

# DO INDIRECT TESTS OF ENGLISH INTONATION

## MEASURE STUDENTS' ABILITY IN PERFORMING ENGLISH NUCLEUS PLACEMENT?

### 1. INTRODUCTION

Indirect tests (pencil and paper tests) on pronunciation are economical to make and easy to mark. They are frequently used in the Japanese SAT (the Japanese version of the American Scholastic Aptitude Test, conducted by the National Centre for University Entrance Examinations), mid-term and final tests at school, and in other kinds of tests administered by test companies to measure students' general pronunciation production ability (henceforth, GPPA) (as opposed to general pronunciation listening ability (GPLA)) in the field of English education in Japan. Lado (1961: 96) pointed out that "these partial production techniques are not thought of as full substitutes for direct production techniques, but they can be used effectively in all those cases in which direct production techniques are not possible or impractical."

Buck (1989), however, doubted the reliability (indicating the stability of test scores, including "internal reliability" below) and validity of word-level indirect tests, and denied the effectiveness of using an indirect test as a substitute for a direct test. Several subsequent similar studies have been conducted to investigate on the effects of word-level (Shirahata 1991, Yamauchi 1992, Inoi 1995) and sentence-level (Yamauchi 1992, Komazawa & Ito 1997) indirect tests and concluded that good performance in indirect tests did not guarantee good performance in direct tests.

Although all these previous studies provided insight into the deficiency of indirect tests, their designs were not without limitations. First, very few attempts have been made at indirect tests on the sentence level. In view of the current situation in Japan, namely, a situation in which questions on sentence stresses are frequently given in the Japanese SAT, more attention needs to be given to the efficacy of indirect tests on sentence stresses. Second, most lacked control for

such intervening variables as English proficiency. English proficiency might be an explanatory variable to predict GPA if the direct tests, indirect tests and English proficiency tests showed a high correlation in their scores. In addition, the number of participants in each study was very limited. Furthermore, some sentence-stress level studies, such as Yamauchi (1992) and Komazawa and Ito (1997), did not provide any detailed information as to how the participants' performance was judged phonetically correct or incorrect. To remedy these methodological limitations, the present study was conducted. It focuses on nucleus placement because we believe this may contribute more to the speaker's intelligibility and successful communication than just lexical or sentence stress (Taniguchi 2001).

## 2. PURPOSE AND RESEARCH QUESTIONS

The research questions are as follows: (1) Does an indirect test work as a reliable and valid method of predicting nucleus placement? and (2) Does English proficiency, as an explanatory variable, predict general nucleus placement performance?

## 3. THE STUDY

### 3.1 PARTICIPANTS

A total of 111 Japanese second-year university students (45 male and 66 female) majoring in English language education participated in the present study. All of them were native speakers of Japanese. Their ages ranged from 19 to 23. They had studied English as a foreign language in Japan for an average of 7 years, mainly through highly controlled formal education in Japan.

### 3.2 MATERIALS AND PROCEDURE

*Materials:* The materials consisted of an English proficiency test, an indirect test and a direct test of nucleus placement. All the instructions on the three tests were given in Japanese, but the participants were not told that the tests were to measure their performance on nucleus placement.

The English proficiency test was taken from Shimizu, *et al.* (2004). It is intended to be used in liberal English courses at the university level. The passage selected for the test, "Harry Potter," was a relatively neutral, expository topic, and it contained 300 words. The test consisted of reading comprehension, vocabulary, grammar, and listening.

The direct test was adapted from an example in Hashimoto and Taniguchi (2003: 108): a dialogue (See Appendix).

The same dialogue was used in the indirect test, in which there were 10 underlined sentences. The participants were asked to choose any number

of appropriate words to emphasize in each underlined sentence. They were allowed 15 minutes for completion.

*Procedure:* The ordering of the tests was taken into consideration. The English proficiency test and the direct test were administered in this order on the same day. The participants were given 15 minutes for the completion of the former.

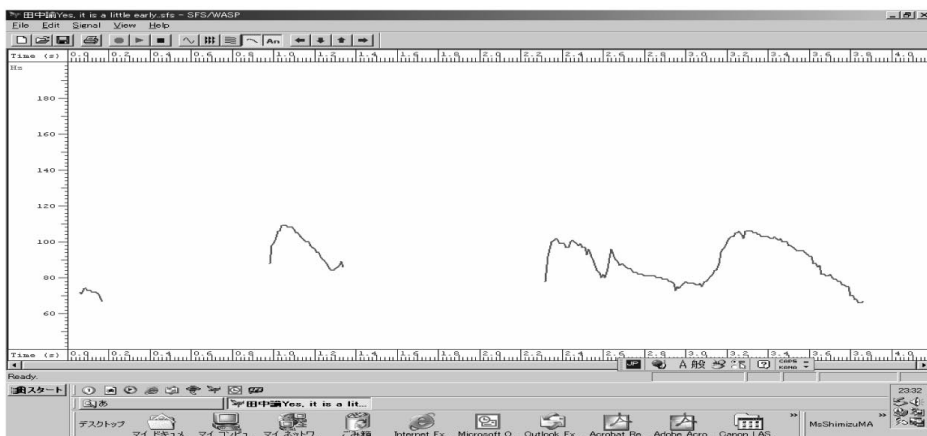
Before starting the direct test, they were given instructions to perform as they would in a real situation and not to write intonation marks or anything else on the test paper. Next, they read the conversation silently for 5 minutes. Then they read it aloud at their own pace. Their reading performance was all recorded on tape.

The indirect test was given to the participants a week later, taking the order effect into consideration, that is, it was intended to avoid the possibility of the participants' remembering the intention of the indirect test to perform better in the direct test.

*Scoring:* The test performance in the direct test was analyzed by one Japanese specialist in phonetics and phonology and one Japanese TEFL specialist. All recorded data were analyzed and judged using *Speech Filing System*, mainly from the perspective of nucleus placement. Below are three typical samples of the participants' nucleus misplacement.

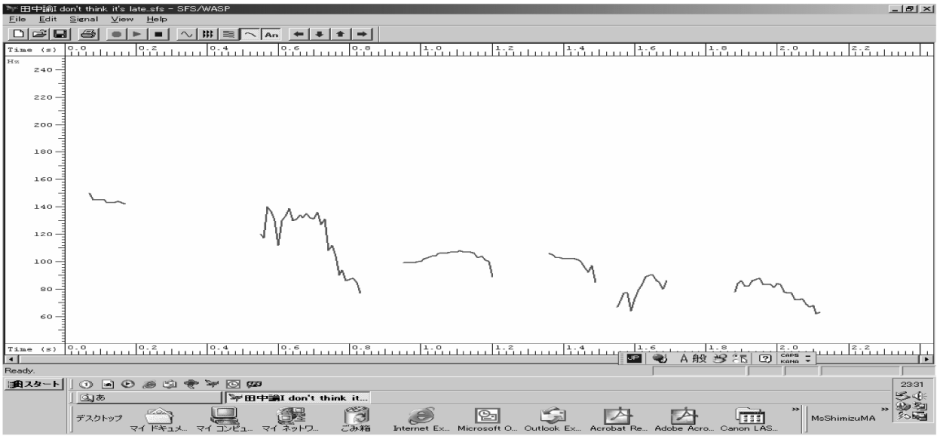
(1) Nucleus placement in this sample was different from what native speakers would normally do. There was a tendency to place a nucleus at the end of each sentence whatever the context, as in Fig. 1, "Yes, it is a little early." as pronounced by one participant in response to "Is 3:30 too early?" in the conversation shown in the Appendix..

Figure 1. "Yes, it is a little early." as pronounced by one subject in response to "Is 3:30 too early?" in the conversation shown in the Appendix.



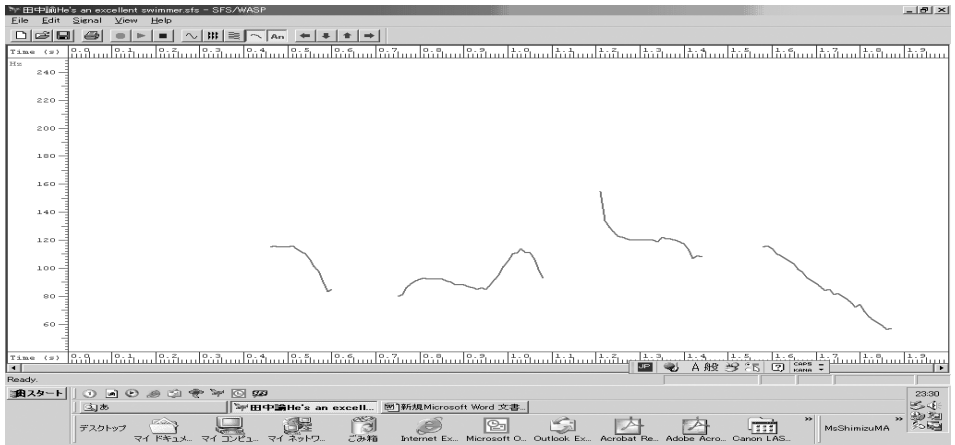
(2) Nucleus placement in the sample in Fig. 2 was **unclear**. Overall, a flat pitch was used; there was no single syllable that stood out clearly from the rest.

Figure 2. "I don't think it's late." as pronounced by one subject in response to "Isn't 4:30 a bit late?" in the conversation shown in the Appendix.



(3) In the sample in Fig. 3, a high flat pitch was used in the syllable where a nuclear tone should be placed, and then a change of frequency occurred in a later syllable where no nuclear tone should be placed.

Figure 3. "He's an excellent swimmer." as pronounced by one subject in response to "He's a very good swimmer!" in the conversation shown in the Appendix.



After a careful examination, the data was analyzed as dichotomous score. All the raw data of the direct, indirect, and English proficiency tests were then entered into SPSS 11.5 for Windows.

## 4. RESULTS

Table 1 shows full scores, mean, standard deviation (s.d.), highest score (max), lowest score (min), and Cronbach's alpha coefficient (which refers to the statistics that indicate the internal reliability).

Table 1. Descriptive statistics of the tests ( $n=111$ )

	English proficiency test	indirect test	direct test
full score	18	10	10
mean	11.69	4.28	0.80
s.d.	2.20	1.34	1.37
max	17	8	9
min	6	1	0
Cronbach's $\alpha$	0.43	0.23	0.78

Beginning with the descriptive statistics of the test results in Table 1, we can see that the means of the English proficiency test and the indirect test were 11.69 (full score =18) and 4.28 (full score=10) respectively. The mean of the direct test, however, was 0.80 (full score=10). Table 2 describes the score frequency of the direct test, which shows the fact that 106 (95.5%) out of the 111 participants fell into the group scoring from 0 to 2 (full score=10).

Table 2. Score frequency of the direct test

score	N	percentile
0	61	55.0
1	30	27.0
2	15	13.5
3	1	0.9
5	2	1.8
7	1	0.9
9	1	0.9
sum	111	100

Most of the participants in this study did not perform as well as they did in the indirect test. This suggests that they were not well trained on performing nucleus placement.

The reliability of the direct test was high ( $\alpha=0.78$ ) and the English proficiency test proved relatively moderately reliable ( $\alpha=0.43$ ). However, the indirect test showed low reliability ( $\alpha=0.23$ ).

Table 3 displays the correlation coefficients among the three tests.

Table 3. Correlations among the tests ( $n=111$ )

	English proficiency test	Indirect test	direct test
English proficiency test	1.00		
indirect test	0.18	1.00	
direct test	0.09	0.15	1.00

There were no significant correlations between the English proficiency test and the indirect test ( $r=0.18$ , n.s.), between the English proficiency test and the direct test ( $r=0.09$ , n.s.), or between the indirect test and the direct test ( $r=0.15$ , n.s.).

In Tables 4 and 5, the correlation coefficients among the test results of advanced and less advanced learners are shown respectively.

Table 4. Correlations among the tests (less advanced learners,  $n=8$ )

	English proficiency test	indirect test	direct test
English proficiency test	1.00		
indirect test	-0.54	1.00	
direct test	-0.05	0.15	1.00

Table 5. Correlations among the tests (advanced learners,  $n=13$ )

	English proficiency test	indirect test	direct test
English proficiency test	1.00		
indirect test	-0.49	1.00	
direct test	0.32	0.19	1.00

In the lower group, there were no significant correlations between the English proficiency test and the indirect test ( $r=-0.54$ , n.s.), between the English proficiency test and the direct test ( $r=-0.05$ , n.s.), or between the indirect test and the direct test ( $r=0.15$ , n.s.).

Even in the higher group, there were no significant correlations between the English proficiency test and the indirect test ( $r=-0.49$ , n.s.), between the English proficiency test and the direct test ( $r=0.32$ , n.s.), or between the indirect test and the direct test ( $r=0.19$ , n.s.).

## 5. DISCUSSION

We now need to examine these findings further in the context of our two research questions; we will discuss these in turn.

(1) *Does an indirect test work as a reliable and valid method of predicting nucleus placement?*

*Reliability:* The reliability coefficient is sometimes referred from the internal reliability perspective. Cronbach's alpha in Table 1 was very low ( $\alpha = 0.23$ ). The low internal reliability suggests that less advanced learners responded correctly to questions that more advanced learners responded incorrectly to, and vice versa. This pattern of response occurred frequently in the indirect test.

*Validity:* The validity of indirect testing can be calculated in terms of concurrent validity, namely, correlation coefficient with direct testing, which showed the highest internal reliability in this study. The coefficient was very low ( $r = 0.15$ , n.s.), which means the indirect test has no concurrent validity.

We conclude that indirect testing cannot in any way be a substitute for direct testing, namely, it does not work as a reliable and valid measure of predicting nucleus placement performance. By way of parenthesis, 125 questions would be needed to obtain 0.8 of Cronbach's alpha coefficient (Spearman-Brown prophecy formula). However, it would be unrealistic for teachers to administer a test with such an enormous number of questions within a limited time.

(2) *Does English proficiency, as an explanatory variable, predict general nucleus placement performance?*

To examine this research question, the correlations among the tests should be considered. English proficiency will explain nucleus placement performance if the correlation coefficient between the English proficiency and the direct test proves high. As is shown in Table 3, however, there was no significant correlation between them ( $r = 0.09$ , n.s.). To examine the details further, the advanced learners (more than 14.99 (mean+ 1.5 s.d.)) and the less advanced learners (less than 8.39 (mean-1.5 s.d.)) were extracted according to their scores of the English proficiency test. It was expected that the higher scores of English proficiency would predict higher nucleus placement performance. Again there were no correlations between the English proficiency test and the direct test in each group ( $r = -0.05$ , n.s. for lower group,  $r = 0.32$ , n.s. for higher group). No preferences could be found on the relationship between English proficiency and production ability.

## 6. SUMMARY

Indirect tests are economical to make and easy to mark. They are frequently used in the Japanese SAT, mid-term and final tests at school, and in other kinds of tests administered by test companies. Attempts to prove the reliability of indirect tests on phoneme discrimination and stress (both word and sentence levels) have been made, but none on nucleus placement.

The present study investigated issues of the reliability and validity of an indirect test and the possibility of an English proficiency test as a substitute for a direct test from the nucleus placement perspective. The results obtained



from this study were: (1) Indirect tests cannot measure nucleus placement performance and (2) English proficiency is not an explanatory variable of predicting nucleus placement performance.

It should be acknowledged that the alpha coefficient in the English proficiency test was estimated low. A high correlation coefficient would be expected between the English proficiency test and the direct test if the participants are experts in English, such as interpreters, translators, and phoneticians. However, it was not true with the participants in the present study, who were students studying English as a foreign language at schools in Japan. It is quite understandable, if we consider their proficiency levels. Our conclusion is that only direct tests are applicable to measure students' nucleus placement performance.

This research also made us realize the responsibility of teachers to receive adequate phonetic training to learn to auditorily discriminate between their students' correct and incorrect performances.

### ACKNOWLEDGEMENTS

Sincerest gratitude is due to Professor John Wells, professor of phonetics at UCL, for his helpful comments on this study. Any shortcomings are the responsibility of the authors.

### REFERENCES

- Brown, J. D. 1996. *Testing in Language Program*. Upper Saddle River, NJ: Prentice Hall Regents.
- Buck, G. 1989. Written Tests of Pronunciation: Do They Work? *ELT Journal* 43, 50-56.
- Komazawa, S. and A. Ito. 1997. A Written Test of Stress in Connected Speech in the NCUEE-test: its Reliability & Validity. *Chuubu Chiku Eigo Kyoiku Gakkai Kiyō* 27, 285-292.
- Hashimoto, M. and M. Taniguchi. 2003. *English Sounds and Oral Communication: Approach to Living English*. Tokyo: English Phonetic Society of Japan.
- Inoi, S. 1995. *The Validity of Written Pronunciation Questions: Focus on Phoneme Discrimination. Language Testing in Japan*. Tokyo: JALT.
- Lado, R. 1961. *Language Testing*. London: Longman.
- Shimizu, Y. et al., 2004. *Exploring Cultural Issues: Practice in TOEIC Test Format*. Tokyo: Seibido.
- Shirahata, T. 1991. Validity of Paper Test Problems on Stress: Taking Examples from Mombusho's Daigaku Nyushi Senta Shiken. *Shizuoka Daigaku Kyoiku Gakubu Kiyō*, 23, 161-172.
- Taniguchi, M. 2001. Japanese EFL Learners' Weak Points in English Intonation. *English Phonetics* 4, 45-54.
- Yamauchi, S. 1992. Validity of Measuring Oral Production Ability on Pencil-and-paper Tests. *Kyushu Eigo Kyoikugaku Kenkyū Kiyō* 20, 72-82.



## APPENDIX

Materials used in the direct and indirect tests (Below is the sample of the indirect test)

- A What time shall we meet tomorrow?  
 B Shall we meet at 3:30?  
 A Q1 How about 4:30?  
 B Is 3:30 too early?  
 A Yes, Q2 it is a little early.  
 B Isn't 4:30 a bit late?  
 A Q3 I don't think it's late.  
 B Well, the party begins at five o'clock.  
 A All right, John. How about four o'clock instead of 4:30?  
 B Q4 I guess that will do.  
 A Have you invited any Olympic athletes to the party?  
 B Yes, Q5 three medallists are coming.  
 A Q6 Who are they?  
 B One of them is Donald Dolphin!  
 A Great! He got six Gold Medals and one Silver Medal.  
 B He's a very good swimmer!  
 A Q7 He's an excellent swimmer!  
 B I think he came to Japan in nineteen eighty-nine.  
 A Yes, and Q8 he came again in nineteen ninety-nine.  
 B Will he come to Japan again next year?  
 A Yes.  
 B Q9 I thought he would.  
 A If I remember correctly, he'll be coming yet again Q10 the year after next.  
 B Will he?

Q1	1 How	2 about	3 four	4 thirty	
Q2	1 it	2 is	3 a	4 little	5 early
Q3	1 I	2 don't	3 think	4 it's	5 late
Q4	1 I	2 guess	3 that	4 will	5 do
Q5	1 three	2 medallists	3 are	4 coming	
Q6	1 Who	2 are	3 they		
Q7	1 He	2 is	3 an	4 excellent	5 swimmer
Q8	1 came	2 again	3 nineteen	4 ninety	5 nine
Q9	1 I	2 thought	3 he	4 would	
Q10	1 the	2 year	3 after	4 next	

## SUMMARY

### DO INDIRECT TESTS OF ENGLISH INTONATION MEASURE STUDENTS' ABILITY IN PERFORMING ENGLISH NUCLEUS PLACEMENT?

This study focuses on nucleus placement in English intonation and attempts to investigate (1) the reliability and validity of indirect tests on nucleus placement, and (2) the credibility of English proficiency tests as a predictor of students' ability in nucleus placement. Three tests were administered to 111 Japanese EFL university students: (a) an English proficiency test, (b) an indirect test on nucleus placement, and (c) a direct test (oral performance test) on nucleus placement.

The results obtained are: [1] the internal reliability of the direct test was very high, but the internal reliability of the indirect test was very low, and [2] there were no significant correlations among the three tests. These results indicate that indirect tests and English proficiency tests cannot predict the students' performance of nucleus placement. This research also made us realize the responsibility of teachers to receive adequate phonetic training to learn to auditorily discriminate between their students' correct and incorrect performances.

**KEYWORDS:** direct test, indirect test, intonation, nucleus placement, reliability, validity, National Centre for University Entrance Examinations.