics). Each group consisted of two classes, and each class was made up of approximately 35 students. Group A was made up of Classes 3 and 5, and Group B was made up of Classes 1 and 2, while Group C consists of Classes 4 and 6. To minimize the instructors' idiosyncratic variables, Classes 1 and 4 were taught by Instructor A, Classes 2 and 5 by Instructor B, and Classes 3 and 6 by Instructor C. See Table 2 for a summary.

**Table 1. Descriptive Statistics for Pre-Test**

Group	M	SD	N
A	47.76	10.93	72
B	48.73	11.56	66
C	48.93	10.85	69

**Table 2. Groups, Classes, Instructors**

Group	Number	Class	Instructor
A	72	3,5	C,B
B	66	1,2	A,B
C	69	4,6	A,C

For further analysis, all subjects were divided into two proficiency groups according to their performances on the pre-test: Upper Ability Group and Lower Ability Group. The Upper Ability Group included 97 subjects whose score on the pre-test was more than 51. A total of 110 subjects was included in the Lower Ability Group, in which each subject scored less than 45 on the pre-test. Heterogeneity among these two groups was confirmed ( $F_{1,205} = 686.07; p < .001$ ). See Table 3 for the statistics.

**Table 3. Descriptive Statistics for Upper and Lower Ability Groups**

Group	M	SD	N
Upper	58.77	6.43	97
Lower	39.37	4.11	110

## Material

The material used in this experiment was an American movie entitled *Ferris Buller's Day Off*, which depicts a high-school boy's truancy from school and his ensuing mischievous behavior. This material was highly popular with the subjects of this experiment and, therefore, seemed to motivate them to study hard. See, for example, Edasawa, Takeuchi, and Nishizaki(1990) and Takeuchi, Edasawa, and Nishizaki(1990) for the relationship between materials and motivation.<sup>2</sup>

## Treatment

During the 13-week period between April and July, subjects in each class took a 45-minute lesson twice a week in a language laboratory. In each lesson, subjects in Group A watched a part of the material several times, and filled in the blanks on the scenario (simple dictation: see Sample 1 for a portion of the scenario used). Subjects in Group B did the same, but were given a scenario in which each blank was accompanied by Japanese translation (dictation with translation: see Sample 2). Students in Group C were given a scenario in which the number of the words deleted was indicated for each blank (clued dictation: see Sample 3). The number of times the subjects watched a portion of the film was controlled to an extent, although some deviation did exist. The instructor's explanation of each portion was pre-determined. We also controlled for other activities used in the classroom, in addition to controlling for variables concerning teaching style, to the best of our ability.

### Sample 1. A Portion of the Scenario for Simple Dictation

Radio: \_\_\_\_\_

\_\_\_\_\_

Right now 75 at the lake front, 74 at Midway, 73 at O'Hare. And now up in the sky...

Ms. Bueller (Ms. B): Ferris, Ferris, Tom!

Mr. Buller (Mr. B): \_\_\_\_\_

Ms. B: Oh, it's Ferris.

Mr. B: What, what's wrong?

Ms. B: What's wrong? \_\_\_\_\_

Mr. B: Ferris?

Ms. B: \_\_\_\_\_

and he's seeing spots....

### Sample 2. A Portion of the Scenario for Dictation with Translation

Radio: \_\_\_\_\_

シカゴの天気は良好です

\_\_\_\_\_

気温は70度後半になるでしょう。

Right now 75 at the lake front, 74 at Midway, 73 at O'Hare. And now up in the sky...

Ms. Bueller (Ms. B): Ferris, Ferris, Tom!

Mr. Bueller (Mr. B): \_\_\_\_\_

どうしたんだい？

Ms. B: Oh, it's Ferris.

Mr. B: What, what's wrong?

Ms. B: What's wrong? \_\_\_\_\_

お願いだからフェリスをみてちょうだい、あなた。

Mr. B: Ferris?

Ms. B: \_\_\_\_\_

熱はないけどお腹が痛むって

and he's seeing spots....

### Sample 3. A Portion of the Scenario for Clued Dictation

Radio: \_\_\_\_\_ (7)

\_\_\_\_\_ (8)

Right now 75 at the lake front, 74 at Midway, 73 at O'Hare. And now up in the sky...

Ms. Bueller (Ms. B): Ferris, Ferris, Tom!

Mr. Bueller (Mr. B): \_\_\_\_\_ (3)

Ms. B: Oh, it's Ferris.

Mr. B: What, what's wrong?

Ms. B: What's wrong? \_\_\_\_\_ (7)

Mr. B: Ferris?

Ms. B: \_\_\_\_\_ (11)

and he's seeing spots....

### Sample 4. Full Script

Radio: It is a beautiful day in Chicago.

Today's temperature's expected to reach the upper seventies. Right now 75 at the lake front, 74 at Midway, 73 at O'Hare. And now up in the sky...

Ms. Bueller (Ms. B): Ferris, Ferris, Tom!

Mr. Bueller (Mr. B): What's the matter?

Ms. B: Oh, it's Ferris.

Mr. B: What, what's wrong?

Ms. B: What's wrong? For Christ's sake, look at him, honey.

Mr. B: Ferris?

Ms. B: He doesn't have a fever, but he says his stomach hurts, and he's seeing spots...

In July, the listening section of the CELT, Form B (henceforth "post-test") was administered.

ed again to all subjects. The 13-week period between the pre-test and post-test seems to be sufficiently long for the subjects to forget the questions of the test.

### Analysis

Table 4 shows the descriptive statistics of the post-test, while Table 5 indicates those of the difference in scores between the pre- and the post-tests. The data was analyzed, using ANOVA in the SPSS package of statistical programs. The analysis confirmed, in all the three groups, a statistically significant improvement between the performances on the two tests ( $F=132.38$ ;  $p < .001$ ). Group differences, however, were not found ( $F(2, 204) = .93$ ;  $p = .34$  n.s.).

**Table 4. Descriptive Statistics for Post-Test**

Group	M	SD	N
A (simple dic.)	56.11	12.16	72
B (dic. with trans.)	54.95	12.58	66
C (clued dic.)	55.94	10.88	69

**Table 5. Descriptive Statistics for Differences between Pre- and Post-Tests**

Group	M	SD	N
A (simple dic.)	8.35	9.25	72
B (dic. with trans.)	6.23	8.12	66
C (clued dic.)	7.01	10.20	69

Then, a MANOVA analysis was conducted by using Treatments (Groups: Tr), Proficiency (Upper vs. Lower: P), and Time (pre-post difference: Ti) as factors. The results of the analysis can be seen in Table 6.

Table 6 suggests that there existed an interaction among Tr, P, and Ti, though the figure itself ( $F=2.31$ ;  $p = .10$ ) did not quite reach the probability level of .05. Since the F figure was close to the .05 critical value, a further ANOVA analysis was conducted. The analysis (Table 7 and Figure 1) confirmed that the interaction was mainly due to the significant improvement of the subjects in treatment Group C of the Lower Ability Group. Figure 1 also shows that the subjects in the Lower Ability Group who were in treatment Group B improved less than the subjects of the same Lower Ability Group who were in treatment Groups A and C. No difference in improvement was found in Upper Ability Group subjects regardless of the treatment group they belonged to.

### Discussion

Dictation was confirmed to be an effective teaching method in this experiment. Subjects in all the three groups show statistically significant improvement in foreign language listen-

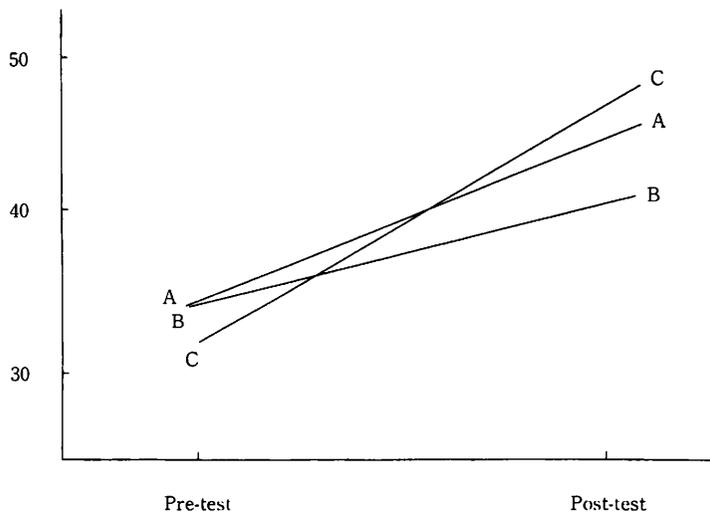
**Table 6. MANOVA Analysis**

	SS	DF	MS	F	<i>p</i>
Between-Subject Effects					
Within Cells	15655.31	201	77.89		
Constant	1137792.32	1	1137792.31	14608.22	.00*
Proficiency (P)	29410.40	1	29410.40	377.60	.00*
Treatment (Tr)	22.56	2	11.28	.14	.87
P by Tr	308.89	2	154.44	1.98	.14
Within-Subjects Effects					
Within Cell	7878.49	201	39.20		
Time (Ti)	5188.70	1	5188.70	132.38	.00*
P by Ti	653.60	1	653.60	16.68	.00*
Tr by Ti	72.21	2	36.11	.92	.40
P by Tr by Ti	181.01	2	90.50	2.31	.10 <sup>+</sup>

\*:  $p < .001$  +:  $p = .10$

**Table 7. ANOVA Analysis**

Group	F	<i>p</i>
A (simple dic.)	2.59	.11
B (dic. with trans.)	2.06	.16
C (clued dic.)	15.52	.00



**Figure 1. Gains between Pre- and Post-Tests**

ing comprehension. This result lends empirical support to the language teachers using this method.

An aptitude treatment interaction (ATI: Cronback and Snow, 1977) was found between the different methods of dictation and the improvement of the subjects' proficiency. One of the reasons the ATI was found only in the Lower Ability Group may be that the material used in this experiment was relatively easy for the subjects in the Upper Ability Group. Informal interviews conducted later with some Upper Ability Group subjects support this inference. If the material used had been difficult enough for the Upper Ability Group subjects, the ATI might have been found in both groups.

In the Lower Ability Group, clued dictation was found to be more effective than the other two types. This may be because the predicting-and-testing process plays a significant role in foreign language listening comprehension (Kohno, 1994). The number of the words deleted in clued dictation may affect the predicting-and-testing process and thereby listening comprehension, while dictation with translation in the subjects' native language can disrupt the process and hinder comprehension. As to simple dictation, it offers comparatively fewer clues for the predicting-and-testing process and, therefore, does not help subjects with relatively low proficiency comprehend the material. The 13-week cumulative effect of this facilitation and hindrance may have caused the difference in the effects of the various methods of dictation.

Dictation seems to require students to make the most of the knowledge available to them, and activate the predicting-and-testing process of listening comprehension. This activation then leads to better understanding of input, which is essential to the improvement of foreign language ability. Dictation, therefore, can prove to be an effective instructional device for foreign language listening.

## **Conclusion**

Before concluding, some shortcomings in this experiment should be pointed out. First, although maximum efforts were made to control the variables concerning the teaching style in each class, there existed some differences. These differences, therefore, might have exerted some influence on the results of this experiment. In this connection, other intervening variables such as the hours students spent at home listening to English and the effects of other English courses the subjects took may have affected the results reported above. Second, the period of the experiment, 13 weeks, was rather brief. This briefness may have had some influence upon the results obtained. Third, the number of the subjects, 207, was small when we divided the number into several treatment/proficiency groups. This small number might have caused some anomaly in the data analyzed. Fourth, no control group was established. This is because the experiment was conducted in an educational institution as part of the required curriculum, so it was impossible for us to have a control group.

With these shortcomings in mind, the following summary is in order: The experiment reported above proves that dictation is effective in foreign language teaching. It also determines that clued dictation is effective especially for learners whose proficiency is relatively low. The reason may be that the clues in clued dictation facilitate the predicting-

and-testing process of listening comprehension. Dictation with translation is less effective than clued and simple dictations in the lower proficiency group. This may be due to the confusing impact of native-language translation on the predicting-and-testing process. As to the upper proficiency group, no difference was found concerning the effects of the different dictations. This may be due to the degree of difficulty of the material used in the experiment.

## Notes

\* This article is a revised version of the paper presented by Professor Namie Saeki of Doshisha Women's College and the present author at the 34th LLA (Language Laboratory Association of Japan) Annual National Conference held at Konan Women's University, Kobe, Japan in 1994. Although the research is a joint effort with Professor Saeki, the analysis and interpretation presented in this article are solely mine, and, therefore, the whole responsibility regarding errors and mistakes should be borne by the present author. Professor Saeki's analysis and interpretation of this research may appear in another journal.

I would like to express my appreciation to my former colleagues, Professor N. Saeki for her cooperation in this study, Professor H. Mine for his help in statistical analysis, and Professor B. Susser for his constructive criticism on the draft of this paper.

1. Two major types of dictation, full (standard) and partial (spot), have often been used by language teachers (Oller, 1979). In this experiment, the latter type of dictation (i.e., partial dictation) and its three different varieties are our focus.
2. See also Edasawa, Takeuchi, and Nishizaki (1992).

## References

- Cohen, A.D. 1980. *Testing language ability in the classroom*. Rowley: Newbury House.
- Cronback, L. J. and R.E. Snow 1977. *Aptitude and instructional method*. New York: Irvington.
- Davis, P. and M. Rinvulcri 1988. *Dictation: New methods, new possibilities*. Cambridge: Cambridge Univ. Press.
- Edasawa, Y., O. Takeuchi, and K. Nishizaki 1990. Use of films in listening comprehension practice. *The IALL Journal of Language Learning Technologies*, 23: 3, 21-34.
- Edasawa, Y., O. Takeuchi, and K. Nishizaki 1992. The effects of teaching methods in using films for listening comprehension practice. *Language Laboratory*, 29, 53-70.
- Harris, D.P. 1969. *Testing English as a second language*. New York: McGraw-Hill.
- Harris, D.P. and L.A. Palmer 1986. *CELT: A comprehensive English language test for learners of English*. New York: McGraw-Hill.
- Heaton, J.B. 1988. *Writing English language tests* (New edition). London: Longman.
- Itakura, T., F. Ohosato, and F. Miyahara 1985. On dictation: Theoretical background, validity as a measure, and classification of errors. (in Japanese) *Language Labora-*

- tory, 22, 3-25.
- Kohno, M. 1992. Hitowa otononagarewo donoyouni rikaisurunoka: Eigo ni yoru communication no kihon mondai. (How human being comprehends the stream of sounds: Basic issues in foreign language communication.) *LLA Kansai Chapter Monograph Series IV*, 13-49.
- Oller, J.W. 1979. *Language tests at school*. London: Longman.
- Rost, M. 1990. *Listening in language learning*. London: Longman.
- Takeuchi, O., Y.Edasawa, and K.Nishizaki 1990. Do films improve EFL students' listening comprehension ability? *Language Laboratory*, 27, 81-98.
- Yoshida, K. 1978. Relation between listening comprehension and dictation in closed circuit TV. *Bulletin of Fukuoka University of Education*, 28: 4, 151-164.
- Yoshida, K. 1981. Relation between cloze test and consolidation in respect to dictation. *Bulletin of Fukuoka University of Education*, 30: 4, 273-289.
- Yoshida, K.(ed.) 1984. *Eigo no risuningu*. (*Listening in English language teaching*.) Tokyo: Taishukan.