Understanding Sentenced-Based Partial Dictation

Hirokatsu KAWASHIMA

要　約

文レベルのパーシャル・ディクテーションの性質を把握するための調査を行った。パーシャル・ディクテーションに用いた英文はすべて15語から成る英文で、聞き取った英文を基に、文頭、文中、文尾の括弧の中に対応する単語（それぞれ3単語ずつ）を書き込む形式をとった。調査の結果、1）文中のディクテーションのパフォーマンスが最も低く、また、2）総合的リスニング能力との関係において、文中のディクテーション能力が最も重要な役割を果たしている可能性が高い、こと等が明らかになった。

Background

In general dictation has been widely used in EFL listening classrooms although it is not the most important element in teaching listening, and most listening textbooks more or less include dictation exercises. The significance of using dictation in the classroom is understandable, because dictation can be easily prepared and conducted in a relatively simple way in the classroom. It may even be possible for instructors to make some English sentences for dictation impromptu in order to increase the variety of classroom activities of the day. More fundamentally, it must be pointed out that dictation has the role of raising learners' consciousness of listening learning in the sense that they can view their listening performance and check their weaknesses instantly. Nation & Newton (2009: p.59) claim that "dictations help language learning by making learners focus on language form of phrase and clause level constructions, and by providing feedback on the accuracy of their perceptions."

In addition, dictation involves aspects of assessment, by which learners' listening proficiency can be measured and assessed. Buck (2001: pp.77-78), for example, supports this possibility by claiming 1) that "dictation works in a number of ways, depending upon how long the segments are and how much they challenge the test takers," 2) that "if the segments are very short and do not challenge the test taker, then dictation is probably testing little more than the ability to recognize simple elements," but 3) that "if the segments are a little longer, it will be testing understanding on a local, literal, linguistic level."

It is unknown how much the above significance of using dictation is recognized in actual classrooms, but all that can certainly be maintained is that the proper and systematic use of dictation may enhance listening learning and make it possible for instructors to assess their learners' listening proficiency. It must be noted, however, that the present

moment sees dictation understood mostly at a general level and that much of the nature of using dictation remains unclear, as a review of the literature shows (e.g., Hio 1983, Sugawara 1999, Rost 2002, Flowerdew & Miller 2005, and Field 2008). For example, the following aspects of dictation have not yet been fully investigated:

1) relationships between general listening proficiency and sentence-based partial dictation
2) relationships between general listening proficiency and constituents of sentence-based partial dictation

Understanding these kinds of relationships will help clarify what *the proper and systematic use of dictation* should be in the classroom, but a number of investigations need to be conducted taking types/lengths/difficulties of passages/sentences into consideration, and their results must be integrated.

Current Study

The author launched a research project in 2007 in an attempt to elucidate the nature of listening sub-skills, in which a series of empirical studies have been conducted. One of them, Kawashima (2009), preliminarily explored the above relationship 1) and reported, for example, that sentence-based partial dictation was significantly related in performance to general listening proficiency ($r = .55$, $p = .000$). The current study, which is also a basic investigation conducted as part of the research project, aims at exploring the above relationship 2) employing the same basic research framework.

The research designs are summarized below, and then some of the major findings are reported.

1. Research Design

1.1 Research Questions

The current study involves two main research questions:

1) Which constituent of sentence-based partial dictation is the most/least difficult to comprehend and dictate?

2) Which constituent of sentence-based partial dictation is the most/least related to general listening proficiency?

1.2 Materials

1.2.1 General Listening Proficiency

In order to measure the subjects' general listening proficiency (referred to as hereinafter as GLP), two sets of listening sections of *STEP Grade 2* tests were used, which had been originally designed to match the level of high school graduates in general and administered on October 8, 1998 and June 18, 2000. Each set had 20 four-option multiple-choice test items, and 40 test items were used in total.

1.2.2 Sentence-Based Partial Dictation

In order to measure the subjects' performance of sentence-based partial dictation at the sentence constituent level (referred to hereinafter as SPD), 12 fifteen-word sentences were prepared, which involved three dictation blanks in each of the RIGHT/MIDDLE/LEFT parts of the sentence. The following SPD) is an example:

SPD) (Misaki) (is) (planning) to study the (history) (of) (England) which is full (of) (interesting) (stories).

All the sentences were recorded onto CD at a self-selected normal speaking rate by a male native speaker of English.

1.3 Subjects

79 first-year students of the general education course at a university in Japan participated in the current investigation.

1.4 Data Collection and Procedure

The investigation was conducted during regular English classes in July, 2008, the main goal of which was to improve the learners' overall listening proficiency.

1.4.1 Measuring GLP

The subjects took two GLP tests at intervals in order to grasp their general listening proficiency and monitor their progress periodically: at the beginning of June and at the end of July. About 20 minutes were allocated for each test, after which the distributed

computer-scored investigation sheets were collected, and then the subjects immediately checked with their sub-investigation sheet whether their answers were correct and grasped their general listening proficiency by the totaled score.

1.4.2 Measuring SPD

The subjects took four SPD tests as part of their regular classroom listening activities, in which they carried out sentence-based partial dictation tasks and checked their weaknesses in listening to English. About 10 minutes were allocated for each class, after which the subjects immediately checked their answers on the distributed investigation sheets and grasped their SPD performance.

2. Scoring and Processing of the Data

All the investigation sheets were collected when each class was over, after which the raw data were scored, examined, and processed for analysis.

2.1 Scoring

With regard to GLP tests, the computer-scored investigation sheets were read and processed by an optical mark reader (*SR-3500*, Sekonic) and a mark reader computer software (*SS kun II*, Software for Education), in which the correctness of each test item was provided with the item scores (0, 1) representing correct and incorrect answers, respectively.

As far as sentence-based partial dictation tasks were concerned, first, the correctness of each dictation item on the investigation sheets was carefully checked with the same item scores (0, 1), and then the total score of each of the three sentence constituent parts of the fifteen-word sentences (RIGHT/MIDDLE/LEFT) was calculated with the item scores (0, 1, 2, 3). If the blanks are filled by a subject as in the following, for example, the learner gets six points in total [2 (RIGHT) + 3 (MIDDLE) + 1 (LEFT) = 6]:

Example Subject's Answers:
(Misaki) (is) (?) to study the (history) (of) (England) which is full (?) (interesting) (?).

Correct Answers:
(Misaki) (is) (planning) to study the (history) (of) (England) which is full (of) (interesting) (stories).

Slight spelling mistakes were expected to be made (e.g., *planing*), but the current study did not regard them as incorrect answers unless they would cause serious semantic confusion.

2.2 Examining Internal Consistency Reliability

The scored data were then examined in terms of internal consistency reliability. The *Cronbach Alpha* coefficients of the two GLP tests (the total number of test items is 40) and those of SPD tests (the total number of test items for each of RIGHT, MIDDLE, and LEFT constituents of the sentence is 12) were measured.[1] Table 1 presents the results:

Table 1: Internal Consistency Reliability by *Cronbach Alpha* Coefficient

	GLP	LCS	MCS	RCS
Number of Test Items	40	12	12	12
Cronbach Alpha Coefficient	.70	.68	.78	.63

LCS: Left Constituent of the Sentence　MCS: Middle Constituent of the Sentence　RCS: Right Constituent of the Sentence

It is generally assumed that *Cronbach Alpha* coefficient should exceed at least 0.7 for reliable analysis, but the current study more or less takes Dörnyei (2007)'s view that "somewhat lower *Cronbach Alpha* coefficients are to be expected" owing to "the complexity of the second language acquisition process (p. 207), and holds that although *Cronbach Alpha* coefficients of the LEFT and RIGHT constituents of the sentence of the SPD tests are 0.68 and 0.63, respectively, failing to reach 0.70, those tests may be employed to a certain degree while paying attention to their limits in terms of internal consistency reliability.

2.3 Examining Normal Distribution

The data of the GLP and SPD tests were also examined in terms of normal distribution, upon which the statistical analyses of the current study were based. *Shapiro-Wiki* tests, whose α value had been set at 0.01, were conducted for this examination.[2] Table 3 presents the results:

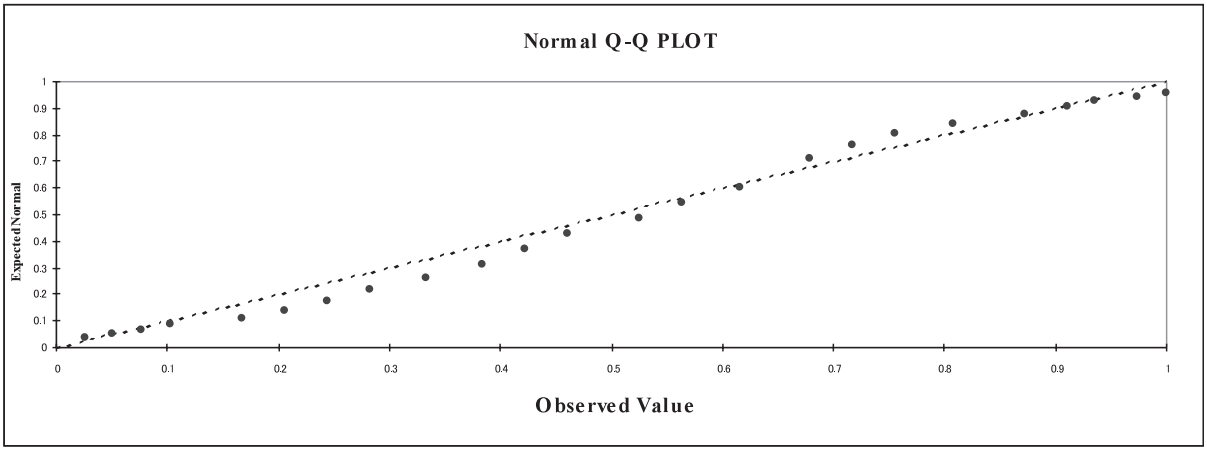
Table 2: Normal Distribution of the GLP and SPD Data (α = .01)

	GLP	LCS	MCS	RCS
W	.98	.98	.96	.98
p-value	.19	.20	.02	.20

LCS: Left Constituent of the Sentence　MCS: Middle Constituent of the Sentence　RCS: Right Constituent of the Sentence

It is statistically shown in this table that each p-value is greater than 0.01 and that the data of the GLP and SPD tests are normally distributed, but attention must be directed at the data of the SPD test (the middle constituent of the sentence), because p-value barely

exceeds 0.01. In order to verify this pure statistical result, the current study examined the normal distribution of this data by looking at the following Normal Q-Q Plot:



The straight line presents what the data would look like if it were perfectly normally distributed, and the actual data is represented by the small squares plotted along this line. The closer the small squares are located to the line, the more normally distributed the data is plotted. The current study assumes that most of the small squares fall along the line though not perfectly, and that the data of the SPD test (the middle constituent of the sentence) is normally distributed as a whole.

3. Data Analysis

The pre-examined data above were then processed for analysis.[3)]

3.1 Descriptive Statistics

First, the minimum, maximum, mean score and standard deviation of the processed data of the GLP and SPD tests were calculated. The results are presented in Table 3:

Table 3: Descriptive Statistics for the GLP and SPD Tests

	GLP	LCS	MCS	RCS
Min.	7.00	13.00	2.00	11.00
Max.	28.00	34.00	26.00	33.00
Mean	18.21	23.54	14.27	23.08
S.D.	5.25	5.06	6.74	4.94

LCS: Left Constituent of the Sentence MCS: Middle Constituent of the Sentence RCS: Right Constituent of the Sentence

This table shows that the subjects' performance of the SPD test may vary with the

three constituents of the sentence, and that it may be the lowest in the middle constituent (mean: 26.00).

3.2 Dictation Performance in the Three Constituents of the Sentence

In order to verify the above tentative results with statistic significance, the current study first examined the data employing one-factor repeated measures ANOVA. The results are presented in Table 4:

Table 4: One-Factor Repeated Measures ANOVA for the SPD Data

Source	df	Sum of Squares	Mean Square	F-Value	P-Value
Sentence Constituent	2	4256.39	2128.19	159.75	.00
Subjects	76	32126.78	32126.78		
Residual	78	1797.55	23.35		

This table shows that the subjects' performance of the SPD tests varies significantly with the three constituents of the sentence [(F (2, 78) = 159.75, p =.00]. Multiple comparisons by *Contrast*, which could be used for repeated measured data, were then conducted in order to examine whether the subjects' performance of the SPD test was the lowest in the middle constituent of the sentence. The results are presented in Table 5:

Table 5: Multiple Comparisons by *Contrast* for the SPD Data

Combination	Difference: Absolute Value	MSe	P-Value
LCS−MCS	9.23	.57	.00
MCS−RCS	8.81	.65	.00
LCS−RCS	0.47	.52	1.00

LCS: Left Constituent of the Sentence　MCS: Middle Constituent of the Sentence　RCS: Right Constituent of the Sentence

This table shows 1) that there is a statistically significant difference in performance between the left and middle constituents of the sentence (MSe = 0.57, p = .00) and between the middle and right constituents of the sentence (MSe = 0.65, p = .00), and consequently 2) that the subjects' performance of the SPD tests is certainly the lowest in the middle constituent of the sentence. Figure 2) illustrates the results on a horizontal line:

Figure 2: Dictation Performance in the Three Constituents of the Sentence

Low　MIDDLE CONSTITUENT > RIGHT CONSTITUENT = LEFT CONSTITUENT　High

3.2 Relationships between General Listening Proficiency and Sentenced-Based Partial Dictation

In order to explore relationships between general listening proficiency and sentence-based partial dictation, their simple correlation coefficients were computed. Table 6 presents the results:

Table 6: Simple Correlation Matrix for the GLP and SPD Data

	GLP	LCS	MCS	RCS
GLP	1	.41*	.55*	.45*
LCS		1	.68*	.58*
MCS			1	.55*
RCS				1

LCS:Left Constituent of the Sentence MCS:Middle Constituent of the Sentence RCS:Right Constituent of the Sentence *<.01

This table shows 1) that partial dictation is significantly related in performance to general listening proficiency at all the three levels of sentence constituent, and 2) that partial dictation in the middle constituent of the sentence is most significantly related to general listening proficiency ($r = 0.55$, $p < .01$). It must be noted, however, that these relationships may be superficial and unstable, because it is possible that they are mere reflections of some unrevealed third relationships (known as pseudo correlations), subsistent in the nature of sentence-based partial dictation. In order to closely examine these relationships between general listening proficiency and sentence-based partial dictation, partial regression analysis was conducted, in which partial regression coefficients of direct relationships between two variables were computed, thereby eliminating the influences of the rest. Table 7 presents the results:

Table 7: Partial Correlation Matrix for the GLP and SPD Data

	GLP	LCS	MCS	RCS
GLP	1	-.01	.36*	.20
LCS		1	.50*	.32*
MCS			1	.17*
RCS				1

LCS:Left Constituent of the Sentence MCS:Middle Constituent of the Sentence RCS:Right Constituent of the Sentence *<.01

This table shows that while it is unrelated to general listening proficiency in the left and right constituents of the sentence ($r = -0.01$, $p < .01$, and $r = 0.20$, $p > .01$, respectively), partial dictation is significantly related to general listening proficiency at the middle constituent ($r = 0.36$, $p < .01$). Considering these partial regression coefficients solely and the nature of statistical significance, the relative strength of relationships in performance between

general listening proficiency and partial dictation in the three sentence constituents could be delineated on a horizontal line as in Figure 3:

Figure 3: Relative Strength of Relationships between GLP and SPD in the Three Sentence Constituents

strong MIDDLE CONSTITUENT > RIGHT CONSTITUENT > LEFT CONSTITUENT weak

4. Summary & Discussion

The above analyses have made clear two major points with regard to sentence-based partial dictation.

4.1 Dictation Performance in the Three Sentence Constituents

First, it has been found that there is little difference in the subjects' dictation performance between the right and left constituents of the sentence, and that such a performance is the lowest in the middle of the sentence. In theory, several other possibilities can be considered regarding the difficulty of dictation performance in the three constituents of the sentence. It is quite conceivable, for example, that dictation performance may be the same in the three constituents of the sentence, or that dictation performance may be the lowest in the right constituent of the sentence.

It is difficult to correctly and comprehensively interpret this research finding with the limited data, but one possible interpretation can be made from the complicated nature of partial dictation itself. First, let us look at the internal aspects of the SPD tests. In these tests, subjects are required in general to carry out four main cognitive activities: to process spoken English sentences, to comprehend them, to store them in their working memory, and to fill in the blanks. It can be claimed that these cognitive activities are complicated in themselves and become even more complicated in interactive situations, and that successful partial dictation depends upon how well each activity can be conducted and how collaboratively each activity can work together with the others.

The first research finding, if viewed from this perspective, may be interpreted as follows:

1) To dictate the middle constituent of the sentence may be difficult, because test takers have to pay attention to its relationships with both right and left constituents in processing and comprehending the whole sentence. This is considered to impose a heavy burden upon the test takers' working memory, causing poor dictation performance.

2) To dictate the left and right constituents of the sentence may be less difficult, because in general test takers pay less attention to the other constituents in processing and

comprehending the whole sentence. This is considered to place less burden upon the test takers' working memory, bringing about good dictation performance.

4.2 Dictation in the Three Sentence Constituents and General Listening Proficiency

Next, it has been found in the above analyses that partial dictation in the middle constituent of the sentence is most related in performance to general listening proficiency. It is likewise difficult to correctly and comprehensively interpret this research finding with the limited data, but based upon 1) and 2) above, the following interpretation may be possible:

3) Dictation in the middle constituent of the sentence is most related to general listening proficiency, because it may embrace the highest level of processing and comprehending spoken sentences, in which test takers must listen to sentences "hard and carefully."

4) Dictation in the right and left constituents of the sentence is less related to general listening proficiency, because a surface, lower level of processing and comprehending spoken sentences may be just conducted here, where test takers do not have to listen to sentences "hard and carefully."

Concluding Remarks

Partial dictation is very useful in the classroom, because it can be scored much easier than full dictation. It must be noted, however, that little has been understood about the nature of sentence-based partial dictation. The current study could offer some insights into this nature in the sense that it has made clear that performance of sentence-based partial dictation and its relationships with general listening proficiency may vary across sentence constituents. Based upon these results, it might be possible, for example, to tell instructors not to overestimate his/her students' performance of sentence-based dictation as far as the beginning of the target sentence is concerned.

The results of the current study, however, should be viewed as tentative and must be reexamined from several experimental perspectives. Firstly, as was stated above, more control of such experimental factors as internal consistency reliability is needed, and careful attention must also be directed at the validity of general listening proficiency. Secondly, more subjects must be used, and their backgrounds should also be considered in terms of grammatical and writing proficiency. Thirdly, the length and difficulty of target sentences and the number of dictation words must be taken into account more systematically. Dictation performance in the right constituent of the sentence may be lowered if the number of dictation words increases, for instance. Finally, the speed at which the subjects listened to target sentences must be controlled. A higher speed is considered to make it more difficult

for learners to process, comprehend, and dictate sentences. Further studies, taking these points into account, will take us closer to a complete map of the nature of sentence-based partial dictation.

Notes

1) *SPSS* (Version 16.0: SPSS Inc.) was used for this examination.
2) *XLSTAT-PRO* (Version 2009: Addinsoft Inc.) was used for this examination.
3) *EXCEL STATISTICS* (Version 5.0: Esumi Inc.), *TAHENRYOU-KAISEKI* (Version 5.0: Esumi Inc.), and *SPSS* (Version 16.0: SPSS Inc.) were used for the analyses.

Acknowledgements

This study is supported by Grant-in-Aid for Scientific Research (1952539) from the Japan Society for the Promotion of Science. The author is grateful to Dr. Richard Hodson, an English lecturer at University of Nagasaki, Siebold for his full support for making the SPD tests and proofreading the first & second drafts.

References

Bachman, L. F. (2004). *Statistical Analyses for Language Assessment*. Cambridge University Press.

Buck, G. (2001). *Assessing Listening*. Cambridge University Press.

Dörnyei, Z. (2007). *Research Methods in Applied Linguistics*. Oxford University Press.

Field, J. (2008). *Listening in the Language Classroom*. Cambridge University Press.

Flowerdew, J. & Miller, L. (2005). *Second Language Listening: Theory and Practice*. Cambridge University Press.

Hio, Y. (1983). Dictation as a Testing Device -Based on Error Analysis-. *Language Laboratory,* 20, 2-16.

Icho, H. (2008). Understanding the Differences between Successful and Unsuccessful Japanese High School EFL Listeners: Focusing on Explanatory Factors. *ARELE*, 19, 1-10.

Kawashima, H. (2009). Understanding Sub-Skills of Listening from the Perspective of Classroom Diagnostic Assessment. *Conference Book: Interface between National Tests and Classroom Assessment*. The 4th Annual KELTA Conference: 2009 International Conference, 102-106

Mackey, A. & Gass, S. M. (2005). *Second Language Research: Methodology and Design*. Mahwah, N. J.: Lawrence Erlbaum.

Nation, I.S.P. & Newton, J. (2009). *Teaching ESL/EFL Listening and Speaking*. Routledge, Taylor and Francis.

Rost, M. (2002). *Teaching and Researching Listening*. Longman.

Seliger, H. W. & Shohamy, E. (1989). *Second Language Research Method*. Oxford: Oxford

University Press.

Sugawara, Y. (1999). Dictation and Listening Comprehension -Does Dictation Promote Listening Comprehension?-. *Language Laboratory,* 36, 33-50.

Takeno, J. & Takatsuka, S. (2007). Factors Affecting Listening Comprehension Ability of Japanese Learners of English. *ARELE*, 18, 1-10.