

Is the Reading Comprehension Performance of Learners of Japanese as a Second Language the Same as That of Japanese Children? An Analysis Using a Cloze Test

Sayoko Okada Yamashita*

Key words: **reading comprehension, native first graders, Japanese as a second language (JSL) learners, cloze test, item analysis**

This study investigates whether there are differences in reading comprehension performance between native Japanese children and learners of Japanese as a second language (JSL). The subjects were public elementary-school first graders ($n=43$), as well as advanced ($n=29$) and intermediate ($n=31$) university students who were studying JSL. A fixed-ratio cloze test with 72 blanks (adapted from a folk tale, entitled *Momotaro*) was used to measure comprehension performance. The statistical analyses included descriptive statistics, reliability estimates, and item analysis. One-way analysis of variance (ANOVA) indicated that there was a significant difference between native children and lower proficiency JSL learners, although, there was no significant difference between native children and higher proficiency JSL learners. However, closer analysis indicated several salient differences between native children and advanced JSL learners. This study demonstrates that the more proficient JSL learners become, the closer their proficiency becomes to having native proficiency. In addition, the cloze procedure appears to have been a good measure of the language knowledge of both native children and JSL learners and also helped to discover individual items that are difficult for JSL learners to master.

INTRODUCTION

The final goal for second/foreign language learners in learning a target language is to master it with native-like proficiency (ACTFL, 1986). If the above assumption is true, advanced second/foreign language learners should perform as well as native speakers on a test that measures their language ability.

Cloze tests, in which subjects fill in words that have been deleted from a reading selection, were first used by Taylor (1953) to assess the readability of texts. A num-

* 山下早代子: Lecturer, International Christian University, Japan.

ber of studies have been done on the effectiveness of cloze as a measure of reading proficiency for native speakers of English (Bormuth, 1965; Miller and Coleman, 1967; Shiba, 1957; Taylor, 1957). Later, extensive studies utilizing cloze in testing ESL students were reported (for overviews of this research, see Oller, 1979; Hinofotis, 1987). Chihara et al. (1977) studied native speakers and ESL students, using a cloze procedure (fixed-ratio deletion), and found that the percent of correct closures (i.e., correct answers written in the blanks) was higher for groups with higher language proficiency than for groups with lower proficiency. In other words, the Chihara et al. study demonstrated that the native speakers and the advanced learners did equally well on the cloze test compared with lower proficiency learners.

Alderson (1980) also reported on a study comparing the performances of native speakers of English with those of non-native speakers. Native and non-native speakers performed similarly on a number of cloze tests that the researcher constructed by manipulating variables in the cloze test design.

Among the research issues raised about cloze tests in the literature, two stand out as particularly relevant to this study. One is the argument about what cloze really assesses. One group of researchers argues that cloze test items are primarily assessing sentence-level linguistic elements (Alderson, 1979; Klein-Braley, 1983; Markham, 1985; Porter, 1983). Another group finds that cloze is a stable, reliable, and sensitive measure of the inter-sentential components of the language (Bachman, 1982; Brown, 1983; Chavez-Oller, Chihara, Weaver, and Oller, 1985; Jonz, 1987, 1990). This researcher was concerned about which position to take. For instance, Jonz's (1990) claim that his analysis of the passage from Bachman (1985) in terms of categorization (inter-sentential analysis) did not match Bachman's analysis for 14 cloze deletions out of 30. Even if the claim of the latter group is true (i.e., that cloze is sensitive to inter-sentential components of the language), the issues seem to be very complicated. Probably, as Brown (1991) stated, cloze items assess a wide range of language points from morphemic and clausal level grammar rules to discourse and pragmatic level rules of cohesion and coherence (p. 2). This researcher shall mainly deal with morphemic and clausal level grammar rules in order to make the analysis simple and clear.

Another cloze research issue deals with types of cloze procedure. Researchers have studied the reliability, validity, mean item facility and discrimination, and usability of various types of scoring methods (Bachman, 1985; Brown, 1980, 1984, 1988, 1992; Chapelle and Abraham, 1990; Darnell, 1970; Feldmann and Stemmer, 1987; Jonz, 1976; Klein-Braley, 1985; Markham, 1985; Shin, 1990). In addition, there are many types of cloze procedures, such as fixed-ratio, rational deletions, multiple-choice, clozen-trophy, and C-test. Among scoring methods, even the fixed-ratio and rational deletion methods have two scoring choices, namely, exact-word scoring and acceptable-word scoring methods.

In the present study, fixed-ratio deletions were applied to measure language traits, more specifically, written grammatical and textual reading competence (after Chapelle and Abraham, 1990). The acceptable-word scoring method was used after Brown (1980) with pilot-test results which were obtained from native university students in

order to compile a glossary of acceptable answers for each blank. The scoring method will be described in detail in the Procedures section of this paper.

The purpose of this study, then, was to compare native Japanese childrens' reading comprehension performance to adult learners of Japanese as a second language (JSL) using a cloze test that was specially developed for this study. In other words, this study investigated whether there is a difference in reading comprehension performance between native Japanese first graders and learners of Japanese as a second language (advanced and intermediate levels) by using a cloze procedure as the measurement instrument. To that end, more specific research questions were formulated as follows:

- 1) Is the cloze test reliable for each group (the first graders, advanced JSL learners, and intermediate JSL learners)?
- 2) What are the item facilities and discrimination indices for the first graders, advanced JSL learners, and intermediate JSL learners' performance?
- 3) Which items discriminate between JSL learners and native Japanese children?
- 4) Is there a significant difference in performance between JSL learners and native children?

Method

Subjects

The subjects in this study were: 1) a group of native first graders (aged 6 to 7) in a public elementary school in Musashino City, Tokyo (n=43); 2) a group of university students in the advanced JSL level at International Christian University (ICU) (n=29); and 3) a group of university students in the intermediate JSL level (n=31) at ICU. They were all sampled as intact classes. The test for the first graders was administered at the end of the school year, which was in March in the Japanese school calendar. It was the time when the children had completed all of their work for the first grade. The JSL subjects were predominantly American. The number of years of Japanese study for the advanced students varied from 2 to 5 years, and they had studied it with varying degrees of intensity. The intermediate level in this study was the level in which the students had just finished studying the grammar and structure of beginning level Japanese. The levels were determined by a placement test administered at the beginning of the school year, which was September for the JSL groups.

The same cloze test was also administered to 33 native Japanese university students as a pilot test to compile a glossary of acceptable answers.

Materials

A cloze test with every ninth character deleted was designed from an old folk tale for children, entitled "*Momotaro*" ("Peach Boy"), taken from Yamashita and Ogawa (1994). The story is well known to native Japanese. All 33 Japanese university students indicated (when asked) that they knew the story after completing the pilot test, and 40 out of 43 first graders indicated that the cloze test was about "*Momotaro*." This particular story was selected because it is familiar to children, and also because

it deals with cultural knowledge, which may play an important role in measuring advanced JSL learners' proficiency.

The method of deleting every ninth character (including kanji) was applied to this test. Ogawa (1992) deleted every seventh character for two cloze tests for a JSL placement test (based on the literature). In that study, one cloze passage was taken from an intermediate JSL text and another was taken from an authentic newspaper article. Both passages made extensive use of kanji, which carries meaning in most nouns and stems of verbs. Since the text in this study was written mostly in hiragana, syllabic characters, it was necessary to have more characters between deletions in order to carry morphological meanings. Thus, the system of deleting every ninth character was used.

A typical cloze test has several sentences intact at the beginning of the passage to provide context (Hinofotis, 1987; 413). In this study, however, the first blank appeared at the ninth character. Since in any passage starting with “*Mukashi mukashi*,” it is obvious to native speakers and advanced learners (if they have native-speaker proficiency) that “[*a*]rutokoroni” follows. In order to keep the blanks uniform throughout, commas, periods, and parentheses were treated as characters. The test was written in hiragana characters and kanji were limited to those taught to first graders according to the *Monbushō* (Ministry of Education) kanji list for elementary students. The first paragraph was as follows (see Appendix A for the complete test and acceptable answers. The original text was written in hiragana and kanji; the test was not given in Roman characters as shown in the example below):

Mukashimukashi, [A]rutokoroni, oji[I] santoobaasan[GA] sundeimashita(.)
Ojisanwamai[NICHI]yamaeshibakarini, [O]baasanwakawaese[N]takumiiki-*
dashita.

**NICHI* is a kanji.

The total number of blanks was 72, including recursive words.¹ The grammatical categories which appeared in the passage are listed as follows (numbers in the parentheses indicate frequencies; categories were established after Makino and Tsutsui (1986)).

- 1) Adjectives (4)
- 2) Adverbs (2)
- 3) Conjunctions (1)
- 4) *ko-so-a-do* (4) (special functions-demonstrative, etc.)
- 5) Nouns—(13)/proper nouns (3×5 each²)
- 6) Onomatopoeiac words (2)
- 7) Particles (8)
- 8) Prefixes (1)
- 9) Special functions (4) (causative 1; passive 1, others 2)
- 10) Verbs—Compound verbs (5)/inflection (7)/stem (6)

¹ “*Obaasan*” appeared in line one. Then it appeared recursively in item 5 as “[*O*]baasan.” This pattern makes an easy guess for JSL learners.

² “*Obaasan*,” “*Ojisan*,” and “*Momotaro*” fit the blanks three times each.

The acceptable-answer scoring method was used based on the native university student (NUS) data. In other words, if a word was replaced by another word which was found equally frequently (more than one-third of the time) in the NUS data, both words were accepted as correct answers. One exception was that the word *soko* [*NI*] (*soko* [*E*] was the exact word) was not accepted even though the native data showed almost equal frequency. This was decided rationally by grammatical analysis.³ Table 1 shows the decisions for acceptable answers.

Table 1 Native Speaker's Usage of Words in Pilot Study

Exact word	%	words appeared in NUS data	n=33
# 9 <i>soko</i> [<i>E</i>]	58%	<i>soko</i> [<i>NI</i>] 42%	(not accepted ³)
#24 [<i>WA</i>] <i>rou</i>	52%	[<i>KI</i>] <i>rou</i> 48%	(accepted)
#27 <i>paku</i> [<i>T</i>] <i>to</i>	52%	<i>paku</i> [<i>RI</i>] <i>to</i> 42%	(accepted)
		<i>paku</i> [<i>N</i>] <i>to</i> 6%	(not accepted)
#35 [<i>SO</i>] <i>no ko</i>	79%	[<i>KO</i>] <i>no ko</i> 21%	(not accepted)
#45 [<i>YA</i>]	93%	[<i>TO</i>] 15%	(not accepted)

A possible word is shown in [].

Procedures

The native first-graders' data were collected from two classes (n=49) of the same public school in Tokyo. The test was administered in classrooms under the homeroom teachers' supervision in March toward the end of the school year. The children should have mastered all hiragana and 76 kanji by this time. As shown in Table 2, the time limit was 30 minutes. Six out of the 49 answer sheets were incomplete. Only complete answer sheets (i.e., n=43) were used in the analysis. Most children spent 20 to 30 minutes according to the teachers. (Note that the native university students in the pilot group spent only 2.5 to 3 minutes to complete this test). The JSL learners

Table 2 Subjects

	Native children (G1)	Adv. JSL	Int. JSL	NSU-Pilot
Valid	43 (88%)	29 (100%)	31 (79%)	33 (100%)
Invalid**1	6 (12%)	0 (0%)	8 (21%)	0 (0%)
know <i>Momotaro</i> **2	40 (93%)	23 (79%)	5 (16%)	33 (100%)
Time spent	30 min.	10-20 min.	30 min.	2-3 min.

**1 Test was incomplete (probably due to time limit).

**2 Knowing the story or not may affect the ability to read and fill in the blanks.

³ The particle, "*e* (*he*)" is described as "a particle that indicates the direction toward which some directional movement or action proceeds" (Makino and Tsutsui, 1986: 116), whereas "*ni*" is described as "a particle which indicates a place toward which someone or something moves." (Ibid., 302). The particle "*e*" seems to carry a more exact meaning (i.e., "directional movement") in this context.

also took the test in a classroom. The advanced students spent about 10 to 20 minutes. Most of the intermediate JSL learners spent 30 minutes. Eight out of 39 intermediate JSL learners did not complete the test, whereas all the advanced JSL learners completed the test. (See Table 2.)

Analysis

The descriptive test statistics in this study included the mean, standard deviation (SD), minimum and maximum scores, range, variance, number of subjects (n), and number of items (k). The K-R20 reliability coefficient was also calculated.

Item facility and item discrimination for each of the 72 items were calculated separately for each group (native first graders, advanced JSL, and intermediate JSL learners). Each group was further divided into three levels (approximately one-third each) for upper, middle, and lower levels for estimating item discrimination.

To determine which specific differences among the group means were significant, a one-way ANOVA was performed, and post-hoc comparisons were made with the Scheffé test. Then, one-way analysis of variance procedures were applied for each of the 72 items separately to see if there were any significant mean differences on specific items. The alpha for all statistical tests was set at .05. Hence, with this many statistical comparisons, the results must be interpreted very cautiously.

Results

The descriptive statistics for the test results of each group are reported in Table 3. The difference between the means of the native first graders and the advanced JSL learners was only 3.46. However, the first-graders' minimum score was 18, which made the range much larger than those of the two groups of JSL learners. The K-R20 statistic indicated that the reliability of the test was higher in the more proficient levels (i.e., the native first graders and advanced JSL learners) than for the lower level. But even in the lower level, a K-R20 of .83 indicates that the test was remarkably reliable.

Table 3 Descriptive Statistics

	1st Grade	Adv. JSL	Int. JSL
k	72	72	72
n	43	29	31
MEAN	62.77	59.31	43.65
SD	9.82	8.43	6.55
MIN	18.00	32.00	28.00
MAX	72.00	71.00	64.00
RANGE	54.00	39.00	36.00
K-R20	0.94	0.91	0.83

The item facility (IF) and item discrimination (the percentage answering correctly in the top 33% minus the percentage for lower 33%) was also calculated for each group. These item statistics are listed in Appendix B.

The overall ANOVA procedure indicated that there was a significant difference somewhere among the individual item comparisons ($F=46.56$; $df=2, 100$; $p<.01$) between the group means. The Scheffé procedure further indicated that, overall, Group 1 (native 1st graders) and Group 3 (intermediate JSL), and Group 2 (advanced JSL) and Group 3 were significantly different. However, Group 1 and Group 2 were not significantly different (i.e., native children and advanced JSL learners were not significantly different). F-ratios were then calculated for each item. Thirty-three items were found to be significant (i.e., items 1, 3, 9, 13, 20, 24, 25, 27, 28, 30, 35, 36, 38, 42, 43, 44, 45, 46, 47, 48, 49, 50, 53, 54, 56, 60, 61, 66, 67, 68, 69, 70, and 72).

Discussion

This study addressed four research questions regarding comparisons between native Japanese children and two groups of JSL university students. The first question concerned the reliability of the test for each group. The K-R20 reliability coefficients indicated that the test had moderate to very high reliability for these three groups. This means that the results of this cloze procedure were consistent regardless of the background of the subjects (i.e., native speakers and JSL learners). In other words, the test was equally consistent for native children and JSL university students, regardless of level. The above findings also suggest that the cloze procedure may be used not only for measuring reading comprehension but also for measuring second language acquisition processes for comparison with native children (although this paper is not dealing with that topic).

The second research question was aimed at determining whether or not there would be differences in item facility and item discrimination for certain groups (i.e., native children and JSL learners). The results indicated that the performances of the native children and the advanced JSL learners were very similar. Interestingly, the advanced JSL learners answered more items correctly than the native children (31 items out of 72, or 43%). Close analysis revealed that: 1) advanced JSL learners were more correct in writing rules than native children on some items (e.g., geminative consonant, or small *tsu*, particle usage of *wa* and *o*, or the distinction between *ra* and *da* in *tabe-[ra]rete*); 2) JSL learners used reasonable guesses or higher-order cognitive strategies for solving problems (e.g., correctly filling in recursive words such as *[O]baasan* after *obaasan* had appeared in a sentence before); 3) classroom learning for JSL may affect performance (e.g., both the intermediate and advanced JSL learners performed better than the first graders in *sentaku [O] shiteiru*, and *tabemasho[U]*, both of which the JSL learners learn systematically according to the grammatical syllabus in the classroom); on the other hand, 4) native children performed better on idiomatic or cultural expressions or words which they often hear and/or read naturally (e.g., *arutokoroni*,

ojiisanto obaasan[GA] *sundeimashita* where many JSL learners put [WA] instead. The conjunction usage, [SU]*ruto sonotoki*, *yamete*[O]*kure*, and *kibi*[DA]*ngo* as a special vocabulary appeared in this particular folk tale were such examples); 5) some expressions are hard to acquire for JSL learners (e.g., usage of the particle for 'purpose' in *oniotaijishi*[NI], and the adverbial usage of adjectives as in *genkiyo*[KU]*aruitekima-shita*); and 6) both native children and JSL learners had difficulty with words that carried delicate shades of meaning compared to the native university students (e.g., usage of [YA] which indicates 'two or more items' being distinctively different from [TO] (two items)). Hence, detailed analysis has revealed clear-cut differences in the proficiency of native children and JSL learners.

The third research question asked if there was any difference in performance on each item according to the group membership (native-children, advanced JSL and intermediate JSL). Seventy-two items were analyzed using one-way ANOVA and 46 percent of the items (33 out of 72) turned out to be significant. Hence, this test can be used to discriminate between native children and JSL learners as well as used as a comprehension test for each group (as stated in question 1).

The last research question was concerned with whether or not there was a significant difference in overall performance between native children and JSL learners. According to the statistical analysis, the performance of native children and advanced JSL learners was not significantly different, but that of intermediate JSL learners was significantly lower than both. This indicates that at the lower level, JSL learners are different from natives and as they get closer to an advanced level, their performance becomes closer to that of native children. However, as indicated above in answering research question 2, a close grammatical and semantic analysis revealed different characteristics between the native children and the advanced second language learners. These results suggest that the acquisition processes of first and second language learners may conceivably be investigated using data obtained on a cloze procedure, which lies beyond the scope of this paper.

CONCLUSION

This study has described characteristics of the reading comprehension performance of native children, advanced JSL learners, and intermediate JSL learners using a cloze test based on a folk tale, *Momotaro*. The findings indicated that the native children and the advanced JSL learners had similar overall linguistic performance, while the performance of the native children and the intermediate JSL learners were significantly different. The close analyses suggested that even between the native children and the advanced learners, there were some salient differences. Moreover, this study found that the cloze test was an effective measure for analyzing reading comprehension performance for both native children and JSL learners. Although this study only dealt with linguistic elements, it appears that cloze tests provide fairly representative samples of the written language, including rule systems at the word, clause, sentence, discourse, and pragmatic levels (Brown, 1991). Hopefully, analyses of the higher level inter-

sentential components of the language can also be conducted on the data from this study in the future.

ACKNOWLEDGMENTS

I am grateful to Dr. J. D. Brown of the University of Hawaii at Manoa for his insightful comments and suggestions. I would like to thank Ms. Fumiko Kawakami of Sakurazutsumi Elementary School in Tokyo for her help in collecting the test samples from the first graders. I also wish to thank my colleagues at International Christian University, Machiko Netsu, Takashi Ogawa, and Yōko Suzuki, for their help in collecting data from their students for this study.

Appendix A

Cloze test "MOMOTARO"

あなたのなまえ()

年 月 日

() にひらがなかんじを一ついれてください。

むかしむかし、(あ)るところに、おじ(い)さんとおばあさん(が)すんでいました。

おじいさんはまい(日)山へしばかりに、(お)ばあさんは川へせ(ん)たくにいきました。

ある日、おばあ(さ)んが川でせんたく(を)していると、そこ(へ)大きなももがどん(ぶ)らこ、どんぶらこ、とながれてきまし(た)。おばあさんは、「おや、まあ、これは(は)大きなももだこと。いえへもってかえ(つ)ておじいさんとい(っ)しよにたべましょ(う)」と(い)って、いえへ(も)ってかえりました。

夕(た)がた、おじい(さ)んは山からかえ(つ)てきました。おじい(さ)んはももを見ると、大(そ)うび(っ)くりし(て)、「おやおや、こ(れ)は大きなももだ。(お)ばあさん、いっし(よ)にたべるとしよう」と(い)って、も(も)を(わ)ろう(と)しまし(た)。(す)るとそのとき、も(も)がまん中(か)らば(っ)と(わ)れて、中(か)ら(か)わい(い)男(の)赤(ち)ゃ(ん)が「おぎ(ゃー)と(う)ま(れ)まし(た)。おじ(い)さんとおばあ(さ)ん(は)と(も)おど(ろ)き(ま)し(た)。け(れ)ど(も)、子(ど)も(が)い(な)い(の)で、(こ)の(こ)を(大)せ(つ)に(そ)だ(て)る(こ)とに(し)ま(し)た。名(な)ま(え)は、も(も)か(ら)ま(れ)た(の)で「(も)も(た)ろう」に(し)ま(し)た。

も(も)も(た)ろう(は)ま(い)日(日)、ま(い)日(日)大(き)く(な)っ(て)い(き)ま(す)。そ(の)こ(ろ)、村(むら)に(お)る(い)お(お)に(が)や(っ)て(き)て(は)、村(むら)人(にん)か(ら)た(べ)も(の)や(き)る(も)の(を)と(っ)て(い)き、村(むら)人(にん)を(大)へ(ん)(こ)ま(ら)せ(て)い(ま)し(た)。そ(こ)で、も(も)も(た)ろ(う)は(お)に(を)たい(じ)し(に)、お(お)に(が)し(ま)へ(い)(く)こ(と)に(し)ま(し)た。(お)じ(い)さん(と)、お(お)ば(あ)さん(は)「そ(ん)な(あ)ぶ(な)い(こ)と(は)や(め)て(お)く(れ)、お(お)に(に)た(べ)ら(れ)て(し)ま(う)よ」と(い)っ(て)し(ん)ば(い)し(ま)し(た)。け(れ)ど(も)、も(も)も(た)ろ(う)は、「だ(い)じ(よ)う(ぶ)、き(っ)と(ぼ)く(が)お(お)に(を)や(っ)つ(け)て(き)ま(す)」と(い)い(ま)し(た)。お(お)じ(い)さん(と)お(お)ば(あ)さん(は)、も(も)も(た)ろ(う)の(た)め(に)、「(日)本(に)一(に)」と(か)い(た)は(た)と、日(に)本(に)一(に)の(き)び(だ)ん(ご)を(つ)く(っ)て、も(も)も(た)せ(て)や(り)ま(し)た。

「日本一」のは(た)と、日本一のきび(だ)んごをもったもも(た)ろうが、げん気よ(く)あるいていきました。

[しつもん]

あなたはこの話をしていますか。 はい→なんという話ですか。

いいえ、しりません。

Appendix B
Item Analysis

GRADE 1 (1)=the native Japanese children (first graders)

ADV. JSL (2)=the advanced Japanese learners

INT. JSL (3)=the intermediate Japanese learners

I= item

IF= item facility

IFT= item facility for the upper group (33%) on the whole test

IFB= item facility for the lower group (33%) on the whole test

ID= item discrimination for an individual item (=IFT - IFB)

IV= item variance

DI= difference index

	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116
GRADE1(1)	1.00	1.00	0.98	0.91	0.91	0.98	1.00	0.86	0.53	0.81	0.98	0.74	0.95	0.91	0.86	0.81
IF	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.79	1.00	1.00	0.93	1.00	1.00	1.00	0.93
IFT	1.00	1.00	0.93	0.87	0.73	1.00	1.00	0.60	0.20	0.53	0.93	0.60	0.87	0.73	0.73	0.67
IFB	0.00	0.00	0.07	0.13	0.27	0.00	0.00	0.40	0.59	0.47	0.07	0.33	0.13	0.27	0.27	0.26
ID	1.00	1.00	0.93	0.74	0.46	1.00	1.00	0.60	0.21	0.06	0.86	0.27	0.87	0.46	0.46	0.11
IV	0.00	0.00	0.02	0.08	0.08	0.02	0.00	0.12	0.25	0.15	0.02	0.19	0.04	0.08	0.12	0.26
ADV. JSL(2)	0.93	0.97	0.83	0.63	1.00	0.97	0.97	0.97	0.38	0.93	1.00	0.93	0.93	1.00	1.00	0.72
IF	1.00	1.00	1.00	0.90	1.00	1.00	1.00	1.00	0.30	0.90	1.00	0.80	1.00	1.00	1.00	0.90
IFT	1.00	1.00	0.50	0.60	1.00	0.90	1.00	0.90	0.30	0.90	1.00	0.60	0.90	1.00	1.00	0.50
IFB	0.00	0.00	0.50	0.10	0.00	0.10	0.00	0.10	0.30	0.10	0.00	0.20	0.10	0.00	0.00	0.40
ID	1.00	1.00	0.50	0.10	1.00	0.80	1.00	0.80	0.00	0.00	1.00	0.40	0.80	1.00	1.00	0.10
IV	0.20	0.03	0.14	0.14	0.00	0.03	0.03	0.03	0.24	0.06	0.00	0.06	0.06	0.00	0.00	0.20
INT. JSL(3)	0.32	1.00	0.26	0.67	1.00	0.94	1.00	0.90	0.10	0.71	0.97	0.74	0.77	0.97	0.94	0.56
IF	1.00	1.00	0.90	0.60	1.00	1.00	1.00	1.00	0.10	0.90	1.00	0.60	0.70	1.00	1.00	0.30
IFT	1.00	1.00	0.20	0.60	1.00	0.80	1.00	0.80	0.10	0.40	0.90	0.60	0.70	0.90	0.90	0.30
IFB	0.00	0.00	0.30	0.10	0.00	0.20	0.00	0.10	0.00	0.10	0.10	0.20	0.00	0.10	0.10	0.00
ID	1.00	1.00	0.30	0.10	1.00	0.60	1.00	0.70	0.10	0.30	0.80	0.40	0.70	0.80	0.80	0.30
IV	0.40	0.00	0.19	0.11	0.00	0.05	0.00	0.09	0.09	0.21	0.03	0.19	0.17	0.11	0.06	0.24
DI1-3	0.07	0.03	0.15	0.08	-0.09	0.01	0.03	-0.11	0.16	-0.12	-0.02	-0.19	0.02	-0.09	-0.14	0.09
DI2-3	0.88	0.00	0.57	-0.04	-0.05	0.04	-0.01	-0.04	0.74	0.22	0.01	0.00	0.16	0.13	-0.08	0.23
DI3-3	0.81	-0.03	0.57	-0.04	0.00	0.03	-0.03	0.06	0.73	0.22	0.03	0.19	0.16	0.13	-0.08	0.14
GRADE1(1)	0.98	0.98	1.00	0.88	0.88	0.95	0.88	0.95	0.84	0.88	0.77	0.98	0.95	0.95	1.00	0.74
IF	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.60	1.00	0.90	1.00	1.00	1.00	1.00	0.90
IFT	1.00	1.00	0.97	0.87	0.67	1.00	1.00	0.87	0.33	0.87	0.93	1.00	1.00	1.00	1.00	0.86
IFB	0.00	0.00	0.03	0.13	0.33	0.00	0.00	0.13	0.33	0.13	0.03	0.07	0.07	0.03	0.03	0.04
ID	-0.07	0.07	0.00	0.06	0.33	0.13	0.33	0.13	0.14	0.10	0.40	0.07	0.07	0.03	0.00	0.03
IV	0.02	0.02	0.00	0.12	0.10	0.04	0.10	0.04	0.14	0.10	0.18	0.02	0.04	0.04	0.00	0.15
ADV. JSL(2)	0.93	1.00	0.83	0.63	1.00	0.97	0.97	0.97	0.48	0.90	0.45	0.97	1.00	1.00	0.97	0.93
IF	1.00	1.00	1.00	0.90	1.00	1.00	1.00	1.00	0.60	1.00	0.90	1.00	1.00	1.00	1.00	1.00
IFT	1.00	1.00	0.70	0.70	1.00	1.00	1.00	0.50	0.00	0.70	0.90	1.00	1.00	1.00	1.00	0.90
IFB	0.00	0.00	0.00	0.30	0.00	0.00	0.00	0.50	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
ID	1.00	1.00	0.70	0.40	1.00	1.00	1.00	0.50	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
IV	-0.26	0.00	0.00	0.34	0.00	0.00	0.00	0.18	0.29	0.09	0.29	0.03	0.00	0.12	0.00	0.06
INT. JSL(3)	1.00	0.97	1.00	0.86	1.00	1.00	0.84	0.07	0.03	0.87	0.06	0.58	0.67	0.48	0.97	0.71
IF	1.00	1.00	1.00	0.90	1.00	1.00	1.00	0.22	0.10	1.00	0.10	0.80	1.00	1.00	1.00	0.80
IFT	1.00	1.00	0.90	0.20	0.80	1.00	0.80	0.00	0.00	0.80	0.10	0.50	0.70	0.20	0.90	0.60
IFB	0.00	0.10	0.00	0.40	0.20	0.00	0.20	0.22	0.10	0.20	0.10	0.30	0.30	0.30	0.10	0.20
ID	1.00	0.90	0.90	0.50	0.80	1.00	0.60	0.08	0.03	0.11	0.06	0.24	0.11	0.33	0.03	0.21
IV	0.00	0.03	0.00	0.19	0.09	0.00	0.14	0.06	0.18	0.03	0.01	0.05	0.03	0.03	0.03	0.21
DI1-2	0.05	-0.02	0.00	0.03	-0.08	-0.05	0.09	0.18	0.35	-0.01	0.32	0.01	-0.05	0.02	0.03	-0.12
DI2-3	-0.02	0.01	0.00	0.60	-0.02	-0.05	0.05	0.48	0.40	0.01	0.70	0.40	0.03	0.03	0.03	0.03
DI3-3	-0.07	0.03	0.00	0.37	0.06	0.00	-0.05	0.68	0.45	0.03	0.38	0.38	0.13	0.38	-0.00	0.22

	I33	I34	I35	I36	I37	I38	I39	I40	I41	I42	I43	I44	I45	I46	I47	I48	I49	I50	I51	I52
GRADE1(1)	1.00	0.93	0.88	0.88	0.95	0.86	0.98	0.88	0.95	0.95	0.85	0.93	0.40	0.93	0.77	1.00	0.66	0.77	0.95	1.00
IF	1.00	0.90	0.60	0.67	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.57	1.00	1.00	0.79	1.00	1.00	1.00	1.00
IFB	1.00	0.90	0.60	0.67	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.43	0.80	0.93	0.79	1.00	1.00	1.00	1.00
ID	0.00	0.20	0.11	0.33	0.07	0.31	0.93	0.93	0.93	0.93	0.93	0.93	0.13	0.20	0.33	0.00	0.93	0.87	0.87	1.00
IV	0.00	0.06	0.24	0.10	0.04	0.12	0.02	0.02	0.04	0.04	0.10	0.06	0.24	0.06	0.15	0.00	-0.12	0.45	0.13	0.00
ADV_JSL(2)																				
IF	1.00	0.93	0.69	0.86	1.00	0.93	1.00	0.97	1.00	0.93	1.00	0.79	0.31	0.62	0.55	0.79	0.34	0.69	1.00	1.00
IFB	1.00	0.90	0.50	0.50	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.50	0.50	1.00	0.90	0.50	1.00	1.00	1.00
ID	0.00	0.90	0.60	0.60	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.10	0.60	1.00	0.50	0.10	1.00	1.00	1.00
IV	0.00	0.06	0.30	0.40	0.00	0.20	0.00	0.40	0.00	0.20	0.00	0.16	0.20	0.20	0.30	0.40	0.40	0.50	0.00	0.00
INT_JSL(3)																				
IF	0.97	0.87	0.32	0.10	0.90	0.68	0.97	0.90	0.94	0.74	0.55	0.55	0.00	0.16	0.06	0.68	0.06	0.19	1.00	1.00
IFB	1.00	0.90	0.50	0.10	1.00	1.00	1.00	1.00	0.90	1.00	0.60	0.50	0.00	0.20	0.20	0.60	0.20	0.50	1.00	1.00
ID	0.90	0.70	0.30	0.00	0.70	0.40	0.90	0.70	0.90	0.40	0.30	0.00	0.00	0.00	0.00	0.50	0.00	0.00	1.00	1.00
IV	0.10	0.20	0.20	0.10	0.30	0.60	0.10	0.30	0.00	0.50	0.30	0.90	0.00	0.20	0.30	0.30	0.20	0.50	0.00	0.00
IV	0.03	3.11	0.22	0.09	0.09	0.22	0.03	0.09	0.06	0.19	0.25	0.14	0.00	0.14	0.06	0.22	0.06	0.16	0.00	0.00
D11-3	0.00	0.00	-0.11	0.02	-0.05	-0.07	-0.07	0.01	-0.05	0.02	-0.12	0.14	0.09	0.31	0.22	0.21	0.62	0.06	-0.05	0.00
D11-3	0.03	0.06	0.26	0.79	0.05	0.18	0.07	0.07	0.05	0.21	0.34	0.77	0.40	0.77	0.70	0.32	0.80	0.67	-0.05	0.00
D12-3	0.03	0.06	0.37	0.77	0.10	0.23	0.03	0.06	0.06	0.19	0.45	0.63	0.31	0.46	0.49	0.12	0.18	0.50	0.00	0.00
	I53	I54	I55	I56	I57	I58	I59	I60	I61	I62	I63	I64	I65	I66	I67	I68	I69	I70	I71	I72
GRADE1(1)	0.44	0.60	0.84	0.91	0.83	0.85	0.88	0.84	0.90	0.93	0.95	0.95	0.66	0.47	0.91	0.72	0.55	0.93	0.93	0.91
IF	0.79	0.66	1.00	1.00	1.00	1.00	0.93	1.00	1.00	1.00	1.00	1.00	1.00	0.90	0.85	0.85	0.90	1.00	1.00	0.90
IFB	0.20	0.77	0.53	0.73	0.80	0.73	0.73	0.60	0.87	0.80	0.87	0.87	0.60	0.07	0.87	0.33	0.20	0.00	1.00	1.00
ID	0.59	0.59	0.47	0.27	0.20	0.27	0.30	0.40	0.13	0.20	0.13	0.13	0.20	0.60	-0.01	0.60	0.60	0.20	0.20	0.72
IV	0.25	0.34	0.14	0.08	0.06	0.10	0.10	0.14	0.06	0.06	0.04	0.04	0.12	0.25	0.08	0.20	0.25	0.06	0.06	0.06
ADV_JSL(2)																				
IF	0.83	0.88	0.93	0.93	0.97	0.87	0.87	0.83	0.83	1.00	1.00	0.97	0.90	0.34	0.45	0.31	0.35	0.45	0.86	0.45
IFB	1.00	0.90	0.90	0.80	0.90	0.90	0.90	0.70	0.70	1.00	1.00	0.90	0.80	0.90	0.90	0.70	0.90	0.90	1.00	0.90
ID	0.60	0.00	0.90	0.80	0.90	0.90	0.90	0.60	0.70	1.00	1.00	0.90	0.80	0.90	0.70	0.70	0.10	0.20	0.60	0.10
IV	0.40	0.90	0.00	0.20	0.10	0.10	0.10	0.40	0.30	0.00	0.00	0.10	0.20	0.60	0.70	0.20	0.10	0.20	0.60	0.10
IV	0.14	0.25	0.06	0.06	0.03	0.03	0.03	0.14	0.14	0.00	0.00	0.03	0.09	0.23	0.25	0.23	0.24	0.25	0.40	0.25
INT_JSL(3)																				
IF	0.26	0.03	0.71	0.42	0.97	1.00	0.84	0.29	0.42	0.87	1.00	0.90	0.61	0.06	0.16	0.06	0.03	0.19	0.77	0.00
IFB	0.70	0.00	0.60	0.10	1.00	1.00	0.80	0.40	0.50	0.90	1.00	1.00	1.00	0.20	0.40	0.20	0.10	0.40	0.90	0.00
ID	0.10	0.00	0.30	0.60	-0.10	0.00	0.10	0.00	0.10	0.00	0.00	0.00	0.20	0.00	0.00	0.00	0.00	0.00	0.80	0.00
IV	0.40	0.03	0.20	0.24	0.03	0.00	0.14	0.21	0.24	0.11	0.00	0.09	0.16	0.06	0.14	0.08	0.10	0.00	0.40	0.00
IV	0.19	0.03	0.21	0.24	0.03	0.00	0.14	0.21	0.24	0.11	0.00	0.09	0.16	0.06	0.14	0.08	0.10	0.00	0.40	0.00
D11-3	-0.39	0.32	-0.09	-0.02	-0.04	-0.09	-0.06	0.01	0.10	-0.07	-0.05	-0.01	-0.04	0.12	0.46	0.41	0.16	0.46	0.07	0.46
D11-3	0.15	0.45	0.32	0.49	-0.04	-0.12	0.05	0.55	0.51	0.05	-0.05	0.06	0.05	0.40	0.25	0.65	0.33	0.74	0.15	0.91
D12-3	0.15	0.45	0.32	0.49	-0.04	-0.12	0.05	0.55	0.51	0.05	-0.05	0.06	0.05	0.40	0.25	0.65	0.33	0.74	0.15	0.91

BIBLIOGRAPHY

- ACTFL. 1986. *ACTFL proficiency guidelines*. Hastings-on-Hudson, New York: American Council on the teaching of Foreign Languages.
- Alderson, J. C. 1980. Native and nonnative speaker performance on cloze tests. *Language Learning* 30: 59–76.
- Bachman, L. F. 1982. The trait structure of cloze test scores. *TESOL Quarterly* 16: 61–70.
- . 1985. Performance on cloze tests with fixed-ratio and rational deletions. *TESOL Quarterly* 19: 535–56.
- Bormuth, J. R. 1965. Optimum sample size and cloze test length in readability measurement. *Journal of Educational Measurement* 2: 111–16.
- . 1967. Comparable cloze and multiple-choice comprehension tests scores. *Journal of Reading* 10: 291–99.
- Brown, J. D. 1980. Relative merits of four methods for scoring cloze tests. *Modern Language Journal* 64: 311–17.
- . 1983. A closer look at cloze: Validity and reliability. In *Issues in language testing research*, ed. J. W. Oller, Jr. (pp. 237–250). Rowley, Mass.: Newbury House.
- . 1984. A cloze is a cloze is a cloze? In *On TESOL '83*, eds. J. Handscombe, R. A. Orem and B. P. Taylor (pp. 109–119). Washington, D. C.: TESOL.
- . 1988. Tailored cloze: Improved with classical item analysis techniques. *Language Testing* 5: 19–31.
- . 1991. What text characteristics predict human performance on cloze test items? *Proceedings of the third conference on second-language research in Japan*. Tokyo: International University of Japan.
- . Forthcoming. *Language testing: A practical guide to proficiency, placement, diagnostic and achievement testing*. New York: Prentice Hall.
- Chapelle, C. A. and R. Abraham. 1990. Cloze method: What difference does it make? *Language Testing* 7, no. 2: 121–46.
- Chavez-Oller, M. A., T. Chihara, K. Weaver, and J. W. Oller, Jr. 1985. When are cloze items sensitive to constraints across sentences? *Language Learning* 35: 181–206.
- Chihara, T., J. W. Oller, Jr., K. Weaver, and M. Chavez-Oller. 1977. Are cloze items sensitive to constraints across sentences? *Language Learning* 27, no. 1: 63–73.
- Darnell, D. K. 1970. Clozentropy: A procedure for testing English language proficiency of foreign students. *Speech Monographs* 37: 36–46.
- Hinofotis, F. B. 1987. Cloze testing: An overview. In *Methodology in TESOL—A book of readings*, eds. M. Long and J. Richards (pp. 412–417). Rowley, Mass.: Newbury House.
- Jonz, J. 1987. Textual cohesion and second-language comprehension. *Language Learning* 37: 409–38.
- . 1990. Another turn in the conversation: What does cloze measure? *TESOL Quarterly* 24, no. 1: 61–83.
- Klein-Braley, C. 1985. A cloze-up on the C-Test: A study in the construct validation of authentic tests. *Language Testing*, 2: 76–104.
- Makino, S. and M. Tsutsui. 1986. *A dictionary of basic Japanese grammar*. Tokyo: Japan Times.
- Markham, P. L. 1985. The rational deletion cloze and global comprehension in German. *Language Learning* 35, no. 3: 423–30.

- Miller, G. R. and E. B. Coleman. 1967. A set of thirty-six prose passages calibrated for complexity. *Journal of Verbal Learning and Verbal Behavior* 6: 851-54.
- Oller, J. W., Jr. 1979. *Language tests at school: A pragmatic approach*. London: Longman.
- Shanahan, T., M. Kamil, and A. W. Tobin. 1982. Cloze as a measure of intersentential comprehension. *Reading Research Quarterly* 17, no. 2: 229-55.
- Taylor, W. L. 1953. "Cloze procedure": A new tool for measuring readability. *Journalism Quarterly* 30: 414-38.
- . 1957. "Cloze" readability scores as indices of individual differences in comprehension and aptitude. *Journal of Applied Psychology* 41, no. 1: 19-26.
- 小川多恵子 (1992) 「プレースメントとしてのクローズテスト」, 『日本語教育論集』8号, pp. 201-14, 筑波大学留学生センター.
- 芝 祐順 (1957) 「読みやすさのはかり方——クローズ法の日本語への適用」, 『心理学研究 (Japanese Journal of Psychology)』28号, pp. 67-73.
- 秦 喜美恵 (1983) 「クローズ法による日本語能力の測定」, 国際基督教大学学報『教育研究』25号, pp. 141-77.
- (1990) 「日本語クローズテストから日本語変形 C テストへ——採点法の問題を中心に」, 『言語文化論集』XI-2, pp. 213-25, 名古屋大学総合言語センター.
- 山下早代子・小川早百合 (1994) 『上級日本語教科書: インタビュープロジェクト』, くろしお出版.