DEVELOPING JAPANESE SKILLS TEST: 
THEORETICAL FRAMEWORK FOR 
A STANDARDIZED PROFICIENCY TEST

A Thesis
Presented in Partial Fulfillment of the Requirements for 
the Degree Master of Arts in the 
Graduate School of The Ohio State University

By 
Masayuki Itomitsu, B.S.

* * * * *

The Ohio State University
1996

Master’s Examination Committee:
Dr. Mari Noda, Adviser
Dr. Galal Walker

Approved by

[Signature]
Adviser

Department of East Asian 
Languages and Literatures
ABSTRACT

This thesis describes the theoretical foundation of the Japanese Skills Test (hereafter J-SKIT) project, which aims to produce a reliable, easily administered and scored standardized proficiency test in Japanese. While there has been a solid increase in the number of students learning Japanese in U.S. high schools and colleges, the field lacks a global measurement of students' proficiency in both the receptive (listening and reading) and productive (speaking and writing) skills, especially for the students at the lower end of proficiency. Availability of such a comprehensive test is needed in the field for the purpose of the placement of students, program evaluation, and follow-up assessment of the students. J-SKIT is designed to respond to these urgent needs by providing a comprehensive proficiency test that would be used for evaluation, comparison, and articulation of students and language programs.

J-SKIT is targeted at students with lower proficiency (Advanced and below in ACTFL Guideline standards), and it is designed to test the learners' proficiency in every-day, non-technical interactions in simulation. The J-SKIT project also aims to produce a comprehensive proficiency test, which would be easily administered in any proficiency-oriented institution and easily graded by the test provider. J-SKIT contains five sections and is divided into two parts: Part 1 consists of Listening Comprehension, Structure, and Reading; Part 2, Speaking and Writing. Each part takes approximately 50 minutes to administer, and the test materials are provided and scored by the National Foreign
Multiple choice questions are used for the receptive skills and structure, and open-ended questions for the productive skills. Each section is designed to maintain a single-skill testing format. Commonly used textbooks in the U.S. as well as other research results are considered to construct the content specification, while the test format is based on the simulated interaction and task performance to measure proficiency — what a test taker can accomplish in the target culture with the target language.

These format and content specifications have been developed based on some basic assumptions and production principles of proficiency-oriented language instructions. J-SKIT assumes that learning the Japanese language is to know and practice using the language in Japanese culture through interaction in a variety of contexts, and such communicative skills should be tested in interaction and task performance simulated by English descriptions and recorded voices.

Scoring for Part 1 (the receptive skills and structure) is done by the scanning machine, and Part 2 (the productive skills) is graded in terms of linguistic performance, socio-cultural appropriateness, discourse/text skills, and pronunciation/orthographic representation. Feedback consists of the earned points in each section accompanied by grading statements for interpreting of the scores. J-SKIT provides more detailed feedback in each skill.

Results of the second edition administration of J-SKIT at The Ohio State University in March, 1996, indicates that the test is a good indication of students' proficiency, and is reasonably effective in discriminating the levels of the students. The analyses also show that there is a correlation between the scores in the receptive skills and the productive skills, and the Structure section score is not as important as the Listening Comprehension and Reading scores to hypothesize the scores in the productive skills. After the discussion for the future development of J-SKIT, it is proposed that J-SKIT promises to serve well for
the need in the field — to evaluate and compare students and programs of Japanese in U.S. high schools and colleges.
Dedicated to my mother
ACKNOWLEDGMENTS

This thesis and the project owe appreciation to many people who were involved in and supported the development of both the project and the thesis. This project was funded through the Department of Education Title VI. I wish to thank, first of all, Dr. Diane Birckbichler of the National Foreign Language Resource Center at The Ohio State University for her guidance and support in the development of the project.

I also wish to thank Ms. Penny Corbett of Exchange: Japan, for providing the opportunity to field-test the first version of the test, and for her support in administering the test.

I am also indebted to Ms. Virginia Marcus of Washington University in St. Louis for her guidance and generous support in administering the first version of the test. The data collected at her institution have been very helpful for the development of the project, and her encouragement and suggestions for the future development of the test have been highly valuable.

I am grateful for Mr. Hiroaki Kawamura, for his assistance in the administration of the second version of the project product, and for his intellectual support and encouragement. Without his comments and support, the completion of the second version and the administration thereof would not have been possible.

I also wish to thank the many people who gave assistance in the administration and scoring of the second edition at The Ohio State University, including Dr. Charles Quinn, Director of the East Asian Language Program, Ms. Kazumi Ikeda and Ms. Yuko Okutsu, Instructors of Japanese, Mr. Hiroaki Kawamura, and Ms. Misako Suzuki, Ms. Asako
Yamaguchi, and Mr. Kazuo Yamada, Graduate Teaching Associates of Japanese. I am also grateful to Ms. Kyoko Omori, graduate student in Japanese language and literature, for her comments and encouragement.

I also wish to thank Ms. Sanae Eda, Mr. Tom O’Conner, and Mr. Atsushi Onoe, who participated in the production of the audio tapes for the test project. Their assistance and comments were invaluable for the development of the test.

I would like to thank the students of Japanese all over the United States, who have taken the two field-test versions, particularly those at Washington University in St. Louis and at The Ohio State University. Without their cooperation this project would not be meaningful and successful.

I am also indebted to Ms. Virginia Replogle of The Ohio State University for the statistical analyses. Her involvement in this project has been very helpful, and will be indispensable to the future development of the test.

I also wish to thank Mr. Jeffrey Angles and Dr. Müge Galin, also of The Ohio State University, for their guidance and support in proofreading the documents. Their excellent skills as proofreaders have improved the stylistic aspect of this thesis tremendously.

I thank Mr. Sheldon Safko for his encouragement and support while I wrote this thesis.

Finally, I wish to thank Dr. Mari Noda, my advisor and the project director, for her intellectual support, encouragement, and examination of every part of the project and the thesis. Her profound knowledge, experiences, and teaching philosophies in the field of Japanese language pedagogy were always fascinating and enlightening to me. Her intellectual support and patient guidance for both qualitative and stylistic aspects of the project and this thesis cannot be appreciated enough.
VITA

March 30, 1970 ................... Born -- Okinawa, Japan

1992 ............................ B.A. in English Language and Literature,
The University of the Ryukyus, Japan.

Northeast Missouri State University, Kirksville.

1994 ............................. B.S. in English (Linguistics Emphasis),
Northeast Missouri State University, Kirksville.

1992 - present ................... Graduate Teaching and Research Associate,
The Ohio State University, Columbus.

PUBLICATIONS


FIELDS OF STUDY

Major Field: East Asian Languages and Literatures
(Japanese Language Pedagogy)
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>Abstract</strong></td>
<td>ii</td>
</tr>
<tr>
<td></td>
<td><strong>Dedication</strong></td>
<td>v</td>
</tr>
<tr>
<td></td>
<td><strong>Acknowledgments</strong></td>
<td>vi</td>
</tr>
<tr>
<td></td>
<td><strong>Vita</strong></td>
<td>viii</td>
</tr>
<tr>
<td></td>
<td><strong>List of Tables</strong></td>
<td>xii</td>
</tr>
<tr>
<td></td>
<td><strong>List of Figures</strong></td>
<td>xiii</td>
</tr>
<tr>
<td></td>
<td><strong>Chapters:</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1. Japanese Language in the Context of the U.S. Pedagogy and Needs for a Standardized Test</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.1. Introduction</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1.2. The Current State of the Field in the United States:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.2.1. Increase in Number of Students and Programs</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1.2.2. Diversity of Japanese Programs and Students</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>1.3. Movement Toward Proficiency-Oriented Instruction</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>1.4. Nature of Learning Japanese in High Schools and Colleges</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>1.4.1. Japanese as a Less Commonly Taught Language</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>1.4.2. Japanese as a Category IV Language</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>1.4.3. The Role of Formal Instruction in Japanese Language Pedagogy</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>1.5. The Issue: The Need for a Standardized Test</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>1.5.1. Current Problems</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>1.5.2. The Need for a Standardized Test</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>2. Review on Currently Available Testing Materials</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>2.1. ACTFL Oral Proficiency Interview (OPI)</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>2.1.1. The Advantages of OPI</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>2.1.2. The Disadvantages of OPI</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>2.2. Simulated Oral Proficiency Interview (SOPI)</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>2.2.1. The Advantages of SOPI</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>2.2.2. The Disadvantages of JST</td>
<td>25</td>
</tr>
</tbody>
</table>
2.2.3. The Disadvantages of Pre-JST ........................................ 26
2.3. SAT II Subject Test in Japanese with Listening (SAT-Japanese) .................................................. 27
2.3.1. The Advantages of SAT-Japanese ................................. 27
2.3.2. The Disadvantages of SAT-Japanese ............................ 28
2.4. Conclusion ............................................................. 30

3. J-SKIT: Basic Assumptions and Project Principles .................................................. 31
3.1. Introduction .......................................................... 31
3.2. J-SKIT: Basic Assumptions ......................................... 31
3.2.1. Assumption I: Language is Used as a Tool for Communication .......................................................... 31
3.2.2. Assumption II: Language Use in Communication Has a Purpose ....................................................... 34
3.2.3. Assumption III: Communicative Competence can be Evaluated Through Task Performance ...................... 35
3.2.4. Assumption IV: Language Use Observes Cultural Conventions and Expectations .................................... 37
3.2.5. Assumption V: Authentic Language Use Must Be Promote and Encouraged ........................................ 38
3.2.6. Assumption VI: Communication Takes Place in Context ........................................................................ 40
3.2.7. Assumption VII: Language Use Must Consider Function, Structure, Vocabulary, and Context ................... 41
3.3. J-SKIT: Project Principles ............................................. 43
3.3.1. Principle I: J-SKIT is Targeted at Students in the United States ......................................................... 43
3.3.2. Principle II: J-SKIT Measures Performance in Simulated Contexts ..................................................... 43
3.3.3. Principle III: J-SKIT Measures All of the Four Skills Separately (Single-skill Testing) .............................. 45
3.3.4. Principle IV: J-SKIT is a Proficiency Test for Students of Lower Levels of Proficiency ............................ 46
3.3.5. Principle V: Tasks in J-SKIT Must be Realistic, Like Something Learners are Likely to Encounter in Daily Interactions in Japan ................................................................. 48

4. J-SKIT: Content Specification .................................................................................. 50
4.1. Introduction: J-SKIT as Function-based, Structurally Organized, and Contextually Realized ...................... 50
4.2. J-SKIT: Content Specification ........................................... 52
4.2.1. Function ................................................................ 52
4.2.2. Structure ................................................................ 55
4.2.3. Vocabulary ........................................................... 59
4.2.4. Context ................................................................ 61
4.2.5. Orthography .......................................................... 63

5. J-SKIT: Format Specification .................................................................................. 64
LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2. Skills Tested in J-SKIT J-SKIT intends to measure learners' ability to perform in Japanese in simulated settings to evaluate students' proficiency in each skill</td>
<td>48</td>
</tr>
<tr>
<td>4. J-SKIT Structure Specification N refers to a noun or a nominal, IA-i to an adjective or adjectival, NA-na for a na-adjective or na-nominal, V for a verb or verbals. A structure with an asterisk indicates that the structure covered in more than 3-5 out of 12 high school books, as reported by Japan Foundation, 1994</td>
<td>55</td>
</tr>
<tr>
<td>6. J-SKIT Orthography Specifications</td>
<td>63</td>
</tr>
<tr>
<td>7. Basic Framework of J-SKIT</td>
<td>67</td>
</tr>
<tr>
<td>8. J-SKIT Maximum Points Distribution for Each Section</td>
<td>78</td>
</tr>
<tr>
<td>10. ANOVA Analysis on Course Level vs. Total Score</td>
<td>93</td>
</tr>
<tr>
<td>11. Tukey HSD Multiple Comparisons of The Ohio State University</td>
<td>95</td>
</tr>
<tr>
<td>12. Correlation between SOPI scores and J-SKIT Part 1 Scores at Washington University</td>
<td>100</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Interaction Competence Model by Neustupny (1982: 53)</td>
<td>33</td>
</tr>
<tr>
<td>2. A Tentative Model for Sentence Production</td>
<td>42</td>
</tr>
<tr>
<td>3. J-SKIT Administration Procedures</td>
<td>65</td>
</tr>
<tr>
<td>4. Least Squares Mean for Each Level</td>
<td>94</td>
</tr>
<tr>
<td>5. J-SKIT Converted Section Scores and Levels of Students at The Ohio State University</td>
<td>97</td>
</tr>
</tbody>
</table>
CHAPTER 1

JAPANESE LANGUAGE PEDAGOGY IN THE CONTEXT OF THE U.S. AND NEEDS FOR A STANDARDIZED TEST

1.1. Introduction

The Japanese Skills Test (hereafter J-SKIT) project has been under development since 1994 as a project of the National Foreign Language Resource Center at The Ohio State University\(^1\). The project aims to produce a comprehensive proficiency test in Japanese designed for students learning Japanese as a foreign language in high schools or colleges. Since the project specifically focuses on the field of Japanese language pedagogy, a brief survey of the current state of Japanese teaching and learning in high schools and colleges would be an appropriate starting point. This chapter attempts to overview the field of Japanese language pedagogy in U.S. high schools and colleges in order to examine the current situation and consequent needs emerging in this particular arena.

1.2. The Current State of the Field in the United States: Survey

1.2.1. Increase in Number of Students and Programs

As of 1993, according to a survey conducted by the Japan Foundation Japanese Language Institute, there are a total of 50,420 people learning the Japanese language in the United States, including 29,516 (59% of the total) secondary and 16,951 (34%)

\(^1\)The J-SKIT project has been established by the Department of Education Title VI center funding.
post-secondary educational level students. More than 600 institutions provide formal language instruction (380 secondary and 181 post-secondary). Of these, 172 secondary and 39 post-secondary institutions started their language program after 1990 (the Japan Foundation Japanese Language Institute, 1995: 82). Considering the low return rate of the survey (36.4%), it is reasonable to expect more students and institutions to be currently involved in Japanese language learning and teaching. The following table summarizes the increase in the number of students and institutions of Japanese since 1990:

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1990</td>
<td>12,667</td>
<td>247</td>
<td>16,818</td>
<td>136</td>
<td>126</td>
<td>4</td>
<td>29,661</td>
<td>387</td>
</tr>
<tr>
<td>1993</td>
<td>29,516</td>
<td>380</td>
<td>16,951</td>
<td>181</td>
<td>3,953</td>
<td>41</td>
<td>50,420</td>
<td>602</td>
</tr>
</tbody>
</table>

Table 1. Number of Students and Institutions in the U.S. Japanese Language Education. Source: The Japan Foundation, 1995: 74-75, 148-149

As indicated in the table, number of the students and availability of the formal instruction of Japanese language in the United States have shown a steady growth since 1990, and this is especially true in the field of secondary education. Before, it seemed that “the United States is unique in the world in placing onto colleges and universities a major part of the burden of educating its citizen” (Brecht and Walton, 1995), but this situation has been changing gradually with the efforts of many conferences, institutions, and

---

2 As mentioned previously, it is reasonable to expect more students and institutions to be involved in Japanese language education in the U.S., considering the low return rate of the survey and also a research by Jorden and Lambert (1992).
individuals. These include the NCSTJ (National Council of Secondary Teachers of Japanese) founded in 1991 to magnify the Japanese language instruction in secondary education and some teacher training programs for high school teachers to further improve the quality of secondary education.

All of the above facts report a significant increase in the number of students and programs. They also seem to endorse an optimism in the field that "enrollments in Japanese have soared in the past ten years and Japanese is apparently in the process of becoming one of the so-called 'mainstream languages' taught in U.S. colleges and universities" (Wetzel, 1995). As the field expands, the consequent diversity among institutions and students becomes even more evident in the current Japanese language pedagogy in the United States and brings another variable.

1.2.2. Diversity of Japanese Programs and Students

After the "boom" of language learning that took place in the 1980s, the diversity of the students and formal language curricula in the United States is even more remarkable today. At the college level, besides the steady growth of the number of regular programs, more and more programs have been offering unique opportunities that are flexible and tailored toward the students' needs and demands. More specialized programs such as Japanese for business persons and engineers are available at various institutions and more summer intensive language programs provide opportunities to learn the language in a short period of time. At the high school level, the significant increase in the number of institutions that started offering a course in Japanese language itself is a remarkable sign of the wider availability of Japanese language instruction. Satellite programs have been developing in some states to enhance long-distance learning, and new textbooks have also been developed, such as *Situational Functional Japanese* (1991, 1992) and *Yookoso!* (1994). These add more options to the choice of the course textbooks, providing

The diversity among programs seems to reflect the diverse interests and needs of students of Japanese, and also suggests the diversity of areas of interaction between Japan and the U.S. Unlike the situation in the early 1980s, Japanese language study today attracts not only Japanologists and businesspersons, but also people in various fields of study interests--whether they are interested in Japanese religion, martial arts, TESL, architecture, art, or *animé*. The Japan Foundation survey verifies this diversity of the study interests of the students of Japanese and reports that the major fields of the students of Japanese in the U.S. and Canada ranges from Japan studies (16.0%) to the humanities (23.8%), social sciences (26.1%), natural sciences (18.3%) and other fields (15.8%) (Japan Foundation Japanese Language Institute, 1995: 38). This wide distribution of student interests suggests that more interaction and interest between Japan and the United States are taking place not only in academic, public, and diplomatic fields but also in various aspects of more practical, personal, and domestic areas. As Brecht and Walton (1995) describe, "communication among Americans [is] domestically characterized increasingly by cross-cultural communication: in the workplace, in the schools, in hospitals, in the courts, in social services, and more importantly in our neighborhood, as peoples of different cultural discourses are trying to make themselves understood to their neighbors." As more cross-cultural interaction takes place in various fields of life in the U.S., an increasing number of the students of Japanese seems to show signs of a national response to a rather rapid change of life and a widespread endeavor to catching up with the subsequent needs of dealing with foreign languages, including Japanese.
1.3. Movement Toward Proficiency-oriented Instruction

Another significant change in the field of Japanese language pedagogy is the ongoing movement toward a proficiency-oriented instruction. "In the past several years, the concept of 'proficiency' and 'proficiency-oriented instruction' has sparked a great deal of interest and discussion among theorists and practitioners in second-language education" (Omaggio, 1985: xi). The emergence of the so-called communicative approach can be seen as a response to the field under this movement toward proficiency, trying to meet the needs and interests of the variety of students who wish to gain practical knowledge and skills in Japanese (Okazaki and Okazaki, 1990: 4-9). Although terms such as "proficiency" and "communicative approach" are sometimes loosely used,³ the field of Japanese language pedagogy is clearly moving toward proficiency-oriented instruction. Proficiency-oriented instruction is summarized as "performance-based approaches that have a common focus on students' ability to use the language [...], that focuses on functional, communicative activities directed toward the goal that students should be able to use their knowledge in real conversation and authentic, contextualized settings" (The College Board, 1996: xii). The J-SKIT project incorporates this definition of proficiency-oriented instruction.

The J-SKIT project also adopts the definition of proficiency which appears in Omaggio (1985: 2) and expands on that of the American Heritage Dictionary of the English Language (1978: 1045), as ""performing in a given art, skill, or branch of knowledge with expert correctness and facility’ and further specifies that the term implies ‘a high degree of competence through training.’"" Proficiency can be further recognized as knowledge and skills in the target language that enables students to actually use the language in communicating with or obtaining information from the language in a

³ See Matsuoka (1991) and Okazaki and Okazaki (1990) for further discussions of the term "communicative approach."
culturally appropriate fashion. This definition of the term will be discussed further in Chapter 3.\textsuperscript{4}

In such a movement, at least two recent epoch-making works deserve nationwide credit for enhancing the basic understanding and encouragement of proficiency-oriented instruction in the field of Japanese language pedagogy. One is \textit{A Framework for Introductory Japanese Language Curricula in American High Schools and Colleges} (hereafter \textit{Framework}) published by the National Foreign Language Resource Center in 1993. This work attempts “to assist the emergence of consensus on the pedagogical principles and practices that might best inform the design and management of introductory Japanese language instruction” (National Foreign Language Center, 1993: v). The other major contribution to the field is \textit{ACTFL Japanese Proficiency Guidelines} (hereafter \textit{ACTFL Guidelines}) published in 1987. This guideline has been created in accordance with the generic guidelines in other languages and is intended to “represent a hierarchy of

\textsuperscript{4} Omaggio (1986: 35-36) further proposes five hypotheses that “relate to the way classroom instruction might best be organized when the students’ ultimate, long-range goal is \textit{Superior} proficiency in a second language. The following is a list of such hypotheses:

Hypothesis 1. Opportunities must be provided for students to practice using the language in a range of contexts likely to be encountered in the target culture.

Corollary 1. Students should be encouraged to express their own meaning as early as possible after productive skills have been introduced in the course of instruction.

Corollary 2. A proficiency-oriented approach promotes active communicative interaction among students.

Corollary 3. Creative language practice (as opposed to exclusively manipulative or convergent practice) must be encouraged in the proficiency-oriented classroom.

Corollary 4. Authentic language should be used in instruction wherever and whenever possible.

Hypothesis 2. Opportunities should be provided for students to practice carrying out a range of functions (task universals) likely to be necessary in dealing with others in the target culture.

Hypothesis 3. There should be concern for the development of linguistic accuracy from the beginning of instruction in a proficiency-oriented approach.

Hypothesis 4. Proficiency-oriented approaches should respond to the affective needs of students as well as to their other cognitive needs. Students should feel motivated to learn and must be given opportunities to express their own meanings in a nonthreatening environment.

Hypothesis 5. Cultural understanding must be promoted in various ways so that students are prepared to live more harmoniously in the target-language community.

The J-SKIT project has been designed along with these hypotheses, and J-SKIT has its own descriptions on the project assumptions. See Chapter 3 for details.
global characterization of integrated performance in speaking, listening, reading and writing” (ACTFL, 1987: 589). The significance and degree of impact of the latter to the field is particularly eminent, and more programs are using the criteria described in the ACTFL Guidelines for program goal statements (intended outcome) and/or entrance requirements. These two critical works are examined in greater detail in later chapters. It is sufficient to say at this point, however, that despite these endeavors to promote proficiency-oriented instruction, it still seems to be the case that diversity among the Japanese programs prevails to a significant degree and consensus within the field is still at the early stage of maturation. This is due to the relatively short history of Japanese language instruction in the United States and the unique nature of learning Japanese. The latter case is analyzed in the immediately following section.

1.4. Nature of Learning Japanese in High Schools and Colleges

Before moving on to the need analysis for the testing project, it is worthwhile to examine the special nature of Japanese language learning and teaching in the United States. This section tries to describe the nature of Japanese formal instruction in high schools and colleges and re-addresses the proposition by Wetzel (1995) that “the difficulty of Japanese for native speakers of English is a factor that is too often ignored when we look at enrollments, attrition, student placement in courses, curriculum and materials, and not least of all testing.”

1.4.1. Japanese as a Less Commonly Taught Language

With the increase in the number of students and institutions, the field of Japanese language pedagogy now seems to be ready to expand on the national scale and to gain attention equal to that of other mainstream foreign languages. There exist, however,
some characteristics unique to Japanese and a few other languages that are often grouped together as Less Commonly Taught Languages (Walker, 1989). The characteristics of these languages necessitate special considerations in teaching Japanese in the U.S. education system. The Less Commonly Taught Languages are conventionally referred to all but the five most commonly taught languages (Spanish, French, German, Italian, and Russian), although Russian is often included among Less Commonly Taught Languages because its critical nature—a language "perceived to be of strategic importance by the United States Government" (Walker, 1989: 111). The Less Commonly Taught Languages are characterized by their unique linguistic features, cultural features, and the consequent difficulty to achieve a high level of proficiency in them (Walker, 1989). It is of course exciting news to the field that the number of students in Japanese has been increasing remarkably. As reported in Jorden (1992: 149), "Japanese is currently increasing more rapidly in the United States than that of any other foreign language."

Nevertheless, as of 1991, still "less than 1% of the whole population of secondary education students in the United States are studying Japanese, and less than 2% of the secondary academic institutions are offering Japanese" (Jorden and Lambert, 1991; cited in Japan Foundation Japanese Language Institute, 1995: 32). "The question of whether Japanese can in fact be mainstreamed is fraught with issues that have not been examined seriously enough," (Wetzel, 1995) and consequently Japanese language testing is one of the areas that deserves closer attention and investigation. More discussion on the nature of Less Commonly Taught Languages, including Japanese, follows.

1.4.2. Japanese as a Category IV Language

One of the most important things to be mentioned here about the nature of learning Japanese is that for the native speakers of English it is more time-consuming to achieve a
certain degree of proficiency in Japanese than other commonly taught languages. This is because of both linguistic and cultural features (Walker, 1989), because the Japanese language is not a cognate of English, and because it contains a distinct writing style and cultural elements which are "truly foreign" to English-speaking students (Jorden and Walton, 1987; Jorden, 1992; National Foreign Language Center, 1992). In fact, according to the classification of the Foreign Service Institute (FSI), Japanese belongs to the most difficult categorization: "Group IV" languages for English natives, which requires almost three times as much time and effort to achieve the same level of speaking proficiency as those "Group I" languages such as Spanish, French, German, and Italian (Liskin-Gasparro, 1982, cited in Omaggio, 1986: 21). According to the study, a learner of Japanese with an average language aptitude is expected to achieve level 2 of the FSI proficiency scale after 1320 hours (44 intensive weeks) of study, whereas a learner of Group I languages reaches the same level of proficiency in only 480 hours (16 intensive weeks) (Liskin-Gasparro, 1982, cited in Omaggio, 1986: 21). Considering that this report is based on the program research for diplomats, who are seriously in need of mastering foreign language skills and are paid to be full-time students of a language, this difficulty of Japanese "has far-reaching implications for teaching and curriculum design" (Wetzel, 1995: 1) in U.S. high schools and colleges.

1.4.3. The Roles of Formal Instruction in Japanese Language Pedagogy

Despite the difficulty and time-consuming nature of learning Japanese, most Japanese programs in high schools and colleges in the U.S. treat it just like any other foreign language in terms of class hours and length of the curriculum. This may be because of the administrations' lack of realization of the significant differences among the Group I and Group IV languages, or the avoidance of dealing with the fact that "this added
difficulty [of Group IV languages] can disturb the equanimity of an institution" (Walker, 1987: 118). In any case, learners of Japanese who approach the language with the expectation that they can acquire an equal proficiency in the language as easily and as fast as in other commonly taught languages are soon to realize that their expectation is wrong. If students are not properly informed about the time-consuming nature of Japanese by instructors, some students may eventually leave the program with disappointment or desperation. Students who are studying Japanese for economic reasons, for example, and who wish to negotiate with Japanese lawyers and businesspersons may give it up "when they [realize] that two years of Japanese is about enough to negotiate whether you want cream and sugar in your coffee" (Steinhoff, 1993: 5). In fact, the average length of study in Japanese in the U.S. secondary education is 380 hours, or approximately 2.8 years (Japan Foundation Japanese Language Institute, 1995: 84-85). Furthermore, in 1983, 78% of the U.S. college students ceased their formal study in two years and did not proceed to the third-year level of instruction (Jorden, 1985, cited in Okamura and Unger, 1987: 120, translation mine). It is a responsibility of the curriculum to explain to the students the nature of studying Japanese and to provide them with realistic goals and expectations toward learning in the limited amount of time in formal instruction (National Foreign Language Center, 1993: 19-21).

This state of affairs results in a very challenging predicament: most of the students of Japanese end up with a low degree of proficiency within the short periods of their formal instruction in high schools and colleges. Given that learning Japanese is so time-consuming, school curriculums can offer only a small portion of the learning environment required for achieving a higher level of proficiency. Unfortunately, sometimes these givens are considered only slightly. It is important to realize the implication of the nature of learning Japanese or any language; in other words, that "language study is a life-long
enterprise that ideally goes on beyond the school years" (Walker and McGinnis, 1995: 1; National Foreign Language Center, 1992: 20). It is crucial for anyone who is involved in Japanese language pedagogy to recognize that "the proper role of formal instruction is to enhance and sustain [the] career" (Walker and McGinnis, 1995: 1) of students who would continue even after the end of the formal curriculum.\footnote{Given that studying Japanese requires a great deal of time and "that one should study [a language] in all skill modalities to a high level of competence presents somewhat a fiction," Brecht and Walton (1995) propose an innovative program that contains more information on theory and observance on intercultural interaction and less practice of the actual language use in the classroom. Neustupny (1991) also proposes similar ideas for a future language program, using a concept of “Japan Literacy.” These innovations would become possible with extensive support from self-managed study using modern technology, including interactive computer software. Such an approach is designed to accommodate more students, to grant more practical knowledge to interact with the Japanese, and to prepare the students for further learning. This thesis, however, is concerned with the current language curricula in a traditional sense and assumes that the classroom hours are spent mostly practicing the actual use of the language and that the main goal of the program is to provide the students the opportunity to practice using the language toward the goal of achieving some degree of proficiency.}

1.5. The Issue: The Need for a Standardized Test

Having overviewed the present situation of the Japanese language pedagogy in the U.S., the following section focuses on the current problems arising from the diversity of the programs and students as well as on the need for a standardized test which can be used in the field at large. It is proposed that J-SKIT, when it is fully developed, would reduce some of the problems and would enhance comparison among students and programs. It is also hoped that the J-SKIT project can strengthen national cohesion and accelerate the ongoing movement toward proficiency-oriented instruction.

1.5.1. Current Problems

As discussed above, only a couple of years of language study is not enough for students to achieve the levels of proficiency to enable them to “apply that knowledge in any significant way” in both the everyday and professional aspects of life (Brecht and
Walton, 1995) and most of the institutions offer only a short length of instruction. Because of this, it is natural to expect that despite formal instruction at a high school and college, most students end up with a degree of proficiency lower than what has been defined as an "Advanced" level of proficiency by ACTFL Guidelines. Consequently, students who wish to continue studying the language and to achieve higher levels of proficiency have to seek ways to continue studying the language when they have completed the original program. Most likely, this involves transferring from one program to another.

This is precisely where problems arise, due to the diversity of the programs mentioned in the previous sections. Even after the creation of national standards by ACTFL and guidelines by the National Foreign Language Center, institutions are currently allowed to employ in their programs whatever procedures they wish (Japan Foundation Japanese Language Institute, 1995: 31), and diversity in teaching methods and content of curricula still prevails. In Japanese and other less commonly taught language programs, as Jorden (1992: 153) points out,

we find the worst and the best [...] in too many institutions, instructors with no training in how to teach a language proceed by individually developed methods based on assumptions, intuitions, and myths that have no basis in fact and never stood up to the test of empirical research. At the other end of the spectrum are the programs taught by trained professionals, whose underlying philosophy, goals, methodology, and classroom techniques are coordinated in the teaching of an organized curriculum, in which everything fits together to produce learners with meaningful competence.

Since there is no solid consensus in the field at this time, this spectrum presents a very wide range in quality, content, and practices among programs. This poses challenges for many students transferring into new programs. The freedom of teaching methodology and the diversity among institutions have had both positive and negative effects in the field. Some positive effects are freedom of choice, flexibility, and a greater chance for
renovation, while some negative effects include the lack of national coherence and
difficulty, if not the impracticality or impossibility, of comparing students from different
programs based solely on the length of study.

Currently many programs are practicing assessment that is "typically linear,
[involving] learning the structure of the language, and are measured by years of
instruction completed rather than by demonstrated competence of the learner" (The
College Board, 1996: 7). However, when some students come from outside of the
program with different experiences, the difference in their knowledge and skills is highly
unpredictable. “Since the content of the program differs among colleges, the knowledge
mastered by the students also differs in a significant degree after the 1-year-long courses.
There are students who cannot read even hiragana nor katakana after a year of instruction.
There are cases which 1 year of study at University A corresponds to two years of study
at University B, and even the credit hours may vary” (Soga, 1987: 12-22, translation
mine). It is apparent that such a linear assessment and/or placement does not work
efficiently, particularly in proficiency-oriented programs.

High school students who come to college with backgrounds obtained by studying
Japanese in high school are another concern that occurs more frequently than in the past.
The gap in curricula between secondary and post-secondary levels shows just as much
diversity as that among colleges. Coordination between the two levels becomes more
crucial than ever (Japan Foundation, 1994: 1) in order to promote a smooth transition
from high school to college education and eventually to establish a coherent eight-year
perspective of language training across secondary and post-secondary education. More
efficient and smooth articulation between the secondary and post-secondary levels of
proficiency-oriented instruction requires a well established means of assessment of
students' proficiency.
The increasing number of the summer intensive language programs increases the possibility of student transfers even more, resulting in further complications in placement and articulation in the field. Furthermore, it is possible for students in a single program to have had different instruction as the other students even within the same program (different instructors, different teaching methodology, or different textbooks). Students with no formal instruction but with an extensive exposure to the language, in Japan or from relatives and friends, also pose additional complexity.

To summarize, Japanese language programs and the knowledge, experience, and skills of students are so different among U.S. institutions that the prospect of advancing cohesion and articulation among them presents a challenge. The field needs an objective test in proficiency that is mutually agreed upon by proficiency-oriented programs in order to enhance the emerging consensus established by ACTFL and the National Foreign Language Center to provide a valid means of placement, and to increase coordination between high schools and colleges.

1.5.2. The Need for a Standardized Test

This diversity in experience and fluidity of students thus makes it imperative for a program to have some means to measure the knowledge and skills of their students and thus accordingly understand the knowledge and skills of each student, in order to place the students in an appropriate level of instruction. Currently, however, each program has its own particular placement regulations or locally produced placement test that is fairly independent of other programs. While in such a diversified field it is inevitable for a program to have its own criteria to maintain its standards and goals, such independence exacerbates the difficulty of establishing a meta-institutional perspective of Japanese language learning and teaching. If the field agrees to proficiency-oriented instruction and sets certain goals of proficiency that would be generally applicable to most of the
programs, the field needs to have a means of measurement which is not dependent on one particular program, textbook, or teaching methodology. This means of measurement would also need to reflect the basic emphasis within the field on proficiency testing. The J-SKIT project aims to produce such a comprehensive proficiency test that meets the current needs of the field.

One of the objectives of the J-SKIT project is to provide a means for placement that would be useful for any proficiency-oriented program. Given that most students learning Japanese in high schools and colleges are below the Advanced level of proficiency established by the ACTFL Guidelines, the project aims at the lower end of proficiency and attempts to provide a finer discrimination among Novice, Intermediate, and Advanced levels in each of four skills: listening, speaking, reading, and writing. Such a program-independent proficiency test would also be useful in comparing, evaluating, and diagnosing students and programs in the United States.

J-SKIT, when completed, will also contribute to further solidify the emerging consensus in the field on proficiency-oriented instruction. With Framework and ACTFL Guidelines published and already available to the field, a reliable and valid proficiency test would enhance the awareness and emphasis on proficiency to further the nationwide movement toward proficiency-oriented instruction.
CHAPTER 2

REVIEW OF CURRENTLY AVAILABLE TESTING MATERIALS

There are some nationally acknowledged test materials available in the United States that have been influential in the field for several years.\(^1\) While there are some program-independent proficiency tests available to the field, none of them provides a comprehensive assessment with detailed feedback and none are conveniently available at a reasonable price for any institution. This chapter focuses on four such currently available tests, namely the ACTFL Oral Proficiency Interview, the Simulated Oral Proficiency Interview, which consists of the Japanese Speaking Test and the Preliminary Japanese Speaking Test, and the SAT II Subject Test in Japanese with Listening. This chapter critically investigates their advantages and disadvantages.

2.1. ACTFL Oral Proficiency Interview (OPI)

OPI has been developed in accordance with the ACTFL Guidelines to examine the speaking proficiency of the students through a short (15-30 minute) oral interview. This interview is designed as a first encounter between a Japanese interviewer and the testee. The conversation is designed to be natural and smoothly flowing as a normal conversation, executed by a licensed OPI interviewer.

OPI consists of four stages of evaluation during the interview: (1) a warm-up (to familiarize the testee with the language and make him/her at ease) (2) a level-check (a set of carefully organized questions to advance step-by-step the level of difficulty and to determine what the interviewee can do with the language in order to find the floor of the proficiency level), (3) probing (finding out what the interviewee cannot do and determine the ceiling of his/her proficiency), and (4) winding-down (confirming the interviewee’s proficiency level by lowering the level of content and providing him/her with a sense of accomplishment) (Makino, 1991: 24-25; NFLRC on-line). In addition to the formal interview, one or two prepared task cards are used in order to do a role-play during the interview (Makino, 1991: 25).

The grading method is holistic, and the judgment of the interviewee’s proficiency is of course based the ACTFL Guidelines whose criteria consist of the four major pillars: (1) task and function, (2) settings and topics, (3) accuracy, and (4) discourse form (Makino, 1991: 18; NFLRC on-line). Novice speakers are characterized by the ability to say learned materials (words or phrases) in a minimal fashion. Intermediate speakers can create with the language and minimally sustain and conclude basic communicative survival tasks in simple ways (sentence-level). Advanced speakers are characterized by their ability to actively participate in the conversation and satisfy a wide variety of communicative tasks such as requirements of school and work situations (discourse level) (Makino, 1991: 20-21).

2.1.1. The Advantages of OPI

Some of the advantages of OPI include the following points. One is that testees are evaluated on the global criteria by a highly trained and qualified interviewer. The testees' speaking proficiency in Japanese is comparable with the same level of global proficiency in
other 31 languages since the ACTFL Guidelines are made to comply with and correspond to the guidelines of other languages (NFLRC on-line). Another advantage is the interactive format of the test. The interview format is an effective way to tailor the content of the conversation toward each testee’s real life and experiences, and the testee could perform in Japanese in a realistic and natural, yet carefully organized conversation while enjoying the real communication. This format also presents an opportunity for the testee to perform linguistic functions that “a straightforward question-answer interview would not permit” (NFLRC on-line).

2.1.2. The Disadvantages of OPI

While OPI is nationally acknowledged as a good tool for measuring proficiency, there are also some drawbacks in the format of the test and the ACTFL Guidelines in which its roots are found.

1) An interview format does not provide situations in which a testee can perform “survival” skills. Here the term “survival skills” means a learner’s skills in a target language are enough to perform essential tasks in order to minimally sustain one’s everyday routines in life, school, or at work. There are a limited number of questions that one would normally ask when meeting a person for the first time. Topics of conversation are therefore restricted to a set of highly predictable ones, such as weather expressions, date, and names of family members, which are some of the so-called "ten desperate questions" (Omaggio, 1985: 16). Obviously, most of the ordinary topics cannot be issues of “survival” in settings where people meet for the first time. OPI uses task cards to do a role-play to compensate for this, but the cards are usually used only once or twice, limiting the range of situations and tasks. This therefore limits the resources for the interviewer to judge if the testee has sufficient communicative skills with which to overcome most simple
survival settings. In addition, the use of task cards wouldn't be normal in the natural conversation that OPI attempts. This limit is inconsistent with OPI's fundamental assumptions because a major criterion for Intermediate speaking proficiency in ACTFL Guidelines is to be able to manage most ordinary "survival" settings. Makino (1991: 25), one of the committee members who developed the ACTFL Guidelines, acknowledges that OPI "should focus more on the role-play, which outrightly deals with communicative competence." He suggests that "OPI should have two modes of testing: the social conversation (shakō-kaiwa) mode, and the function/task mode."

(2) OPI does not test reading or writing. The objective of OPI is primarily to assess the speaking proficiency of the students, and the evaluation is based on characteristics defined in the speaking section of the ACTFL Guidelines. Listening skills do not receive evaluation or feedback, even though it is a major component in the interview and is required in order to speak appropriately.

(3) OPI does not test the speaking skills exclusively. Since OPI is conducted in a conversation, it inevitably includes both speaking and listening. However, the evaluation is made on speaking modality only, even though ACTFL Guidelines does have specified criteria for each of the four skills. Of course most actual interactions with people would involve two or more skills, and OPI is a way to conduct a test in such a realistic manner. However, if OPI is designed to test only the speaking skills and yet takes place in the interview format as described above, the listening component would interfere with the judgment of speaking skills. OPI is thus a multi-skill test intended to evaluate a single skill.

(4) Some portions of the ACTFL Guidelines criteria are also disputable. One of the areas of controversy is that the guideline establishes the four criteria--task/function, settings/topics, accuracy, and discourse-- and matches each level in one criterion with the
same level of the others with a one-to-one correspondence. For example, according to the ACTFL Guidelines, students with a low ability in accuracy would therefore be considered to be equally low in all of the other three areas: function/task, settings/content, and discourse form. This is because the guidelines were created to establish a "global" evaluation criteria. However, this rather simplified equation is questionable when considering the report by Higgs and Clifford (1982), cited in Omaggio (1985: 20, 23-24), on the relative contribution model. They hypothesize and are verified by actual OPI interviewers that "the relative contribution of the factors such as vocabulary, grammar, pronunciation, fluency, and sociolinguistic competence was not constant across proficiency ranges." Speech samples from OPI interviews are undoubtedly complicated, and errors that a testee makes involve various components or subsets of communicative competence. To lump all the components together with one criterion--such as Novice-High--is good in generating a general or global feedback but not very useful in diagnosing the skills of the students, or programs which produce students of such skills. Trade-offs are inevitable, but questions about the practicality of such holistic evaluation procedures still remain.

There are a few more concerns regarding the ACTFL Guidelines, specifically its handling of discourse and socio-cultural skills. Makino (1987) has explained that Novice students function at the word level, Intermediate students function at the sentence level, and Advanced students at the discourse level, and that some fundamental factors as hesitation noises and aizuchi show up only at the Advanced level according to the ACTFL Guidelines, which created some disputes in the field. Proficiency-oriented instruction should teach any necessary discourse skills at the earliest stage possible to create a natural flow in the target language and to teach students such naturalness as they perform successfully. It is not the case that discourse skill emerges only after mastering most of the basic grammar at the sentence level. As Szatrowski (1986a, 1986b, 1987) points out, such
fundamental discourse skills should be taught and learned at the very early stages of Japanese language study (cited in Kokuritsu Kokugo Kenkyūjo, 1989: 4).

There is some discrepancy about the socio-cultural elements of speech proficiency in the document as well. Cultural understanding in language, such as control over stylistic registers (non-polite, humble-polite, and honorific polite) appear only in descriptions of Advanced-Plus and Superior students, yet discussing one’s personal history, including terms for family members (onisan, for instance) which would require the use of an appropriate stylistic register, appear as early as the Intermediate-Mid level in the ACTFL Guidelines. It seems that the ACTFL Guidelines, because of its desire to cooperate with guidelines of other languages, has not yet concretely digested the difficulty of mastering the stylistic register system and other items of cultural significance in Japanese which may not be found in other languages.

(5) OPI is expensive to administer, and it is not suitable for an evaluation of an entire program. OPI involves individual interviews by a licensed OPI interviewer, and the interviewer must physically be at the site of testing. One OPI test costs $115, and if a program has 50 students, it takes $5750 and approximately 20-25 hours to test all of the students, not to mention the transportation and sustenance cost for the evaluator.

(6) OPI cannot give a detailed differentiation between the levels of proficiency at the beginning level. This last criticism is the most significant of all. Even after the ACTFL adopted the Government Scale by the Foreign Service Institute (0 to 5 scale) and provided a finer distinction as Academic Scales--dividing 0 (no or memorized proficiency) into Novice-Low, Novice, and Novice-High, 1 (elementary proficiency) to Intermediate-Low, Intermediate, Intermediate-High, and 2 (limited working proficiency) to Advanced and Advanced-Plus (Omaggio, 1985: 10-15; Takamizawa, 1987: 63-65, 74; National Foreign Language Center, 1992: 23)--this does not fully distinguish between the levels of students
of Japanese at the lower end of proficiency. "Students at the end of one year of college study are just as likely to attain a rating of Novice-High on ACTFL OPI as students at the end of their second year. The scale does not differentiate beginning students of Japanese sufficiently to serve as a useful assessment tool at this level" (National Foreign Language Center, 1992: 23). Recalling that most of the students of Japanese are exactly at this beginning level, the practicality and usefulness of OPI are questionable for most institutions.

The ACTFL Guidelines and OPI are both very influential to the field and their contribution cannot be underestimated. However, it is evident that "there are several points to be improved [...] and the original plan should be revised and strengthened" (Makino, 1991: 31).

2.2. Simulated Oral Proficiency Interview (SOPI)

SOPI, as its name indicates, is a simulated version of OPI, which is also based on the ACTFL Guidelines. The Center for Applied Linguistics (CAL) has developed the Japanese SOPI consisting of a Japanese Speaking Test (JST) and a Preliminary Japanese Speaking Test (Pre-JST) to "evaluate the level of oral proficiency in Japanese attained by American and other English-speaking learners of Japanese" and to “emulate the direct face-to-face OPI used by U.S. government agencies and ACTFL in a semi-direct format” (Center for Applied Linguistics, 1992a: 2). Both tests are approximately 45 minutes long, use cassette tapes and a booklet as a medium. Students are to listen to and follow the directions provided by the master tape and record their speech samples on the other tape. The booklet provides both directions and visual cues (pictures, etc.) to elicit their speech. The basic structure of the OPI--warm-up, level-check, probe, and wind-down--is present in SOPI as
well. The recorded tape is returned to the test provider and feedback will be sent to the institution (Center for Applied Linguistics, 1992a: 6-11).

What makes it possible for SOPI to take on this very handy format is the concept of simulation, the closest form of replication of real communication within given resources and a given environment. It is true that SOPI is not a real communication—what the students actually do is to understand the tasks given in English and perform them in Japanese by talking back to the tape. Nevertheless it is reasonable to utilize the idea and assess the students' communicative competence by the speech samples obtained by simulation. Without the idea of simulation, the only truly realistic test of students' survival skills would be having an evaluator follow a testee and record speech samples as the testee displays his or her actual survival skills in Japan. Cassette tapes and booklets are not as good as multimedia computers to handle simulated communication, yet they are more affordable, useful, and also allow for more standardized evaluation. SOPI offers a reasonably adequate medium for its purpose. The concept of simulation, which is also used in J-SKIT, is further discussed in Chapter 3.

JST has five subsections: Personal Conversation (warm-up), Detailed Description, Picture Sequence, Topical Discourse, and Situational Discourse. The Pre-JST has six subsections: Personal Conversation, Introduce Yourself, Prompted Discourse, Situational Discourse, Unprompted Discourse, and Requesting Information.

2.2.1. The Advantages of SOPI

Since directions and questions are prerecorded in English, the test administrator need not understand Japanese (NFLRC on-line). All the test supervisor needs to do is to receive the test materials, follow the administration procedures given in the instruction manual, proctor the actual test, and send all the materials back to the test provider. Not only can
SOPI be administered on an individual basis, but it is also administerable to groups if the test site has access to a language laboratory (Center for Applied Linguistics, 1992a: 6-7). This format greatly increases the accessibility of the test for more institutions and decreases the burden on the institution in terms of time and cost to administer the test. SOPI is also very useful as a standardized proficiency test because all students can take the test and the difficulty level among the different versions of the test is roughly the same.

SOPI also contains more varied tasks and situations than OPI, and learners of Japanese are tested with the same kind of topics and tasks, unlike the individually tailored OPI. In terms of comparability and usefulness as a placement test, it seems that SOPI is the most effective of all the currently available tests in the United States to measure proficiency in speaking. In JST, the format of actual test tasks is comprised of a single skill except for the ungraded warming-up part in which the students have to rely solely on listening comprehension to answer the questions. The use of English seems to be carefully handled by using performative verbs (persuade, describe, etc., in cue sentences) to minimize the risk of students translating literally from English.

The grading is holistic, in accordance with ACTFL Guidelines and done by trained raters. The interrater reliability of SOPI is very high in other languages (.88 and up) and it is "reasonable to assume that CAL's Japanese tests would have similar high interrater reliability" (Center for Applied Linguistics, 1992a: 24). In addition, although no direct correlation study has yet been conducted about the Japanese version of SOPI, it may be reasonable to assume that a high correlation exists between SOPI and OPI (Center for Applied Linguistics, 1992a: 26).
2.2.2. The Disadvantages of JST

Most of the criticisms on OPI mentioned in the previous section applies to SOPI as well. There are some additional concerns with SOPI, which can be summarized as follows. JST and Pre-JST are reviewed separately.

(1) JST is effective to test Intermediate-High level skills and above, but is not effective for students at beginning levels. This is a major drawback of both JST and OPI. The practicality and usefulness of SOPI as a placement test is therefore questionable. Almost all of the tasks assume long, paragraph-length monologues which require a high level of discourse skills. This is a concern because narration is a skill which would be successfully performed only by students of Intermediate-High and above proficiency. “This limitation poses a problem for its use in secondary schools, since many secondary students are at Novice level of [oral] proficiency” (Stansfield, 1994: 61).

(2) Most of the given tasks in JST require a high level of imagination and creativity, which are extra-linguistic skills. There are some sample tasks in the Examinee Handbook that may be very challenging to some people just because they have no ideas to talk about. “Explain the steps to follow in finding and buying a bicycle” (Center for Applied Linguistics, 1992b: 4), for example, is a good task to make testees speak about their own ideas, but it also has the potential risk of failure if a student cannot think of anything to say to perform the given task. A few other tasks may be something a testee has not thought about, or has no knowledge about.

(3) Collected sample speech may not represent real situations that students have experienced. In relation to the concern above, there is no way to tell whether testees are honestly talking about themselves or wisely avoiding certain structures or vocabulary. Again, there is too much freedom of creativity in SOPI.
(4) Some of the situations involve an encounter with Japanese in the United States, possibly within the cultural context of the United States, and a few tasks are not culturally plausible in a Japanese context. If the Advanced-Plus and Superior students are defined as easily understood by the native speakers of Japanese who are not accustomed to interacting with foreigners, JST's setting in the US is incongruent with this definition. Also, some cultural details that are assumed to be attained by students of very high proficiency are not tested in JST because of such settings.

2.2.3. The Disadvantages of Pre-JST

In order to test students with proficiencies lower than Intermediate-High, CAL has also developed Pre-JST, which "consists only of tasks at Novice and Intermediate levels" (Stansfield, 1994: 61). Pre-JST has the following drawbacks:

(1) Subsections 1 and 5 involve listening comprehension, and this is not a suitable format for a single-skill test. To emulate the real conversation, Pre-JST contains two parts that require students to both listen and speak in Japanese. It is thus a multi-skill test intended to evaluate a single skill, just like OPI.

(2) A few questions in Pre-JST seem to be tailored towards students in a formal education and not for learners who come from outside of a formal institutional education. For example, Sample questions in the Examinee Handbook, in which a testee is asked to talk about courses he/she took last quarter, his/her favorite course, and courses he/she plans to take next quarter, etc., are not very suitable for learners of Japanese who are not formally taking courses at school. In addition, such questions would definitely involve some high-level vocabulary such as names of courses like anthropology, algebra, art history and so forth.
To summarize, both JST and Pre-JST are very effective, validated speaking proficiency tests in Japanese and are easily accessible to many learners of Japanese. Except for a few problems in the types of questions and format, SOPI test materials comprise the multi-institutional test to measure learners’ proficiency in speaking, in terms of cost, accessibility, ease of administration, and validity.

2.3. SAT II Subject Test in Japanese with Listening (SAT-Japanese)

SAT-Japanese, formerly known as the Achievement Test, is administered by the College Board and Educational Testing Service (ETS). SAT-Japanese is one of the 11 Foreign Languages Subject Tests, an “one-hour test that consists of 80-85 multiple-choice questions and is written with secondary school curricula in mind” (The College Board, 1994: 79). It has subsections of Listening Comprehension, Usage, and Reading Comprehension “to measure the ability of students to engage in communication in Japanese in a culturally appropriate way” (College Board, 1994: 79). Scoring is adjusted statistically through a recentering process with a 200-to-800 scoring scale, and provides comparability to the students who took the Subject Test previously (The College Board, 1994: 5).

2.3.1. The Advantages of SAT-Japanese

The goal statement of SAT-Japanese clearly indicates the proficiency-oriented test assumptions, emphasizing "the importance of experiencing language as it is used by native speakers in a variety of settings. [The] students grasp emotionally as well as intellectually that the foreign language and the cultures in which it is spoken form a constellation of meanings that is as rich and varied as [the students'] own" (The College Board, 1996: xiv). Furthermore, SAT-Japanese is currently the only multi-institutional standardized test
available in the US that measures receptive skills and grammar. In addition, the test does not require an expert in Japanese at the test site, similar to JST and Pre-JST.

SAT-Japanese has other very interesting strengths. The test seems to respect the situational and cultural elements of communication. One sees this in the objective statement and also in the following statement that "the questions represent situations one might readily encounter and reflect realistic and commonplace communication" (College Board, 1994: 79). All the subsections are carefully constructed to avoid mixtures of modality to ensure the single-skill testing. The Listening Comprehension section contains realistic dialogues with a context header for each dialogue. The Usage section represents simplex sentences in three systems of representations (Japanese orthography with hurigana, Hepburn Romanaization, and modified Kunreishiki). The Reading section displays various text types with a context header, and contains various writing styles (horizontal, vertical, handwriting, etc.).

2.3.2. The Disadvantages of SAT-Japanese

SAT-Japanese is one of the most highly developed standardized proficiency tests in the U.S., but it has the following points to be considered as drawbacks.

(1) SAT-Japanese does not test speaking and writing. It only tests receptive skills (listening and reading) plus knowledge of grammar in Japanese. Such a format does not enable a testee to know how well one can perform in Japanese using speaking and writing skills. One suspects that the reason behind this format is suspected to be the complex and difficult nature of grading productive skills.

(2) The Usage section, unlike the two other sections, does not provide contexts, and a few of the sample questions are therefore ambiguous. This seems to be a major

    Pan to gohan to ---- ga ii desu ka?
(a) dore
(b) nani
(c) dochira
(d) dono

The answer is (c), according to the Official Guide. The assumed "typical" context would be somewhere at home, restaurants, or other food-providing places and someone is inquiring as to which one (of the two) the addressee prefers--rice or bread. However, (b) can be more appropriate in another rare but plausible context, where people are making a purchase at a department store and one is asking what would be good to buy, in addition to bread and a package of cooked rice. Since the question does not contain the context, there seems to be no way to know which context is posited in the question. The answer, therefore, remains unclear.

The usefulness and practicality of such questions is also questionable. It is extremely difficult for a learner of Japanese to think about all the plausible situations and contexts to eliminate all the options which should not be grammatically correct in any of the contexts. Such a task is something one would do only in the testing environment and not in an ordinary situation. The Usage section should include a header to explain the situation or context, as the other two sections do, to make the test complete.

(3) SAT-Japanese is only for high school students and is administered once a year at participating institutions only. This limitation of accessibility is also a significant criticism. Even though SAT-Japanese is administered by ETS and has a potential of development both in the quality of the test questions and in the creation of nationwide database to ensure
validity and reliability, the present condition of the test hinders it from reaching a wider population.

2.4. Conclusion

As stated earlier, none of these well established, currently available test materials provides a proficiency test in the four skills, provides detailed feedback or is available at a reasonable price for any institution at its convenience. With the advantages and disadvantages of these tests in mind, the J-SKIT project aims to produce a more comprehensive proficiency test in Japanese that measures how well a learner can use the language in a culturally appropriate manner in a variety of settings.
CHAPTER 3

J-SKIT: BASIC ASSUMPTIONS AND PROJECT PRINCIPLES

3.1. Introduction

J-SKIT has been developed to serve as a standardized proficiency test to evaluate students and programs and also to strengthen the emerging consensus in the diversified field. In order for J-SKIT to be useful in a multi-institutional context, it is important to have a clear, global description of the assumptions and principles of proficiency-oriented instruction in Japanese language pedagogy, so that the test will measure the implementation of such assumptions and principles in a particular program. This chapter describes basic assumptions about language and the project principles of J-SKIT, providing a theoretical framework used for the current version and future development of the project. Having the Framework and the ACTFL Guidelines as two major grounding publications, J-SKIT also incorporates theories and ideas from recent research in Japanese language pedagogy. While the assumptions and principles are not tied to any teaching methodology or textbook, they are based on recent research into the state-of-the-art teaching philosophy.

3.2. J-SKIT: Basic Assumptions

3.2.1. Assumption I: Language is Used as a Tool for Communication

Language is a set of arbitrary codes human beings share in a community in a given culture, and these codes are used as a tool for understanding the world and conveying
messages among one another. Language demonstrates how an individual has acquired the ability to recognize experience, to categorize items and experiences by naming features, and to synthesize these in culturally specific ways. Language is thus an embodiment of perspective conditioned by culture: a complex system to speculate, to organize ideas, and to seek knowledge as well as communicate them with one another. The J-SKIT project mainly focuses on the communicative aspects of language use.

As Neustupny (1988: 21) points out, the act of communication is a complicated process, based on substantive behaviors (*jisshitsu kōdō*, or non-communicative socio-cultural activities) as a driving input. This input as a purpose of communication promotes communication, and it is communicated as a message in specific settings or context, through a variety of principles, maxims, strategies, or rules” (Neustupny, 1988: 21, translation mine). The following is the modified model proposed by Neustupny (1982: 53) for performance:
"In conclusion, Japanese language education cannot exist only as linguistic education, nor just as communication education. It must include all the aspects of linguistic, communication, and interaction education" (Neustupny, 1988: 21, translation mine).

Teaching how to communicate with the target-native people should therefore incorporate more information than just a grammatical explanation in order to develop a skill of communicative competence. It is also pointed out by Savignon (1972: 8, cited in Omaggio, 1986:4) that "communicative competence may be defined as the ability to function in a truly communicative setting--that is, in a dynamic exchange in which linguistic competence must adopt itself to the total informational input, both linguistic and paralinguistic, of one or more interlocutors." "Learning to use a language requires learning a pattern of behavior. [...] The ability to use language to fulfill communicative needs cannot be achieved through the rote memorization of isolated utterances or facts" (National Foreign Language Center, 1993: 11).
3.2.2. Assumption II: Language Use in Communication Has a Purpose

When we humans use a language to communicate with others, it usually emerges from a motivation. Why did we decide to use language? The underlying presence of a motive is particularly visible in the productive use of a language--speaking and writing. This motivation or purpose may or may not be consciously recognized. Finocchiaro (1983) uses the term communicative purpose and explains the theoretical basis of the functional-notional approach as follows:

A Functional-notional approach to language learning places major emphasis on the communicative purpose(s) of a speech act. It focuses on what people want to do or what they want to accomplish through speech. Do they want to introduce people to each other? Do they want to invite someone to their home? [...] The above are simple examples of the functions of language which all human beings wish to express at one time or other (Finocchiaro, 1983: 13).

In other words, functions of the speech and other modalities emerge from needs and these needs which serve as purpose of communication, are communicated in a specific setting or context (Neustupny, 1988: 21).

For this reason, there should also be concern in developing language instruction curricula with the selection of types of functions which are often associated with types of speech acts. A list of communicative purposes in the curriculum should center on fundamental and universal needs of human beings when people interact with members of society. This point is further discussed in Assumption VII.
3.2.3. Assumption III: Communicative Competence can be Evaluated Through Task Performance

Communicative competence, as stated in Assumption I, is not static or fixed, and is closely related to the idea of task performance. Although the phrase "task performance" is often used to discuss classroom activities, the phrase sometimes lacks clear definition. This section discusses the two words, "performance" and "task", separately.

"Performance," in a broader and more trans-disciplinary sense, can be defined as a skilled behavior. Performance requires practice, as in fields such as ballet, theater, music, but "any sort of performance requires a leap on the part of the performer. Grasping material intellectually, or in a restricted context, does not guarantee that one will be able to perform it when the occasion to do so presents itself" (O’Connor, 1996: 174). The instantaneous and improvisational nature of performance, while extracting vital information out of a specific context, requires repeated practice to gain "automaticity in performance [, which] permits a performer to accomplish the leap between conception and actualization of material for performance, by granting them the freedom to identify fully with the situational context of the moment" (O’Connor, 1996: 186). To develop skills in performance, one needs to "situate the body into the thick of the matter, to exercise the capacity to conceive of ideas through the body and its involvement in a specific spatial and/or social context" (O’Connor, 1996: 177).

In the field of language pedagogy, performance involves skilled behavior using language. It is important to have a theoretical distinction between communicative competence and performance in a more specific sense, as Savignon (1972: 9), cited in Omaggio (1986: 5), proposes: "competence is what one knows. Performance is what one does. Only performance is observable, however, and it is only through performance that competence can be developed, maintained, and evaluated." Further, Omaggio (1986: 7)
summarizes the work by Campbell and Wales (1970), Hymes (1972), and Munby (1978) as an “integrative view of communicative competence with four major components: (1) grammatical competence, (2) sociolinguistic competence, (3) discourse competence, and (4) strategic competence.” Performance using language in real communication is a rather complicated act, for one needs to utilize the competence or knowledge in these four areas as bases of behavior, and to integrate them into performance on the spot.

The notion of “task” is defined by Richards, Platt and Webber (1986: 289) in the Longman Dictionary of Applied Linguistics, cited in Okazaki and Okazaki (1990: 77), as "an activity or action which is carried out as the result of processing or understanding language (i.e., as a response). [...] Tasks may or may not involve the production of the language. [...] The use of a variety of different kinds of tasks in language teaching is said to make language teaching more communicative, since it provides a purpose for classroom activity which goes beyond the practice of language for its own sake.” This definition includes many activities, and they are subcategorized into five types of activities by Okazaki and Okazaki (1990: 87): (1) activities that involve an information gap among students (the question-and-answer format, the dialogue format which involves variations of existing dialogues, and the format which utilizes pictures), (2) the role play format, (3) the debate format, (4) the "You're my wife" format (searching for the best option from various information), and (5) project work format (translation mine). In order for tasks to be effective, activities should be carefully examined so that they “involve everyday activities to which a learner can easily relate oneself. In this way, learning is facilitated by the active cognitive mechanisms of human beings—using analogy, inference, verification, and modification of learning process through trial and error” (Okazaki and Okazaki, 1990: 132, translation mine).
It is assumed that the use of the language, and communicative competence enabling a student to engage in communication can be observed through task performance. All of the performance data involves the four aspects of communicative competence (grammatical, sociolinguistic, discourse, and strategic competence), and evaluation procedures should consider those aspects of performance.

3.2.4. Assumption IV: Language Use Observes Cultural Conventions and Expectations

A culture is a complex set of standards, assumptions, conventions, value systems, and expectations shared among members of a specific community that influences all the areas of their activities. Culture embraces all lives within the community and provides a certain unity, wholeness, and identity to the group.

In the field of language pedagogy, the notion of acquired culture is of particular interest, as opposed to learned (aesthetic, information, and skill) culture (Jorden, 1992: 114). "Acquired culture" is defined as "a kind of culture, which native speakers acquire subconsciously as they are socialized within their native society. This deep-level mindset [...] is the driving force that determines attitudes toward time and space, styles of analysis and interaction, and identification of self" (Jorden, 1992: 114). Language, communicative behavior, and interactional practices are all parts of this abstract notion of culture, which may be synonymous with "the thick of the matter" of O'Connor (1996: 177). "It is essential to teach the target language IN culture [...] in order for a learner] to become an informed foreigner who can function in Japanese society in a way that does not make Japanese feel uncomfortable" (National Foreign Language Center, 1993: 9, 16). To achieve this objective, one must be familiar with the ways that the target language is used by the target-native people in the target culture, and follow their expectations and conventions when using the language to interact successfully and appropriately with
target-native people. This includes the appropriate use of speech style, according to the social and personal relationship of the interlocutors and genre in which communication takes place.

"Language and culture are inseparably bound; therefore, complete comprehension during any type of intercultural communication depends upon the participants' awareness of the social and cultural significance of the word and expressions employed. Language is used to convey meaning, but meaning is determined by the culture" (Chastain, 1988: 298). This concept of cultural appropriateness is of crucial importance, especially for students of Japanese, for many native Japanese people are not accustomed to interacting with foreigners.

It is also important to note that there are subgroups which share subcultures within a community of larger numbers. Different behaviors, including language use, can be observed in different subgroups, according to the differences in region, generation, sex, and social roles shared by the group. The J-SKIT project assumes the culture of Japanese educated adults who are not accustomed to interacting with foreigners. This assumption would include a large number of native Japanese people and would be the group that students of Japanese would be most likely to encounter. It is also assumed that this culture would function as the dominant culture when a foreigner interacts with native Japanese adults; thus the interculture between Japanese natives and a foreigner is very close to the assumed target culture¹.

3.2.5. Assumption V: Authentic Language Use Must be Promoted and Encouraged

In connection with Assumption IV above, language use must be appropriate within the context of the target culture, and authentic language use must be promoted and encouraged

¹ For the notion of interculture, see Walker (1989).
from the very beginning of instruction in order to achieve the long-term goal of being able to use the language to function in the target culture.

The term "authentic materials" is sometimes used without a clear definition, but it often refers to the language materials "originally produced by and intended for native speakers of the target language" (Frye and Garza, 1992; cited in Noda, 1994: 159). This restricted notion, however, is not very practical for the Japanese taught in the U.S., because of the limited resources available in the U.S. due to the Japanese language's status as a less commonly taught language. It is also a problematic notion because it would exclude all of the linguistic samples produced by non-natives, including students and non-native teachers of Japanese. It is important that students should be able to function in the target culture in ways such that the target-natives belonging to the assumed culture would feel comfortable in communicating with them. From a pedagogical perspective, the J-SKIT project takes a different approach toward authenticity of the language, following Noda (1994) and the National Foreign Language Center (1992).

In the J-SKIT project, authenticity in language refers to quality in language samples: they may or may not be produced with specific pedagogical purposes for the learners of the language, yet they should not be distinguishable from speech or writing that target-native speakers would actually use in a given context of their ordinary social interactions with other target-natives (National Foreign Language Center, 1993; Noda, 1994). Authentic language samples should sound natural, appropriate, and acceptable to the target-natives in a given context, and should not be something an educated target-native would find linguistically peculiar or abnormal. The use of gaijingo, a special dialect prepared for and used by foreigners only, should therefore be avoided (Jorden, 1989: 14-15). Authentic language samples must meet the expectations and conventions of the target culture.
There should be concern for authenticity of tasks and activities in the classroom, as well. "Activities can be authentic whether they utilize artifacts [, recorded, authentic materials in the narrow sense of Frye and Garza (1992) mentioned earlier,] or newly created stories, or scripted materials, so long as they establish plausible stories" (Noda, 1994: 163). This notion includes not only activities among the target natives but also ones that a learner of a language would be likely to encounter frequently when interacting with target natives. The latter activities are called "contact settings" and "contact topics" by Neustupny (1982). Language use in such contact settings, however, should observe the expectations and conventions of the culture in order to be authentic. It is noted that "the authenticity does not lie in the artifact itself, but in the story it creates" (Noda, 1994: 168).

3.2.6. Assumption VI: Communication Takes Place in Context

Context is an important factor in communication that must be taken into consideration while designing language instruction and testing projects. The term “context” refers to "ongoing discourse as well as some particular circumstances or situation" (Omaggio, 1986: 91). This notion minimally includes the time and place of interaction; age, sex, social roles and status plus personal relationship of the interlocutors; the topic, purpose, and modality of communication. "The idea that language learning should be contextualized is certainly not new in language teaching, at least from a theoretical point of view" (Omaggio, 1986: 91).

The cultural assumption described earlier serves as an underlying context, and realization of each specific context in J-SKIT is considered to be an interaction between a learner of Japanese and a group of people sharing the assumed culture mentioned in the previous section. In order to determine whether certain behaviors or language samples are authentic, one must consider the specific context in which they appear. "Communicative
competence is context-specific, in that communication always takes place in a particular context or situation. The communicatively competent language user will know how to make appropriate choices in register and style to fit the particular situation in which communication occurs" (Savignon, 1972: 9, cited in Omaggio, 1986: 5). The performance data obtained in J-SKIT is evaluated in terms of authenticity within the specific provided context.

3.2.7. Assumption VII: Language Use Must Consider Function, Structure, Vocabulary, and Context

Function, structure, vocabulary, and context of language are all equally important in language processing, both in production and comprehension, and therefore these elements deserve equal attention. Here the term “function” refers to a variety of speech acts with which the term is conventionally associated, particularly in sentence production. "In uttering any sentence, a speaker necessarily performs a literal act which is conventionally associated with the type of sentence uttered" (Geis, 1996: 9). There are many ways to label the types of utterance, and J-SKIT adopts the categorization of Van Ek (1980), cited in Finocchiaro (1983: 63-64), which is provided in Chapter 4. “Structure” refers to grammaticality or accuracy in syntax, and it includes constructions of both sentence and discourse level. “Context” includes both situations and topics of communication.

These four domains serve as the major pillars for the J-SKIT project, although simple one-to-one mapping between one item in one domain and the other in another domain have been avoided since all four elements of language are inextricably related. To associate one structural pattern of a language with one speech act, for example, is impractical and meaningless, if not impossible. As Geis proposes, "The thesis that there exists or should
exist a mapping from individual utterances to primary speech acts must be abandoned" (Geis, 1996: 10).

The relationship among these four pillars is a complex and vaguely known one, and further research in cognitive science and psycholinguistics could clarify the relationship. Nonetheless, it would be helpful to posit a tentative model for sentence production (speaking and writing) here. The following is a modification of theory proposed by Finocchiaro (1983: 28). The model serves to illustrate a possible way of capturing the relationship among function, structure, vocabulary, and context.

![Tentative Model for Sentence Production](image)

Figure 2. A Tentative Model for Sentence Production

"Different communicative purposes, situations, and topics lead us to adopt our messages so that they will be most clearly understood. [...] The language that we actually produce changes when these elements change, because we have learned to adjust our language use to be appropriate for the conditions in which we use it" (Finocchiaro, 1983: 13).

It is necessary to make some assumptions regarding each of the four domains in order to create organized and effective testing materials. Content specifications for function, structure, vocabulary, and context are provided in Chapter 4.
3.3. J-SKIT: Project Principles

With these seven assumptions about language use and instruction as a theoretical base, J-SKIT observes the following project principles to develop the specific test format of the present version. This section describes the basic project principles in relation to the project assumptions. A more detailed description on the actual appearance of the test is provided in Chapter 5.

3.3.1. Principle I: J-SKIT is Targeted at Students in the United States

As stated in Chapter 1, J-SKIT is designed to function as a comprehensive test of proficiency that reflects the basic emerging consensus on proficiency-oriented programs in the field. Such a test is necessary in order to enhance the emerging consensus on such programs, to provide a valid means of placement, and to pave the paths of articulation between high schools and colleges in the United States. The administrative procedures and test format are created to reach both high school and college students learning Japanese in the United States. Both the format and content specifications of the test should reflect the present situation of Japanese language pedagogy in U.S. education. Finally, the test administration and the actual test involves use of English, which, it is assumed, both the test administrators and test takers can read and understand without trouble.

3.3.2. Principle II: J-SKIT is Designed to Measure Performance in Simulated Contexts

J-SKIT contains many tasks for test takers to perform (Assumption III), and the task performance is evaluated with certain grading criteria specified in Chapter 6. In order to measure test takers' performance, J-SKIT contains many types of contexts that provide stages for performance.
As Neustupny points out, "Japanese language education needs to go beyond the classroom in a traditional sense as much as possible, since classrooms do not contain socio-cultural purposes, i.e. input, for communication" (Neustupny, 1988: 21). He is right to further mention that "in role plays and in the so-called communicative approach, most teachers artificially create settings [in the classroom]. Such activities do not provide opportunities to practice how to handle real input in communicative activities" (Neustupny, 1988: 22, translation and emphasis mine). Even so, Japanese language instruction, as a less commonly taught language in the United States, cannot provide ample opportunities for the students to engage in real communicative activities outside of the classroom because of the lack of resources. In many regional institutions, it is often the case that the teacher is the only one who speaks Japanese in that community. In addition, standardized test administration requires students to be in a certain place free from distractions for test security reasons.

While acknowledging that such a test environment cannot provide real communicative settings and input, J-SKIT adopts the notion of simulation, the closest means of replicating real communication within given resources and a given environment. It is assumed that utilizing the notion of simulation to assess the students' communicative competence by the speech samples obtained through simulated task performance is practical and reasonable. As mentioned in Chapter 1, the only truly realistic test of students' communicative skills is one in which an evaluator follows a testee and records every speech sample as the testee tries to really communicate with target-native people without the idea of simulation. In addition, simulation makes it possible to incorporate many situations and topics which the interview format cannot fully incorporate.

Valette (1994: 15-16) presents the five-step performance-based model of proficiency: (1) guided observation, (2) guided analysis, (3) guided practice, (4) simulated
performance, and (5) performance. Making an analogy to a football game, she explains as follows:

In Step 4 [, simulated performance,] students have the opportunity to use the new words, phrases and patterns they have learned in simulated conversational exchanges and role play situations. [...] Much as one might try to have these exchanges resemble real-life conversations, [...] teachers and students both recognize that the context is artificial. This type of language practice is similar to football scrimmage practice. [...] Scrimmage gives the players the opportunity of running their plays in a game-like context where the focus is on performing effectively.

It is important to note that simulated context is more effective if the artificially created situations and communicative input are very close to authentic language use (Assumption V) so that the simulated performance would be readily applicable and the "leap of the performer" (O'Connor, 1996: 174) from simulated practice, whether it be a dress rehearsal or football scrimmage, to the real performance, act or a game, be kept to a minimum.

3.3.3. Principle III: J-SKIT Measures All of the Four Skills Separately (Single-skill Testing)

The format of J-SKIT is carefully designed to test proficiency in each of the four skills separately. This is not just to follow the field’s conventional distinction of the four skills, but to strengthen the consensus and understanding of proficiency in each skill as characterized in the ACTFL Guidelines. The single-skill format also diagnoses students' strengths and weaknesses more effectively. Even though we use multiple skills in real communication (conversation of course requires both listening and speaking skills), "unfortunately, when students err in these compound-skill tasks, the reason for their error is sometimes unclear. [...] Tests that measure one skill at a time are therefore necessary complements to multi-skill tests. Only single-skill tests pinpoint student deficiencies in a
useful manner" (National Foreign Language Center, 1992: 27). Following Principles I and II above, the directions and context for performance in J-SKIT are provided in English and in pictures to create a simulated context for performance, requiring only one skill in the target language.

While the single-skill test format is more suitable for the detailed diagnosis on proficiency, it is also important to acknowledge that it is a constrained condition for the test production. More control over the input for and nature of tasks (with descriptions in English) is inevitable in the single-skill format, and the test production has to be operated in a more restricted manner. At the same time, the tasks in J-SKIT should be realistic and authentic. The J-SKIT test production aims to satisfy both requirements.

3.3.4. Principle IV: J-SKIT is a Proficiency Test for Students of Lower Levels of Proficiency

J-SKIT is a proficiency test, as opposed to an achievement test. A “proficiency test,” by definition, "measures a student’s ability with no assumptions as to how the ability has been attained" (National Foreign Language Center, 1992: 10). "It is not possible to prepare for a proficiency test in the same way one might study for an achievement test, since the tasks presented and elements included are not limited to a specific body of material" (Omaggio, 1986: 9). Proficiency tests are based on the notion of outcome observable in a form of performance, i.e., what the testees can or cannot do in the target language (Assumption III). Following the ACTFL Guidelines, the nationally recognized performance proficiency standards, J-SKIT evaluates students' performance, "measuring against that of a native speaker" (National Foreign Language Center, 1992: 10).

It does not mean, however, that J-SKIT as a proficiency test is assumption free. In addition to the seven assumptions mentioned earlier, J-SKIT is designed under the
assumption that a majority of test takers are of lower proficiency. This is reasonable given that most of the students learning Japanese are of lower proficiency as noted in Chapter 1. The tasks and questions in J-SKIT are created and arranged to test the fundamental competence and performance of learners of Japanese. In addition, since J-SKIT is based on the four major pillars stated in Assumption VII (function, structure, vocabulary, and context), instead of those of ACTFL OPI (task and function, settings and topics, accuracy, and discourse form), J-SKIT establishes its own test objectives. A detailed content specification comes later in Chapter 4. It is sufficient to mention here that, in general, the J-SKIT project adopts the notion of generic Intermediate proficiency characterized by the ACTFL Guidelines and focuses on students' "survival skills" -- everyday, simple communication in contact situations. Following ACTFL Guidelines, the J-SKIT project describes intended tested skills as follows:
Speaking -- to be able to create with the language by combining and recombining learned elements; to ask and answer questions; to initiate, sustain, and bring to closure a wide variety of basic communicative tasks that would minimally satisfy the requirements of school and work situations; and to narrate and describe with paragraph-length connected discourse.

Listening -- to be able to understand the main ideas and factual details from exchanges and simply connected aural texts involving a variety of everyday, frequently encountered topics in forms of dialogues, monologues, short announcements, telephone messages, and so on.

Reading -- to be able to understand the main ideas (skimming) and factual details (scanning) from simply connected texts encountered daily, such as menus, memos, signs, simple articles and letters; to read both hiragana and katakana, and basic, frequently used kanji.

Writing -- to be able to minimally satisfy simple, practical writing needs; can write simple documents, such as short messages, memo, notes, or letters, using the writing systems (hiragana, katakana, and frequently used kanji) accurately and appropriately.

Table 2. Skills Tested in J-SKIT

J-SKIT intends to measure learners' ability to perform in Japanese in simulated settings to evaluate students' proficiency in each skill

3.3.5. Principle V: Tasks in J-SKIT Must be Realistic, Like Something Learners are Likely to Encounter in Daily Interactions in Japan

In connection with Principles II and IV above and also Assumptions IV, V, and VI, every simulated task and context presented in J-SKIT is designed to be realistic and authentic. J-SKIT focuses on "survival" settings such as ordering foods, finding accommodation, reading signs, and giving and asking directions, and also considers contact situations (Neustupny, 1988), since "the most important thing for learners of Japanese with limited proficiency is to know how 'foreigners' would behave in a post office, for example, and to know how they should face the challenges in communication when they encounter such a situation" (Neustupny, 1988: 22, translation mine). J-SKIT
incorporates situations that students would be most likely to encounter frequently in Japan, and successful interaction in such situations would be vital to accomplish fundamental substantial behavior through one’s survival skills in Japan.
CHAPTER 4

J-SKIT: CONTENT SPECIFICATIONS

4.1. Introduction: J-SKIT as Function-based, Structurally Organized, and Contextually Realized

As mentioned in Assumption VII in the previous chapter, J-SKIT takes function, structure, vocabulary, and context as its four major supporting pillars in its composition. The elements of each area are discussed in this chapter specifically in terms of their content. It is important to note, however, that each list in this chapter should not be considered to be closed or fixed: the content specification presented in this chapter is subject to change as more research and studies become available. In addition, it is necessary to update these content specifications constantly as the language changes through time.

All test materials in all of the sections of J-SKIT are created with the content specifications in mind and finalized in the form of a task. The arrangement of materials is determined after each task has been checked for its authenticity and testability. The arrangement and organization procedures consider the following defining elements of task specificity by Brecht and Walton (1995):
(1) Nature of the Task(s): What is the nature of the communication task or tasks? What is the function which is being performed?
(2) Communication Mode: Which modes does the task require? Interpersonal, interpretative, presentative, or combination?
(3) Skill Mode: Which skills are required? Speaking/listening, listening only, reading, writing, all of these skills, or combinations of several?
(4) Level of Competency: For the specified skill or skills, what is the level of competency required by the task for each?
(5) Task Frequency: How frequently needed is the skill? Only occasionally, a few times a year for trips, or every day?
(6) Task Criticality: How critical is the task?

Of all the specificities above, (1) is specified in the Function Specification of the J-SKIT; (2) is considered to be interpersonal in the Speaking and Writing sections and to be interpretative in the Listening and Reading sections; (3) is cleared with the single-skill test format (Principle III). (4), (5), and (6) are considered with the four J-SKIT pillars: function, structure, vocabulary, and context.

In general, the J-SKIT material organization operates with basic criteria: emphasizing the common over the rare, arranging the simple before the complex, and progressing from the concrete to the abstract (National Foreign Language Center, 1993: 32-33). For the judgment of common vs. rare, contextual information and the nature of task are regarded to be the primary factors on which the decision is based. As for complexity, structural information is considered as the primary source for decision. In terms of frequency of the task, both functional and contextual information are taken into consideration as primary deciding factors. In other words, fundamental organizational principles of J-SKIT can be summarized as function-based, structurally organized, and contextually realized.

The specific content specification is provided below. While the major textbooks used in the U.S. are considered and examined, the content specification is constructed so that J-SKIT would maintain a proficiency test which is not tied to any specific program, institution, textbook, or teaching methodology (Principle IV).
4.2. J-SKIT: Content Specifications

4.2.1. Function

The function of spoken language can be regarded as a speech act that the speaker wishes to accomplish by using language. Here speech acts refer to more basic and universal intentions of human beings carried with a speech, while functions are instantiations of the speech acts within the given culture and language. Thus a specific speech act as a purpose of communication can be achieved with words, phrases, sentences, or discourse in the form of function. Speech acts are not to be considered to be language-specific: even though the function, or the process of communication, varies significantly among different communities, languages and cultures, the inputs of communication (exemplified as types of speech acts) which are used to perform substantive and socio-cultural behaviors can be said to be universal (Neustupny, 1988: 11-19). In other words, it needs to be recognized that "actions like requesting, offering, promising, making assessments, and the like are social as opposed to linguistic actions. [...] Once it is recognized [...], the temptation to associate the performance of particular speech acts with the uttering of sentences having particular linguistic forms diminishes greatly" (Geis, 1995: xii).

There are many ways to label types of speech acts. The J-SKIT project adopts the classification by Van Ek (1980), cited in Finocchiaro (1983: 63-64). These categories "cover most of the communicative functions that would be included in a regular, basic course [...]. In harmony with the spiral approach advocated, the language beyond [the basic level] program may be more complex and sophisticated but the communicative purposes to be uttered or written [...] would be very similar to those expressed" (Finocchiaro, 1983: 63). Following is a list of categories (Van Ek, 1980, cited in Finocchiaro, 1983: 63-64):
1. Imparting and seeking factual information: identifying, reporting (including describing and narrating, correcting, and asking)

2. Expressing and finding out intellectual attitudes:
   Expressing agreement and disagreement
   Inquiring about agreement and disagreement
   Denying something, accepting an offer or invitation
   Declining an offer or invitation
   Inquiring whether offer or invitation is accepted or declined
   Offering to do something
   Stating or inquiring whether one remembers or has forgotten something or someone
   Expressing or inquiring whether something is considered possible or impossible
   Expressing or inquiring about capability or incapability
   Expressing or inquiring whether something is considered a logical conclusion (deduction)
   Expressing or inquiring how certain/uncertain one is (others are) of something
   Expressing or inquiring whether one is/is not obliged to do something
   Expressing or inquiring whether others are/are not obliged to do something
   Giving and seeking permission to do something
   Inquiring whether others have permission to do something
   Stating that permission is withheld

3. Expressing and finding out emotional attitudes:
   Expressing and inquiring about pleasure, liking
   Expressing and inquiring about displeasure, dislike
   Expressing and inquiring about surprise, hope, satisfaction, dissatisfaction
   Expressing or inquiring about disappointment, fear, or worry
   Expressing and inquiring about preferences
   Expressing gratitude and sympathy
   Expressing and inquiring about intention
   Expressing and inquiring about want and desire

4. Expressing and finding out moral attitudes:
   Apologizing
   Granting forgiveness
   Expressing and inquiring about approval and disapproval
   Expressing appreciation
   Expressing regret
   Expressing indifference

5. Getting things done (suasion):
   Suggesting a course of action (including the speaker)
   Requesting, inviting, or advising others to do something
   Warning others to take care or to refrain from doing something
   Instructing or directing others to do something

6. Socializing:
   Greeting people
   Meeting people
   Introducing people and being introduced
   Taking leave

Table 3 (continued)

Attracting attention
Beginning a meal
Propose a toast

In the production of the J-SKIT test materials, tasks in the Listening, Speaking, and Reading sections have included as many of the above categories as possible. The five major areas of categories have been given equal attention and distribution in the test. Tasks in Writing sections should be selected from these categories above, although more attention will be given to orthography, as specified later in this chapter.

4.2.2. Structure

All of the four skills tested in J-SKIT assume most of the basic sentential patterns and frequently occurring discourse devices. As is the case with functions, there are several ways to categorize elements in structure, and the manner of categorization and linguistic terminology vary among textbooks. J-SKIT works with the results of the research by Kawarazaki et al. (1992) on the sentence patterns and grammar presented in five major Japanese language textbooks popularly used in Japan and the United States\(^1\) plus Japanese: the Spoken Language (Jorden with Noda, 1987), which contains detailed structural information. The research by the Japan Foundation (1994) on structures in high school textbooks has also been incorporated.\(^2\) Most of the structures that appeared in both

---

\(^1\) The research is based on the following five textbooks: Nihongo Shoho by Kokusai Koryu Kikin, Shokyū Nihongo by Tokyo Gaikoku Daigaku Fuzoku Nihongo Gakko, Nihongo no Kiso by Kaigai Gizyutsusha Kenshu Kyokai, Bunka Shokyū I and II by Bunka Gaikoku Gakko Senmon Gakko, and An Introduction to Modern Japanese by Mizutani and Mizutani.

\(^2\) The research is based on the following twelve textbooks: Basic Japanese Level 1-2, Bunka Shokyū Nihongo I, Hello in Japanese I-II (Konnichiwa Nihongo), Hiroko-san no Tanoshii Nihongo, Japanese 1 and 2 (SERC), Japanese Book 1-3 (Alfonso), Japanese Made Possible Vol. 1&2, Japanese Now Vol. I-II
documents are covered in *A Dictionary of Basic Japanese Grammar* (Makino and Tsutsui, 1986). The following is a list of all of the structural patterns and grammar that appeared in three or more textbooks, according to Kawarazaki et al. (1992) and the Japan Foundation (1994).

Note: *N* refers to a noun or a nominal, IA-*i* to an adjective or adjectival, NA-*na* to a *na-*adjective or *na-*nominal, *V* to a verb or verballs. A structure with an asterisk indicates that the structure is covered in more than 3-5 out of 12 high school books, as reported by the Japan Foundation, 1994.

*\(N\) desu / de wa arimasen
*N* deshita / de wa arimasen deshita
*N* da / *N* datta (plain form)
*N* de
*N* deshô / de wa nai deshô
*IA* + *N* desu
*IA desu / A-ku nai desu / A-ku arimasen
*IA-katta / IA-ku nakatta
*IA* (plain form)
*NA* + *N* desu
*NA desu / de wa arimasen
*NA deshita / de wa arimasen deshita
*NA da / NA datta (plain form)
*NA deshô / de wa nai deshô
*IA-kute, IA desu* (connecting adjective sentences)
*NA de, IA desu* (connecting *na-*adjective sentences)
*N o *V-masu
*N* wa *V-masu / masen
*N* ga *V-masu / masen
*V-mashita / -masen deshita
IA-ku *V-masu / IA-ku naru / IA-ku suru
NA ni *V-masu / naru / suru
*V-te, V-te, V-masu* (sequence of actions)
V-te, V-masu (condition, method)
*V-te kudasai
*V-nai de kudasai
*V-te kara
V-ru mae ni

Table 4. J-SKIT Structure Specification (continued)

Table 4 (continued)

*V-te imasu (repeated action)
*V-te imasu (continuing result/stative)
*V-ru (plain form)
*V-ta (plain form)
V-ru N (modifying noun)
V-ta N (modifying noun)
V-te kimasu / ikimasu
*V-te okimasu (preparation)
*V-te mimasu (trial)
V-te simaimashita (perfect action)
V-te simaimashita (regretable action)
V-te / -de (reason, cause)
*V-nai de
V-zu ni
*V-masen ka (invitation)
*V-mashō (invitation)
*V-mashō ka (offer, proposal)
V-nasai
*V-(yo) (to omoimasu) (intention/volitional)
V-ru tsunori desu (intention)
*V-ta koto ga arimasu (experience)
V-ru koto ga arimasu
*V-ru koto ga dekimasu
*V-ru koto ni shimasu
V-ru tokoro desu
(phrase) hō ga (predicate)
N bakari / V-te bakari
--ka mo shiremasen
--ka dō ka
*V-(ra)remasu (potential)
V-ta mama
*V-nagara
*V-te mo it desu (permission)
*V-te wa ikemasen / dame desu (prohibition)
V-nakute mo it desu (acceptance)
V-nakereba narimasen (obligation, justly action)
-- ni V-(ra)remashita
-- ni -- o V-(ra)remashita
-- ni V-(ra)rete imasu (inanimate subject)
-- ni V-(sa)semasu
-- o V-(sa)semasu
-- ni V-(sa)sete kudasai
*yarimasu / agemasu
*moraimasu / itadakimasu
*V-te yarimasu / agemasu

(continued)
Table 4 (continued)

V-te moraimasu / itadakimasu
*V-te kuremasu / kudasaimasu
*V-tari V-tari simasu
*--to, --
*--tara, --
*--toki, --
--ba, --
--nara (--ga ii)
--rara, --
--temo / demo, --
o-V ni narimasu (honorific)
--o V-(ra)remasu (honorific)
o-V simasu (humble)
--de gozaimasu (polite)
*--sō desu (hearsay)
*--sō desu ("looks," aspect, condition)
--sō na /sō ni
--rashii desu
--yō desu (as if)
--yō desu (inference)
*--tai desu
--tagarimasu / --tagatte imasu
*-- ni N ga arimasu (existence)
N wa -- ni arimasu (place of existence)
N to N
-- to -- to de wa (comparison)
-- to iimasu (quotation)
*-- to omoimasu
-- to issho ni (companion)
-- to onaji desu (object of comparison)
-- to aimasu (object of action)
-- to iu N
-- to ka
-- de (location of action)
-- de (number of people for action)
-- de (extent): zentu de
-- de (range): kore de ii
--de (instrument): basu de iku
--de itiban --
--de (raw material): --de tsukuru
*--de (method): Nihongo de --to iu
*--de (connective): --wa --de, --wa --desu
*--ni (location of existence): --ni --ga aru

(continued)
---ni (time of action) --ji ni V-masu
---ni (ratio): (duration) ni (times)-kai
---ni (goal): --ni narimashita
*---ni (goal location): --ni iku
---ni / to (partner of action)
---ni (entering location): --ni noru
---ni (direction of action): --ni iu
---ni (recipient of action)
---ni (time of action) --ji ni
---ni (location of result/installment): --ni kaku
---ni (purpose of use) --ni tsukau
---ni (stopping location): --ni tomaru
*---o (object): --o kudasai
---o (area of action): --o toru
---o (departing location): --o deru
*---no (type/kind): Nihon-go no sensei
*---no (possession): --san no kasa
*---no (location): reizoko no naka
*---no (condition): cha-iro no uwagi
*---no (nominal): nagai no mo aru
*---no/ga (relative clause/sentence modifier): --no/ga katta hon
*---no (explanation/the extended predicate) V-ru/-ta no desu
---no o
---no ni
---no ga
---no wa
*---ga (agent/subject): --ga--
---ga (feelings): --ga suki / kirai desu, --ga hoshii desu, --ga shimasu
*---ga (affect): --ga jôzu / heta desu, --ga wakarimasu
*---ga (conjunction): dekimasu ga, --
*---e (direction): --e ikimasu
*---ka (interrogative): -- desu ka
---ka (unidentified: hon ka nôto
---ka (unidentified: nani o V-ru ka V-masen
---ka (unidentified: nani ka --o shite imasu
*---wa (topic/subject)
---kara (place of origin): Thai kara kimashita
---kara (time) nan-ji kara
---kara (order) ato kara V-masu
---kara (raw material): nani kara tsukurimasu ka
---made (time goal): nan-ji made
---made (place goal): doko made
---yori (comparison): --yori --no hô ga --

(continued)
Table 4 (continued)

--mo (addition)
--shi: --mo --shi, --mo --da
*--kara, -- (cause/reason)
*--no de, -- (cause/reason)
--no ni, --
--ya
goro
gurai
--shika V-masen
--ne
--yo
--wa
--de mo (potential capacity): kodomo de mo
--de mo (unidentified): kohi de mo

Creating the list above involved many difficulties due to the inconsistent research standards and procedures, and some of the asterisked structural patterns are of the inferred ones. To create a more complete list of structural specifications, the list above will be revised so that it would incorporate more specific and concrete grammatical notions using linguistic terminology consistently.

This list is of particular use as a guideline for the Structure section in J-SKIT. Most of the structures in the list appear throughout the test, and all the ones with asterisks will appear in J-SKIT. The asterisked structures constitute approximately 40% of the test.

4.2.3. Vocabulary

Creating Vocabulary Specification presents various problems for the Structure Specification, for "vocabulary is an almost infinite system, and it is a personal system as well" (Chastain, 1988: 397). In order for J-SKIT Vocabulary Specification to be suitable and fair to most students of Japanese, two basic principles have been established as guidelines for the selection of vocabulary items. One principle is that all of the vocabulary should represent basic, non-technical vocabulary of everyday encounters. The other is that
the selection of vocabulary should not come entirely from a particular textbook so that J-SKIT can be a more objective, multi-institutional proficiency test.

Based on these two principles, J-SKIT adopts Kokuritsu Kokugo Kenkyûjo's research results published in 1984, *A Study of Fundamental Vocabulary for Japanese Language Teaching*, to construct the Vocabulary Specification. Their research established a Basic Vocabulary list of 2,000 words, each of which were examined and selected by 22 linguistics and language teaching specialists. The list has been constructed "to gain proper and valid standards for vocabulary which learners of Japanese from abroad (as a second language) should use in the beginning of their study to build a foundation for entering in study and research in their specialized fields or in professional training" (Kokuritsu Kokugo Kenkyûjo, 1984: 4, translation mine).

The final list of J-SKIT Vocabulary Specification has been created based on Basic Vocabulary chosen by Kokuritsu Kokugo Kenkyûjo (1984). From this, approximately 1,300 word and phrases that overlap with at least five of the following documents were selected: Okamoto (1944), Kato (1963, 1964), Kabashima and Yoshida (1971), Bunkashô Kokugoka (1971, 1975), Neustupny (1977), Kokuritsu Kokugo Kenkyûjo (1979), and Shibe (1980). In addition, the J-SKIT includes English-based borrowed words that are commonly used in modern Japanese (Principle I).

The J-SKIT test materials were created using vocabulary in the J-SKIT Vocabulary Specification. The vocabulary in the list appears in all sections of the test. The complete list of J-SKIT Vocabulary Specification is provided in Japanese orthography in Appendix C to avoid misunderstanding with other homonyms.

Of the vocabulary specification, numbers (including both the Chinese-derived and Japanese numeral systems, and frequently used classifiers such as -ji, -hun, -hon, -mai, -satsu, -en, -doru, -nen, etc.), days of the week, months of the year, parts of the body, and family terms (in-group and out-group terms) are considered to be of crucial importance.
Vocabulary within these categories should be incorporated in the Listening and Speaking sections.

4.2.4. Context

Context, which includes both topics and situations of communication, is a very broad concept and creating a list with specific, concrete instances would be highly difficult, if not impossible. To construct a context specification, J-SKIT adopts the study by Finocchiaro (1983: 71-74) on the socio-cultural themes and categories of topics. It is noted, however, that the list serves as representative categories of the context employed in the J-SKIT test material production. In general, all the contexts appearing in J-SKIT are concrete and familiar, and also they are the ones which a learner of Japanese would be likely to encounter most frequently when interacting with Japanese natives in Japan (Principle V). The underlying context is interaction between a learner of Japanese and Japanese natives which takes place primarily in Japan (Assumption IV, V, and VI). The following is the list of J-SKIT Context Specification.

1. About the Individual Learner -- introductions and identification
   A. Greetings, leave-takings, introductions
   B. Identification of self and others
   C. Address, age, phone number, where feasible

2. About the Immediate Classroom
   A. Names and location of parts of the room
   B. Names of instructional materials
   C. Identification of activities (reading, writing, etc.)
   D. The program (hours for various subject areas, activities in and out of class)
   E. Common classroom expressions

Table 5. J-SKIT Context Specifications (adopted from Finocchiaro, 1983: 71-72)

(continued)
Table 5 (continued)

3. About the School and Workplace
   A. Location of rooms and special places in the institution
   B. People in the institution (names, functions, special services)
   C. Rules and regulations (times of arrival, etc.)
   D. School activities such as club programs, general organization, assembly programs, newspapers, and magazines
   E. Workplace activities such as duties, customer services, etc.

4. About the Family
   A. Members
   B. Relationship and ages
   C. The home
      a. Rooms and their use
      b. Furnishings
      c. Cleanliness
      d. Safety
   D. Occupations of various members
   E. Meals
   F. Daily health routines
   G. Clothing
   H. Recreational activities

5. About the Immediate Community of the School, Workplace, and Home
   A. Homes
   B. Nonresidential buildings (offices, movies, library, etc.)
   C. Transportation facilities (directions, tickets)
   D. Communication facilities (telephone, mail, newspaper, radio, television)
   E. Consumer services (stores, banks, etc.)
   F. Local government agencies (post office, police station, etc.)
   G. Places of recreational interest (parks, libraries, theaters, outdoor cafes, etc.)
   H. Educational opportunities
   I. Places of worship
   J. Formulas used in telephoning
   K. Current events

Contexts exemplified above should function as a guideline for J-SKIT test material production. Function, structure, and vocabulary employed in such contexts should be authentic and plausible as a story (Assumption V). The last category of the list, About the Immediate Community of the School, Workplace, and Home, should be incorporated into the test to ensure "survival" settings.
4.2.5. Orthography

Orthography is another factor involved in Reading and Writing sections of J-SKIT. Although it does not consist the major pillars of the project, for efficient test material production J-SKIT establishes orthography specification as follows:

Arabic numerals
Diacritics for sentence composition (ten, maru, kakko, kagikakko, etc.)
Hiragana (all the characters and their possible combination, including the small tsu, ya, yu, yo, nigori, maru, vowel lengthening)
Katakana (all the characters and their possible combination, including the small tsu, ya, yu, yo, nigori, maru, -- (vowel lengthening) and the innovative representation used frequently by Japanese natives)
Kanji (basic characters which appear in J-SKIT Vocabulary Specification (Appendix C)

Table 6. J-SKIT Orthography Specifications

63
CHAPTER 5

J-SKIT: FORMAT SPECIFICATIONS

5.1. Introduction

This chapter describes the skeleton of J-SKIT and specifically focuses on the overall organization as a whole, and the format and appearance of each part of the test. Based on the project assumptions and principles, J-SKIT is designed so that it can be easily administered in most institutions, and so that it can be easily scored and handled. Each part and item in J-SKIT is arranged and presented to ensure single skill testing (Principle III).

5.2. J-SKIT: General Format Specification

5.2.1. Registration and Administration

J-SKIT was developed and managed by the National Foreign Language Resource Center at The Ohio State University (OSU NFLRC). Institutions and individuals who wish to take the test should contact OSU NFLRC through a supervisor to receive the testing materials and accompanying Supervisor's Guide by mail. For test administration, a supervisor is responsible for the test security, the actual administration of the test, correspondence to OSU NFLRC, returning the test materials, and receiving scores and feedback. A detailed Supervisor's Guide enables institutions and individuals to administer the test without any special knowledge of Japanese. The administration procedures are summarized in the chart below:
Figure 3. J-SKIT Administration Procedures

<table>
<thead>
<tr>
<th>OSU NFLRC Test Department</th>
</tr>
</thead>
<tbody>
<tr>
<td>Institutions/Individuals</td>
</tr>
<tr>
<td>designated supervisor</td>
</tr>
<tr>
<td>securing a test site, preparing equipment, safekeeping test materials</td>
</tr>
<tr>
<td>administrating the test, collecting test materials</td>
</tr>
</tbody>
</table>

---contact for test registration---
<---sending a registration pamphlet---
---request for test registration---
---notification of the test mode for Speaking---
<---sending test materials and Supervisor's Guide---

---clarification of procedures and other necessary correspondence---

---sending back all the test materials---

<---sending individual scores and feedback---

scoring, keeping database

5.2.2. Target Population

J-SKIT is designed primarily for American high school and college students of Japanese and for other English-speaking individuals (Principle I). All the directions, questions, and answer choices are provided in English except in the Structure section. It is assumed that test takers have a good enough reading skill in English to follow the test procedures. J-SKIT is also targeted at students of lower proficiency (below Advanced as described in the ACTFL Guidelines), designed to measure students' proficiency in Japanese in all the four skills to provide feedback on students' ability and progress, and to be used as a placement test and program evaluation (Principle IV).
5.2.3. Organization of the Test

There are two parts to the J-SKIT: Part 1 includes three subsections (Listening Comprehension, Structure, and Reading) and Part 2 includes two subsections (Speaking and Writing). Each part has a separate booklet and a different testing format.

Part 1 is designed to measure the testee's receptive skills (listening and speaking) and knowledge of syntax within the context of Japanese culture. Part 1 takes approximately 50 minutes to administer; 5 minutes for general administration prior to the test, 10 minutes for Listening Comprehension, 15 minutes for Structure, and 15 minutes for Reading. There are 20 questions in Listening Comprehension, 15 for Structure, and 10 (5-7 texts) for Reading. All of the questions in part 1 are multiple-choice with four answer choices for each question. Separate answer sheets are provided with the test booklet to mark answers, which will be scored by a scanning machine. A prerecorded cassette tape containing directions and test materials for Listening Comprehension accompanies the booklet.

Part 2 contains Speaking and Writing subsections to measure students’ productive skills in Japanese within the context of Japanese culture. It takes approximately 50 minutes to administer. In the Speaking section, testees are to read the directions and are given tasks in the booklet in English, and then they record their speech samples in Japanese on tape. There are six subsections to the Speaking section: Speech Samples, Self Introduction, Information Gathering, Telephone Message, Story Telling, and Giving Personal Opinions. These tasks have their own specific purposes and procedures, which are described in detail later in this chapter. The accompanying prerecorded tape contains directions and cues in Japanese along with the booklet. The Speaking section can be administered on an individual basis, and also with a group of testees if an institution has access to a language laboratory.
In the Writing section, students are to read the directions and perform tasks by writing in Japanese. There are two main tasks in the Writing section, one focusing on orthography and the other on textual composition. Testees are to write in Japanese on a separate answer sheet.

The basic framework of J-SKIT is summarized as follows:

**Part 1 (50 minutes)**

Medium: booklets, a cassette tape (for General Administration, Listening Comprehension, and Structure), and answer sheets (all provided by OSU NFLRC)

Test format: multiple-choice

Equipment: a cassette tape player, #2 pencils, and erasers

- General Administration (5 minutes)
- Listening Comprehension (20 questions, 10 minutes)
- Structure (15 questions, 15 minutes)
- Reading (10 questions, 5-7 texts, 15 minutes)

**Part 2 (50 minutes)**

Medium: booklets, a cassette tape (for General Administration and Speaking), recording tapes (for Speaking), and an answer sheet (for Writing) (all provided by OSU NFLRC)

Test format: simulated conversation (speaking), and task performance (writing)

Equipment: a cassette tape player and a cassette tape recorder (for individual administration), language laboratory (for group administration), pencils, and erasers

- General Administration (5 minutes)
- Speaking (6 subsections, 25 minutes)
- Writing (2 subsections, 15 minutes)

**Table 7. Basic Framework of J-SKIT**

Ideally, one should take both parts of J-SKIT on the same date, Part 1 followed by Part 2. However, it is also possible to take the two parts on different dates, or to register and take only one part of the test, according to the needs and circumstances of the test taker.

The format mentioned above serves as the basic framework of J-SKIT. More detailed format specification on each section is as follows.
5.3. J-SKIT: Format Specification for Each Section

5.3.1. Listening Comprehension

This section tests students' ability to comprehend utterances in Japanese, including monologues and dialogues. The utterances are recorded on tape at a natural speed in order to sound natural and to provide authentic speech stimuli (Assumption V). Each speech stimulus is short enough to minimize the memory burden on the testees, and is given only once to provide realistic task conditions (Principle III, Assumption V). Both male and female voices are presented in hyōjungo --the standard dialect used in Tokyo and the dialect most commonly taught in the U.S. (Assumption IV). All the speech stimuli are recorded and provided on the tape.

The tape contains, for each item, a header in English, Japanese speech stimuli, an English question, and answer choices. The headers and questions are also written in English on the booklet so that the test takers can specifically listen only to the stimuli (Assumption II). Single skill testing is ensured by providing everything in English on the tape except the Japanese speech stimuli, and by printing no Japanese on the booklet. Each speech sample is followed by one question. There is a 20-25 second pause between each item, during which testees read the answer choices and choose the answer.

Testees are to listen to each speech stimulus and answer the questions on the basis of the content of each stimulus by choosing one of four answer choices. The booklet contains a header to describe the context (Assumption VI), a question, and four answer choices for each item, which are all written in English (Principle III). Only the booklet contains the answer choices.

Speech stimuli include various speech styles such as formal, casual, direct, honorific, and humble (Assumption IV) as well as various types of speech such as conversations,
monologues, announcements, TV/radio broadcasts, etc. All of these might be encountered daily (Principle V). All of the tasks are considered to be noninteractive. If it is a conversation, the task is overhearing others' conversation and not participating in it. Questions are to test general understanding as well as detailed comprehension.

The questions must be realistic as a task. For example, a question which asks the location of the conversation is avoided as an unrealistic task in the J-SKIT, because such information should be normally present and apparent when one is listening to a conversation.

The 20 questions cover the function, structure, vocabulary, and context specifications as much as possible. Specifically, of all 20 questions, at least two questions deal with the social relationship between the speakers as reflected in the speech style to test the non-grammatical communicative behavior and socio-cultural understanding (Assumption I and IV). Selection of the items consider all of the five subcategories of the Function Specification, and at least one question should represent each subcategory. At least three questions should test the crucial vocabulary (numbers, days of the week, months of the year, parts of body, and family terms) specified in the Vocabulary Specification. The 20 questions are arranged on the basis of overall difficulty, and sequenced from easy to difficult.

5.3.2. Structure

This section tests students' understanding of syntax and grammatically correct usage of Japanese. Testees are to read and listen to incomplete sentences and fill in the blank with one of four answer choices that is grammatically correct.

This section does not test students' "skills" in Japanese; the task the students are asked to do in the Structure section is something students might do mentally in actual
communication, but not as a goal of communication. Nevertheless, this section is included in J-SKIT to check the knowledge of Japanese grammar (or communicative competence), the result of which is useful in many programs when placing students. This section may be eventually eliminated from J-SKIT after obtaining more data and/or after discovering an alternative, more effective way to test grammar in Japanese.

All 15 sentences respond to a question or initiate a conversation in a specific context. To minimize the influence of the receptive skills, all of the sentences and answer options are provided both on the tape and in the booklet (Principle III). The tape also serves to pace the students. Furthermore, the written representation of Japanese sentences and answer choices are in three systems of representation so that each testee can choose the most familiar system: Japanese orthography with hirigana on every kanji, the modified kunreishiki romanization, and Hepburn Romanization.

Each item contains an English header to describe the situation and context (Assumption VI), a Japanese incomplete sentence with a blank, and four answer choices in Japanese that potentially fill in the blank. The distracters are designed so that they will be ungrammatical or inappropriate only when they are combined with the incomplete sentence. There is a 20-25-second pause between each item, during which testees read the four answer choices, choose one, and mark the answer on the answer sheet.

Like the Listening Comprehension section, speech style is considered in the Structure section. Of the 15 questions, at least one question deals with the appropriate style choice in Japanese. Items include various styles and situations, and the recording of the sentences and answer choices on tape should be both male and female voices. Sentences are kept short, realistic, and natural, to minimize the influence of the receptive skills yet to enhance the authentic use of the language (Assumption V, Principle III). Sentences are ordered so that they constitute a continuum from structurally simplex to more complex.
5.3.3. Reading

This section tests students' comprehension skills in reading texts written in Japanese orthography. This section does not use the prerecorded tape. On the booklet, each item consists of an English header to describe the source and/or nature of the text (Assumption VI), Japanese text, and 1-4 questions in English with 4 answer choices in English for each question. Texts are produced so that they would closely resemble something a student would be likely to encounter in Japan (Assumption V, Principle V). To maintain authenticity, there is no *furigana* that accompanies the text (Assumption V). Students are to read the Japanese text and answer the questions by choosing one of the four answer choices and marking on the answer sheet. Testees can pace themselves in this section within the given 15 minutes.

Each writing system (*hiragana*, *katakana*, and *kanji*) is considered when producing the test materials and selecting the problems. Of all the texts, one or two are specifically designed to test students' reading skills in *katakana*, and another one or two to test reading skills in *hiragana*. Thus students are able to score 25-30% of the questions if they can read *kana*, even without any knowledge of *kanji*. Both skimming and scanning skills are tested, as well as the detailed understanding of the text (Principle V). Types of the texts should include various formats such as notes, menus, signs, advertisements, letters, journals, articles, business cards, and so forth. In addition to the typed text (in various fonts), there should be at least one handwritten text. There should also be at least one text that is written vertically.
5.3.4. Speaking

This section measures the testees' oral communication skills in Japanese (Assumption I). For individual test administration, this section requires a cassette tape player to play the prerecorded master tape which contains directions and simulated exchanges (Principle II), as well as a cassette tape recorder to record utterances made by the testee. If a testing institution has a language laboratory, this portion of the test can be administered to a group of testees. In this case, the testing institution should notify OSU NFLRC to receive one of the two modes of test administration procedures that would be most suitable for a given facility: (1) one master tape with a number of separate recording tapes (the master tape is played by the master booth, and the testees listen to the tape from the main booth and record their utterances onto the recording tape set in individual booths) or (2) a number of master tapes set in individual booths (to which testees listen while they record their utterances on the same tape in a different tape track). This section takes approximately 25 minutes and is paced by the master tape.

There are two parts to the Speaking section: Part A consisting of Speech Samples, and Part B consisting of Task Performance. Part B is further divided into five subsections: Self Introduction, Information Gathering, Telephone Message, Story Telling, and Giving Personal Opinions. Each subsection has a slightly different format, focusing on particular aspects of the speaking skill.

In Part A of the Speaking section, the Speech Samples section, testees are to either initiate a conversation or to respond to a question by using relatively short, sentence-level utterances. There are 15 items in this section. For each item, there is a header to describe the situation and interlocutor (Assumption VI) and a cue in English for utterances (Assumptions II and Principle I), which is provided in both the booklet and the master tape. The English cue specifies what a testee is expected to say, thus controlling the speech
more specifically. If a testee is to respond to a question, the question is provided in Japanese on the tape, after the English cue to simulate a conversation (Principle II). The cue in English is designed to minimize the translation of the English description. The testee is to say what is specified by the English cue in the context described in the header, within the given pause (12-20 seconds).

This section measures the basic functional, structural, and vocabulary performance of the testees (Principle V). The questions consider the Structure, Function, Vocabulary, and Context Specifications and are created accordingly. Of all the 15 questions, there should be at least five questions for initiating a conversation to encourage more interactive conversation. There are at least two questions that specifically test the distinction of word classes (nouns/nominals, i-adjectives/adjectivalas, na-adjectives/na-nominals, and verb/verbal) to measure structurally accurate performance. There should be at least two questions that involve basic vocabulary--numbers (including both the Chinese-originated and Japanese numeral systems, and frequently used classifiers such as -ji, -hun, -hon, -mai, -satsu, -en, -doru, -nen, etc.), days of the week, months of the year, parts of body, and family terms (in-group and out-group terms)--and their appropriate use. Most of the 15 questions are distal-style (-desu/-masu) speech, but there should be some questions that elicit direct-style speech (-da) and a few targeting honorific- and humble-style speech.

The five subsections in Part B of the section, except for the Information Gathering section, aim to elicit longer, discourse-level utterances with less control by English cue sentences. In the Self Introduction section, a testee is to introduce oneself in a simulated context (Principle II), which is described in English. This task performance is to measure one's skills in the section, About the Individual Learner Context Specification. In addition to one's name and hometown, the testee is to talk about experiences with studying the Japanese language, motivation for studying the language, desire as to what to do with the
language, or future plans while visiting Japan, all within the given pause (approximately one minute). This section focuses mainly on fluency or smooth delivery of speech. Context is described in English to elicit formulaic/ritualistic expressions and other conventional elements of speech used in rather formal self-introductions (Assumptions III, IV, and V).

The next Information Gathering section is designed to gauge a testee's awareness and skills in different speech styles. While aiming at relatively short utterances with less control by an English cue, a testee is to perform essentially the same task three times, each of which is to a different interlocutor. Appropriate use of the distal, direct, humble and honorific speech style based on non-grammatical (social) communicative behavior is the main focus of this section (Assumption I). The context is described in English to specify the nature of each task, and it is established so that such a sequence of tasks with different interlocutors would be realistic and authentic (Assumptions II, III, and V).

In the following Telephone Message section, one is expected to leave a message on an answering machine in a given pause of 45 to 60 seconds. This task performance would include skills in telephone courtesy, and functions such as giving directions to one's place, invitation, or response to an invitation. The utterance is guided with a picture in the booklet, and a testee is expected to follow the English cue and the picture when performing the task. This section is to mainly measure one's skills in discourse and conventions in a telephone conversation.

The following Story Telling section is also designed to measure one's skills in discourse. A testee is expected to tell a story according to what is portrayed in the sequence of pictures. The pictures depict a sequence of activities with a twist or an unexpected event in the middle, resulting in the classic four-part structure of a story: introduction, development, turn, and conclusion. Based on the pictures, one tells a story as if the
depicted events were what one has experienced. The course of events are selected to represent ordinary, everyday events. There is a pause of 90 seconds within which one should conclude the story.

Finally, in the Giving Personal Opinions section, one is asked to provide his/her opinion on a certain issue and also his/her ideas as to how the present condition of the issue can be improved in the future. The context and the issue are described in English, and a testee is to provide his/her opinion after the question in Japanese on the tape. There are 90 seconds for each task performance (providing opinion and providing ideas for improvement) during which the testee should finish what one has to say. The selection of the issue is done so that it would not be too open-ended: the issue should represent familiar, everyday topics (such as having a car in the U.S., watching TV, favorite places or pets, and so on), and the question is designed to elicit a choice between two positions (like-dislike, good-bad, support A or B, for example). This advanced-level task performance as described by the ACTFL Guidelines standards concludes the Speaking section.

These six subsections should include representative samples from the Function, Structure, and Vocabulary Specifications while covering representative contexts from each of five subcategories of the Context Specifications.

5.3.5. Writing

The last section of J-SKIT measures learners' practical writing skills in Japanese. There are two tasks in this section: Completing a Form and Writing a Letter. While the former concentrates more on orthography, the latter is to measure compositional skills in Japanese. Testees can pace themselves within the given 15 minutes to finish the two tasks.

In the Completing a Form section, there are directions, description of the context, and a description of the task, all written in English in the booklet (Assumptions II, III; Principle
II). One is expected to fill in a form in Japanese based on the information given in the booklet. In order to elicit the desired orthography, the given information (the English version of the form) controls the content of writing. This form of task performance is chosen since it is an appropriate and authentic task where one does not have to write in a sentence- or text-length format. All of the three writing systems (hiragana, katakana, and kanji) should appear in this task, as well as other elements specified in Orthography Specification in Chapter 4.

In the Writing a Letter section, directions, context, and the nature of the task are all written in English in the booklet (Assumptions II, III; Principle II). A testee is to write a letter of request or gratitude, two of the major frequently used functions in writing. While the previous Completing a Form section tests one's knowledge and skill in producing orthography, this Writing a Letter task focuses more on the textual composition, the appropriate use of formulaic expressions, and other conventions when writing a letter (Assumptions III, IV, V, and VI).

5.4. Conclusion

J-SKIT is designed to realize the project assumptions and principles stated in Chapter 3 and to incorporate the function, structure, vocabulary, context, and orthography specified in Chapter 4. J-SKIT also considers the current predicament of Japanese language pedagogy in the United States and utilizes booklets and cassette tapes to create a simulated single-skill proficiency test which is readily applicable and easily administered in most proficiency-oriented programs. For more concrete examples of the J-SKIT format specification, see the script of recorded directions and example questions provided in Appendix A.
CHAPTER 6

J-SKIT: GRADING PROCEDURES AND CRITERIA

6.1. Introduction

This chapter describes the grading procedures and criteria developed for the second version of J-SKIT administered in March, 1996 at The Ohio State University. The grading procedures have been modified and improved from the first version of J-SKIT produced in 1995 as the format of the test has been changed. The present grading statement was created and used for the actual grading of the second version of the test.

This chapter focuses mainly on the grading procedures and criteria for Part 2 of J-SKIT (Speaking and Writing), as the scores of Part 1 are machine-scanned. It is to be noted, however, that Part 1 of the second version of J-SKIT uses a scoring procedure in which each correct answer is given one point, each incorrect answer is given minus one-third (-1/3) point, and each unanswered item is given zero (0) points. This grading procedure was adopted to ensure the validity of the test by providing an environment in which students answer only when they are confident about the item, thus eliminating random guesses.

The maximum overall score for J-SKIT is 100 points, which are distributed as follows:
Listening Comprehension -- 20 points (raw score = 20)
  Structure -- 20 points (raw score = 15; converted into 20-point scale)
  Reading -- 20 points (raw score = 10; converted into 20-point scale)
  Speaking -- 20 points (raw score = 12; converted into 20-point scale)
  Writing -- 20 points (raw score = 12; converted into 20-point scale)

Total -- 100 points

Table 8. J-SKIT Maximum Points Distribution for Each Section

Part 2 of J-SKIT, Speaking and Writing, was graded by two trained graders to minimize the effects of subjectivity. All of the graders had experience in teaching Japanese as a second language, either as a Graduate Teaching Assistant, or as an Instructor at The Ohio State University. The handout on grading procedures and criteria were given to each grader and an explanatory meeting was held before the actual grading. To further ensure a bias-free evaluation, the names of the testees were kept confidential so that the graders would not know whose test they were evaluating, or to which level the testee belonged. Details of grading procedures and criteria for the Speaking and Writing sections follow.

6.2. Grading Procedures and Criteria for Speaking

The second version of J-SKIT contains fifteen tasks for Speech Samples, four subtasks for Self Introduction, three subtasks for Information Gathering, two subtasks for Telephone Message, one for Story Telling, plus two subtasks for Giving Personal Opinions. All twenty-seven tasks are graded in terms of the four criteria: Linguistic Performance, Socio-Cultural Appropriateness, Discourse Skills and Delivery, and Pronunciation. Linguistic Performance includes the performance of grammatical and strategic competence, Socio-cultural Appropriateness includes sociolinguistic competence, Discourse Skills and Delivery include discourse and strategic competence of Campbell and Wales (1970), Hymes (1972), and Munby (1978), respectively (Assumption I). Each
criterion has a maximum of three points, and the twelve points of the maximum raw points are converted into a twenty-point scale.

All of the grading procedures are conducted with the assumption that the recorded speech samples take place in the context of interaction between the testee and native speakers of Japanese who are not accustomed to interacting with foreigners (Assumption IV). All of the criteria are based on the comprehensibility of the speech and appropriateness of speech style.

One note is that all four criteria are established in order to evaluate different aspects of task performance. If a problematic element is penalized under one criterion, the same item will not be penalized under another criterion for the same performance: For instance, if a testee misuses certain discourse devices and thus alters the message, that error is penalized under Linguistic Performance, but the fact that one used it is positively evaluated under Discourse Skills and Delivery criterion and will not be penalized there. This rule does not hold to all the cases, however. If a testee performs a task with speech that sounds unrecognizable as Japanese and thus does not communicate any information, this would be penalized under both Linguistic Performance and Pronunciation criteria. Except for such cases, evaluation procedures follow the basic rule: different criteria evaluate different aspects of task performance.

6.2.1. Linguistic Performance

This criterion measures how successful each performance is in terms of choices and accuracy in structure and vocabulary. This criterion is based on the comprehensibility of the performed task by target-natives, and closely looks at the transactional significance of performance. It "reflects what the utterance contributes to satisfaction of the conditions of interaction structures and instantiation of their domain predicates" (Geis, 1995: 11), or the degree of how successfully the function of the given task is communicated to the listener,
temporarily disregarding the stylistic elements, speed of delivery, and quality of speech sounds. Every performed task is evaluated by this criterion, and the average score of all of the performed tasks is calculated. This criterion involves function, structural accuracy, and vocabulary choices (Principle V).

A full score of three points is given to a performance whose choice of structure and vocabulary is accurate and appropriate for the given task in terms of function. In terms of structural accuracy, the performance demonstrates a confident command of the language with no or few grammatical errors. In terms of content and choice of vocabulary, the performance fully expresses the given task and the utterance resembles what a native speaker would most likely say in the given context.

Two points are given to a performance whose choice of structure and vocabulary is adequate to perform the given function. The student’s choices may not be the best, but the intention or communicative purpose of the utterance would be easily understood with the help of contextual information and a little work on the part of the listener. In terms of structural accuracy, the performance demonstrates a good command of the language with few grammatical errors which would not seriously damage the comprehensibility. In terms of content and choice of vocabulary, the performance sufficiently covers the given task. A lack of information would not affect the completion of the given task, though it may require some work on the part of the interlocutor.

One point is given to a performance whose intention or communicative purpose is difficult to understand due to a limited command of the language. In terms of structural accuracy, the performance contains serious grammatical errors that would lead the unsympathetic listener to misunderstand the intention of the utterance or to experience difficulty in interpreting the utterance. In terms of content and choice of vocabulary, the performance only partially covers the given task, and this incomplete information or misuse of vocabulary would result in altering the given task or suspension of understanding.
No points are given to utterances whose performance is totally irrelevant to the given task, and that would leave even a sympathetic listener at a loss or confused. With the function unachieved, such a performance demonstrates that the given task is clearly beyond the capacity of the testee in terms of both structure and vocabulary. This level of performance is also characterized by utterances that indicate that the testee has given up on performing the task (such as dekimasen, wakarimasen, sumimasen, dame desu, etc.), or no utterance at all in Japanese.

6.2.2. Socio-cultural Appropriateness

The criterion of Socio-cultural Appropriateness measures the appropriateness of performance in terms of speech style and other stylistic components in the given context (Assumptions IV, V and VI). This criterion, as opposed to transactional significance measured in Linguistic Performance, focuses on the "interactional significance that reflects what it contributes to face-work (Goffman 1967)" (Geis, 1995: 11). This criterion also has a maximum of three points. All of the students’ work is evaluated and the average score of all of the performed tasks is calculated. Unperformed tasks, which are given a zero in Linguistic Function, are not evaluated and do not affect the final average point.

Three points are given to a performance which demonstrates successful choices of speech style and other stylistic elements, including the proper use of honorific, humble, neutral levels of politeness, distal and direct, and careful and casual speech styles as appropriate to the given situation and social roles of the speech participants. This performance utilizes the speech style that most educated native speakers of Japanese would choose in a given situation. This level of performance demonstrates an excellent command of the system of speech styles.

Two points are given to a performance whose choice of speech style is adequate. The response may not be the best choice, but the choice of acceptable, inoffensive language is
in most of the situations similar to that of a native. The choice of style may be also considered to be too polite for the given situation and typically characterized by the use of the distal style (-masu and desu style) to a close friend, or use of humble style to a co-worker when such style is not expected. This level of performance demonstrates the testee's awareness of different speech styles, and there may be an error in terms of the formation of particular polite expressions but which would not interfere with comprehensibility.

One point is given to a performance which demonstrates a limited command of speech styles. The choice of speech style is most likely unsuitable and inappropriate for the given situation, which might be considered rude by some native speakers of Japanese.

Zero points, rarely given, are for a performance whose choice of speech style and other stylistic elements would most likely be considered to be extremely rude, characterized, for example, by the use of the direct imperative form to a teacher or supervisor. This level of performance would most likely offend even a sympathetic listener.

6.2.3. Discourse Skills and Delivery

This criterion measures how successfully the speaker could convey the message in a smooth and coherent fashion. Since all of the tasks in J-SKIT are given a certain pause within which testees are to finish their task performance, this criterion focuses on how many different types of discursive devices are employed in the entire Speaking section. After the selection of the test items, a specific list of the discourse devices is created.

The second version of J-SKIT contains the following 11 targeted discourse devices: (1) the use of appropriate hesitation noises (anò, etto, sô desu nê, sonô, nan to iimasu ka, etc.), including ne as in sore de ne; (2) sentence connectors indicating consequence (soshite, sore de, etc.); (3) sentence connectors indicating contradictory or opposite consequence (kedo, ga, de mo, shikashi, etc.); (4) discourse devices indicating a sequence
of activities (V-te, V-te; V-te kara; sore kara; V-ta ato de, etc.); (5) discourse devices indicating cause or reason (no de, kara, sore de, naze nara, naze ka to iu to, sô iu wake de, etc.); (6) discourse devices indicating exemplification (tatoeba, N ya N, to ka, nado, nanka, V-tari V-tari suru, etc.); (7) discourse devices indicating purpose (V-ru tame ni, V si ni, V-ru mokuteki de, etc.); (8) use of deictics (sô iu no wa, sono koto, honya e itte, soko de..., etc.); (9) use of sentence modifiers (kinô Tanaka-san ga katta hon, etc.); (10) use of the extended predicate (plus kedo, ga) to establish the discourse background; and (11) discourse devices indicating summarization (to iu wake de, tsumari, etc.).

The maximum three points are given to a testee who uses nine or more items of the discourse elements listed below. Three points indicate that the performance demonstrates very smooth and coherent discourse skills which make the conversation easy to follow. Two points are given to someone who uses seven or more items, who demonstrates fairly smooth and coherent discourse skills, but sometimes lacks appropriate discourse devices. One point is given to someone who uses four or more items, who demonstrates limited skills in discourse that would sometimes make the conversation rather choppy and incoherent. No points are given to someone who uses three or fewer items, who demonstrates few discourse skills that would in most cases render the speaker unable to recombine the sentences in a coherent fashion.
6.2.4. Pronunciation

This criterion is based on measurements of the accuracy of speech sounds, including suprasegmentals (accent and intonation). Each task performed in the Speech Samples section is evaluated and the average score is calculated.

Three points are given to a performance that demonstrates accurate and proper pronunciation, accent, and intonation that are almost indistinguishable from those of native speakers of Japanese. The performance includes no or few errors in pronunciation, accent, and intonation that would render the conversation comprehensible to Japanese natives who are not accustomed to interacting with foreigners.

Two points are given to a performance which demonstrates fairly good pronunciation, accent, and intonation with a few easily noticeable errors. Such errors, however, would not affect comprehension and would still be easily comprehensible to native speakers of Japanese.

One point is given to an utterance that demonstrates acceptable pronunciation, accent, and intonation with some easily noticeable errors that might occasionally be difficult to understand and thus affect the comprehensibility of the utterance. This level of performance contains some consistent errors that obviously reflect the influence of the first language.

Zero points, rarely given, are for a performance which contains many errors in pronunciation, accent, and intonation that would most likely interfere with the listener's comprehension, or which is unrecognizable as Japanese speech.

6.3. Grading Procedures and Criteria for Writing

Each of the two tasks in the Writing section is evaluated with the four criteria: Orthographic Representation, Linguistic Performance, Socio-cultural Appropriateness, and Text Composition. Each criterion has maximum of three points, and the twelve points of
the maximum raw points are converted into a twenty-point scale. Except for Orthographic Representation, all of the criteria follow the same evaluation procedure as in the Speaking Section, explained again below.

6.3.1. Orthographic Representation

This criterion measures how accurately and appropriately the writing symbols are used. Here the writing symbols mean the three writing styles plus other writing conventions specified in the Orthography Specifications in Chapter 4. Both tasks are evaluated in terms of overall task performance and the average score is calculated.

Four points, the maximum score, is given to a performance with accurate and appropriate use of writing symbols and orthographic conventions with no or few errors. This level of performance demonstrates an excellent knowledge of writing conventions and the written samples closely resemble those of native speakers of Japanese.

Three points are given to a performance which demonstrates a fairly accurate and appropriate use of the writing symbols with some errors, although such errors do not interfere with comprehensibility. This level of performance demonstrates an excellent command of both hiragana and katakana with no or few mistakes, but the lack of kanji knowledge is evident.

Two points are given for an acceptable command of writing, which contains some errors in writing, including a few that would interfere with comprehensibility. This level of performance is characterized by a few errors in hiragana and katakana, but the use of kanji is restricted to only a few, simple characters.

One point is given to a performance which demonstrates a poor command of writing, which contains many errors in writing, including the mixture of the kana systems that would most likely interfere with comprehensibility. This level of performance is also
characterized by a limited, if present at all, use of only a few, highly frequent kanji characters.

No points are given to writing samples which are considered to be unintelligible; i.e., samples which demonstrate knowledge of only a few Japanese characters or none. The text may be entirely in romanization, or sporadically may use of kana in a romanized text. No performance also receives zero points.

6.3.2. Linguistic Performance and Socio-cultural Appropriateness

The same evaluation criteria in the Speaking section is also used in the Writing section. The overall task performance is evaluated, and the average scores of the two tasks is calculated. A maximum score of three is given to an excellent performance that fully expresses the given task, and resembles what native speakers of Japanese would write. Two points are for a fairly good task performance with a few grammatical errors or lack of information, but the overall intention of the text is still comprehensible to native Japanese readers. One point is for a performance whose grammatical errors or lack of information is so severe that the intention of the task is either obscured or misinterpreted by the readers. No points are given to a performance which indicates that the task is clearly beyond the capacity of the testee, and the intention of the text is likely not conveyed.

The Socio-cultural Appropriateness criterion used in the Writing section is also very similar to that of the Speaking section. However, the formal nature of writing affects the expectations of the reader, and the grading criteria holds slightly higher expectations with regard to levels of politeness. The maximum three points are given to a performance which demonstrates an excellent command of the writing courtesies with sophisticated use of different levels of politeness. Two points are given to a performance characterized by the consistent use of the distal (desu/masu) style with a few uses of honorific, humble, and neutral politeness, although the overall effect is not rude or offensive to the reader. One
point is given to a performance whose use of style might be regarded to be crude or inappropriate by some native Japanese readers. No points are given to a performance whose use of style is most likely to be considered rude by most native Japanese readers.

Unperformed tasks, which receive a zero in Orthographic Representation, are not evaluated in either Linguistic Performance or Socio-cultural Appropriateness.

6.3.3. Text Composition

This criterion evaluates how successfully the text is organized in a smooth and coherent fashion. After the selection of the tasks, the use of the following eleven text devices are important: (1) the format of the letter, including the seasonal greetings, opening remarks, and closure; (2) ritualistic expressions, such as *yoroshiku onegai shimasu*, etc.; (3) textual devices indicating contradictory or opposite consequences (*ga, shikashi*, etc.); (4) sentence connectors indicating consequence (*soshite, sore de*, etc.); (5) the use of sentence modifiers; (6) the use of the extended predicate (*no da* and its derived forms); (7) text devices indicating cause or reason (*no de, kara, sore de, naze nara, naze ka to iu to, sō iu wake de*, etc.); (8) textual devices indicating purpose (*V-ru tame ni, V si ni, V-ru mokuteki de*, etc.); (9) the use of deictics (*sō iu no wa, sono koto, honya e itte, soko de...*, etc.); (10) textual devices indicating summarization (*sō iu koto de, tsumari*, etc.); and (11) textual devices indicating conditions or provisions, including *moshi yoroshikereba, dekireba*, etc.

A performance containing nine or more of the textual devices above is given a maximum three points, six or more is given two points, three or more is given one point, and two or less is given no points.

6.4. Conclusion

The four criteria for both Speaking and Writing reflect the production assumptions and are employed by graders experienced in teaching Japanese as a second language in the
United States. More accurate, objective grading was promoted by providing all the grading procedures in a form of printed handouts and also by having an explanatory meeting by the producer of the grading procedures and criteria prior to actual evaluation. Assigning two graders for one testee, calculating the average points of the two graders, and concealing the names and the level of the testees also ensures objectivity. After the raw scores were calculated, they are converted into a twenty-point scale to be equal to other parts of the test. The overall converted scores, as well as the raw scores and subcategory points were inputted in a spreadsheet and submitted to a statistician for the further analysis of the test.
CHAPTER 7

J-SKIT: STATISTIC ANALYSES

7.1. Introduction

This chapter mainly describes statistical data and analysis of the second version of J-SKIT administered in March, 1996. After the completion of all the test scoring, the data was submitted to a statistician to evaluate the effectiveness of the J-SKIT. The test scores were also combined with the Demographic Survey conducted by OSU NFLRC for some of the research questions.

There are some points and problems reported. One is that “although some of the original data were converted to a form that equated the number of items on each section of the test, [...] raw scores were used in the analysis due to the difficulty explaining converted scores. All results are exactly the same whether raw or converted scores are used, so talking about them in either raw or converted form is an individual decision.” Second is that “since this data is non-experimental in that no variables were manipulated, no conclusions about causation can be made. We can only talk about the relationship between and among variables, not that one causes changes in the other.” The third point is that “many analyses were done due to the exploratory nature of this project. In actual hypothesis testing, only a limited number of tests are done to avoid capitalizing on chance.”

---

1 I am indebted to Ms. Virginia Replogle for her help in the statistical analyses presented in this chapter. All of the quotation marks in this chapter indicate the citation from her report.

89
Finally, the fourth problem is that “the small number of participants, especially in some of the course levels, causes concern as to whether the estimate is really representative of the group.” All of the above points should be taken into consideration for the interpretation of the statistical analyses presented below.

Also, since the improved second version of J-SKIT has been administered only at one institution; it is yet too early to speculate the validity of J-SKIT as a multi-institutional standardized proficiency test. More data must be collected from more students at various institutions to determine the possible use of the J-SKIT as a proficiency test for most of the proficiency-oriented programs.

7.2. Administration Data

The second version of J-SKIT was created in the Spring Quarter at The Ohio State University, after the creation and critical investigation of the first version. The first version of J-SKIT was created in 1995 and administered as a pilot field test at more than 70 different institutions for more than 100 participants. The data obtained by this first version were used as one of the contributing information for the placement of the participants in Language: Japan 1995 program, an intensive summer language program in Gujo-Hachiman, Japan. The same first version was also administered at Washington University in St. Louis in April, 1995, to compare the test scores with Language: Japan participants, and also with the SOPI speaking proficiency level the students obtained in the same month. After collecting the data, the first version of J-SKIT was revised in format and grading procedures, and renovated as the second version.

The second version of J-SKIT was administered from February 27th to March 8th, 1996, at The Ohio State University as a pilot testing of the project. All of the students in the entire Japanese language program at this institution participated in taking the test as a
requirement of course fulfillment. Of all the students, 114 students actually took both Parts 1 and 2 of the test.

Part 1 of J-SKIT was administered on February 27, 1996 for the entire group of participants simultaneously, except for a few participants who could not take the test on the same day. Almost all of the participants took this part of the test in the same classroom at the same time. The few who did not take the test on that day were asked to take it with Part 2 of the test.

Part 2 of J-SKIT was administered during the Examination week of the Spring Quarter, the first week of March, 1996. A language laboratory was chosen as a test site, and groups of participants (two to twenty) took the test at the allotted time slot for which each participant had signed up. Participants had more than 10 time slots from which they could choose the most convenient time. Due to the conditions of the laboratory, Part 2 of the test was administered in Mode 2: a number of master tapes were set up in individual booths, and testees listened to them while recording their utterances on the same tape in a different tape track.

The following chart indicates the population of the participants.
<table>
<thead>
<tr>
<th>Course Number</th>
<th>Course Title</th>
<th>population</th>
</tr>
</thead>
<tbody>
<tr>
<td>102</td>
<td>First-Year Japanese</td>
<td>62</td>
</tr>
<tr>
<td>205</td>
<td>Second-Year Japanese</td>
<td>14</td>
</tr>
<tr>
<td>310/311</td>
<td>Intensive Second/Third-Year Japanese</td>
<td>10</td>
</tr>
<tr>
<td>508</td>
<td>Third-Year Japanese</td>
<td>7</td>
</tr>
<tr>
<td>603</td>
<td>Classical Japanese</td>
<td>3</td>
</tr>
<tr>
<td>611</td>
<td>Fourth-Year Japanese</td>
<td>13</td>
</tr>
<tr>
<td>693</td>
<td>Fifth-Year Japanese</td>
<td>5</td>
</tr>
</tbody>
</table>

TOTAL N = 114

Table 9. J-SKIT Version 2.0 Participants

Of all the courses above, 603 is a course for students who have advanced skills in Japanese. All three participants in the course were American graduate students studying Japanese literature or linguistics in the Department of East Asian Languages and Literatures.

7.3. Statistics and Discussions
7.3.1. Course Level Versus Total Score

A number of questions were asked with regard to the J-SKIT statistics. This section describes the effectiveness of J-SKIT as an indicator of students’ proficiency, and usefulness of J-SKIT as a placement test. This section also includes the result in examining various parts of the test.

---

2 The questions include the correlation between course level versus total score, number of languages other than Japanese studied by students versus total score, relationship among various parts of the test, pre-collegiate experience in studying Japanese versus total score, and so on. The possible additional analysis questions will be presented in section 7.4. of this chapter. Due to the time constraints, however, some of the questions which were already addressed were not statistically analyzed at the time of the completion of this thesis.
"The first research question asked if the test was able to discriminate between course levels. An ANOVA was done due to the categorical nature of the level variable. The following is the SYSTAT output. Where LEVELS is the course level, TOTAL is the total raw score on the test [, maximum 69 points], and N is the number of participants:

Levels Encountered During Processing Are: LEVELS

102 205 310 508 603 611 693

2 Cases deleted due to missing data.

Dependent Variables: TOTAL N: 114  Multiple R: 0.843  Squared Multiple R: 0.710

Analysis of Variance

<table>
<thead>
<tr>
<th>SOURCE</th>
<th>SUM-OF SQUARE</th>
<th>DF</th>
<th>MEAN-SQUARE</th>
<th>F-RATIO</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>LEVELS</td>
<td>20998.258</td>
<td>6</td>
<td>3499.710</td>
<td>43.628</td>
<td>0.000</td>
</tr>
<tr>
<td>ERROR</td>
<td>8583.303</td>
<td>107</td>
<td>80.218</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 10. ANOVA Analysis on Course Level vs. Total Score

"Since F-Ration is significant below the .05 level (the p-value in the table reads ‘0.000’), at least one of the Level means of Total score is different from the rest. [...] The following is the list and a bar graph of the means on Total score by Level, where SE is the Standard Error and N is the number of participants:
<table>
<thead>
<tr>
<th>LEVELS</th>
<th>LS MEAN</th>
<th>SE</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>102</td>
<td>18.028</td>
<td>1.137</td>
<td>62</td>
</tr>
<tr>
<td>205</td>
<td>31.785</td>
<td>2.394</td>
<td>14</td>
</tr>
<tr>
<td>310</td>
<td>41.248</td>
<td>2.832</td>
<td>10</td>
</tr>
<tr>
<td>508</td>
<td>31.134</td>
<td>3.385</td>
<td>7</td>
</tr>
<tr>
<td>603</td>
<td>66.330</td>
<td>5.171</td>
<td>3</td>
</tr>
<tr>
<td>611</td>
<td>47.277</td>
<td>2.484</td>
<td>13</td>
</tr>
<tr>
<td>693</td>
<td>54.986</td>
<td>4.005</td>
<td>5</td>
</tr>
</tbody>
</table>

Figure 4. Least Squares Mean for Each Level

"In order to determine which means differ from the others, a Post-Hoc test must be done. A Tukey test [was] used in this case to give critical values of differences among the means. The following is the SYSTAT output of this test:
Table 11. Tukey HSD Multiple Comparisons of The Ohio State University

<table>
<thead>
<tr>
<th></th>
<th>102</th>
<th>205</th>
<th>310</th>
<th>508</th>
<th>603</th>
<th>611</th>
<th>693</th>
</tr>
</thead>
<tbody>
<tr>
<td>102</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>205</td>
<td>0.000</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>310</td>
<td>0.000</td>
<td>0.152</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>508</td>
<td>0.007</td>
<td>1.000</td>
<td>0.258</td>
<td>1.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>603</td>
<td>0.000</td>
<td>0.000</td>
<td>0.001</td>
<td>0.000</td>
<td>1.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>611</td>
<td>0.000</td>
<td>0.000</td>
<td>0.682</td>
<td>0.004</td>
<td>0.020</td>
<td>1.000</td>
<td></td>
</tr>
<tr>
<td>693</td>
<td>0.000</td>
<td>0.000</td>
<td>0.085</td>
<td>0.000</td>
<td>0.595</td>
<td>0.660</td>
<td>1.000</td>
</tr>
</tbody>
</table>

"By comparing this matrix of probabilities to the standard .05 point, the following results are found:

"102 is different from all other levels. 205 is different from 102, 603, 611, and 693. 310 is different from 102 and 693. 508 is different from 102, 603, 611, and 693. 603 is different from 102, 205, 310, 508, and 611. 611 is different from 102, 205, 508, and 603. 693 is different from 102, 205, and 508.

"While the test is able to differentiate most of the levels from each other, there are some with which it has trouble. [...] One note of caution that should be kept in mind throughout the report is that some of the Levels have extremely few people in them (for example, 603 has only 3). This tends to increase the standard error of this group, which could obscure potential differences."

Given that the Japanese program at The Ohio State University provides a proficiency-oriented instruction, sharing the same assumptions with the J-SKIT Basic Assumptions, the results of the analysis indicate that the J-SKIT is effective in discriminating the course level of students at the institution, and thus it is useful as an indicator of students' proficiency and as a placement test. The low scores of the 508 level were discussed informally among instructors and graduate teaching associates of the program after the finding of the results. Some of the discussants mentioned that the mean of the total score
of the test, even though it seems lower than the expected one, does seem to indicate the students' proficiency since some of the students were not doing well in regular classroom performance. Although this opinion is not supported in statistical analyses, it is worthwhile to mention this as a possible reason for the unexpected result of the test. As for the effectiveness of J-SKIT, the statistical analysis indicates that J-SKIT is effective in differentiating students' proficiency and would be of potential use as a placement test and as a tool for program evaluations.

Although no statistical analyses have yet been done regarding a possible correlation among the section scores and the level of the participants, the following graph on the converted section scores and levels of students is presented, suggesting that there would be a statistically significant correlation.
7.3.2. Relationship Among Various parts of the Test

"These analyses begin to address the question of how various parts of the test relate to each other and to the total score of the test." After the correlation among Part 1, Part 2, Listening Comprehension, Structure, Reading, Speaking, and Writing are examined, "various Simultaneous and Stepwise Regressions were done" to reveal the relationship by seeking "the best-fitting line or linear equation (as indicated by the least error) to a set of data." The result "clearly shows that in trying to predict Part 2 test scores from scores on Part 1, Structure scores are not nearly as important as Listening and Reading scores. In fact, not including them makes the Adjusted Squared Multiple R (the amount of variance in
the dependent variable accounted for by the linear combination of the independent variables) slightly higher due to the reduction in the number of these predictor variables.”

It is hypothesized that the regression line for Speaking and Writing by part 1 scores are:

Predicted Speaking score = 4.068+(0.212) (Listening score)+(0.179) (Reading score)

Predicted Writing score = 0.578+(0.215) (Listening score)+(0.375) (Reading score)

The predicted Speaking score equation “accounts for a little over one-third of the variability in Speaking scores.” As for the predicted Writing score equation, “the amount of variance in Writing score that is accounted for by [the equation] is 0.556.”

The above statistical analyses indicate the possibility of eliminating the Structure section from the J-SKIT to improve the statistical accountability by increasing the number of questions for both the Listening Comprehension and Reading sections. Considering the fact that the Structure section does not test students’ "skills" in Japanese and the task the students are asked to do in the Structure section is something students might do mentally in actual communication, but not as a goal of communication, eliminating the entire Structure section would also strengthen the realization of the test as a Japanese “Skills” Test. The analyses also indicate that, at least for students at The Ohio State University, even if they did not have the opportunity to take Part 2 of the test for various reasons, their Speaking and Writing ability could have been hypothesized by using the regression equations above. This alternative would be helpful, considering the time-consuming nature of the Part 2 evaluation procedures.

7.3.3. Remaining Statistics Research Questions

There are many more questions that should be posed for statistics research of the project. In order to further assess the validity and effectiveness of the J-SKIT, the following questions must be addressed and examined statistically for the administration of the second version: (1) within the same course, correlation between the total score and the
final course grade obtained by each student; (2) correlation between the course level and scores in each section of the test; (3) item analysis of the test as to which test items more successfully differentiate the levels of the students, and how many items are necessary to have the statistical validity of the analyses; and (4) inter-rater reliability for the Part 2 scores.

As mentioned elsewhere in this chapter, the data obtained through the administration are not yet enough to hypothesize and verify the effectiveness of J-SKIT as a standardized proficiency test used in a multi-institutional fashion. In order to further the development of the J-SKIT project, a few more test administrations at various institutions (including high schools) are highly desirable to obtain more data.

It is also necessary to evaluate the validity of the J-SKIT by comparing the test scores with other proficiency tests such as OPI, SOPI, SAT-Japanese, to examine the correlation among the test results. For a pilot study, the result from the first version of the J-SKIT at Washington University in St. Louis in 1995 is presented below. Although no statistical analyses have yet been done, the following chart suggests that there are some correlations between the students’ J-SKIT Part 1 scores and the proficiency level obtained by the same students in the same month in SOPI. The chart also indicates a further possibility of assessing students’ proficiency level in speaking by using the J-SKIT Listening Comprehension, Structure, and Reading scores.
<table>
<thead>
<tr>
<th>Classification by SOPI Proficiency Level</th>
<th>J-SKIT Part 1 Score Mean (max. 45 points)</th>
<th>number of students</th>
</tr>
</thead>
<tbody>
<tr>
<td>Novice-Mid</td>
<td>17.500</td>
<td>4</td>
</tr>
<tr>
<td>Novice-High</td>
<td>22.222</td>
<td>9</td>
</tr>
<tr>
<td>Intermediate-Low</td>
<td>22.286</td>
<td>7</td>
</tr>
<tr>
<td>Intermediate-Mid</td>
<td>29.333</td>
<td>3</td>
</tr>
<tr>
<td>Intermediate-High</td>
<td>32.556</td>
<td>9</td>
</tr>
<tr>
<td>Advanced</td>
<td>36.667</td>
<td>3</td>
</tr>
</tbody>
</table>

Chart 7.4. Correlation between SOPI scores and J-SKIT Part 1 Scores at Washington University
CHAPTER 8

J-SKIT: FOR FUTURE DEVELOPMENT

As described earlier in Chapter 1, the objective of the J-SKIT project is to produce a comprehensive proficiency test in Japanese for the students learning Japanese as a foreign language in high schools or colleges in the United States. Based on Basic Assumptions on proficiency-oriented instruction and Project Principles, two versions of the J-SKIT have been developed and administered in various institutions.

In addition to the need for more data collection mentioned in the previous chapter, there are more questions with regard to the future development of the J-SKIT. For Part 1, the first concern is the organization of the test. Due to the time constraints (two hours) for easy administration, there is a limited number of questions and tasks that can be included in the test. Considering the less importance of the Structure section in the regression equations shown in the previous chapter, eliminating the section and increasing the number of questions and tasks for both Listening Comprehension and Reading sections should be considered for the future revision of the test.

For part 2, the allotted time period for each task in the Speaking section should be reconsidered as well, for at least in a few tasks most of the students, even for the advanced ones, could not finish their utterances.
The grading procedures for no utterances could be improved as well. With the present grading procedures it is difficult to differentiate in score the “no utterances” and “utterances that do not complete the task in comprehensible manner.” It might be a good way to instruct the students to say “no answer” for tasks clearly beyond their proficiency. The question is whether one should value any kind of utterances (even the unstructured, slipshod ones) or only the utterances that indicate the students’ proficiency in more significant ways.

The description of the speech participants in the Speaking section (such as a friend, colleague, and supervisor) also need to be revised. Just words on the interlocutors are not enough for students to imagine the surrounding contexts. In order to better simulate the real conversation, it is recommended that the Speaking section should include the introduction section that includes pictures and descriptions of all the participants who appear in the entire speech section, and the same picture should appear for each task. This helps the students to realize to whom they are talking and what speech style they should use when talking to each person.

The Text Composition criterion in Writing evaluation is another item of concern. Even the best writing samples collected in the administration of the second edition did not meet all of the items listed in the criterion, indicating that the included items are too many. In order to come up with a more appropriate item choice, the Text Composition criterion must include only the items found in the writing samples by native speakers of Japanese collected during to the test creation. The participation of the Japanese natives as a control group of the test must be included for the future administration of the test.

The selection of core items (“equators”) in all the sections needs to be done in order to create more versions of the J-SKIT and maintain the difficulty level of each version. It is suggested that the scoring system should be changed to one similar to SAT-Japanese, which is adjusted statistically through a recentering process with a 200-to-800 scoring
scale, and provides a comparability in relation to the students taking different version of the J-SKIT.

An attempt to describe the level of proficiency as in ACTFL Guidelines is another possibility for the future development of the test project. This can be done after more data in the comparison of J-SKIT scores with other tests are collected.

There should be a concern for the affective aspects of the testees. For students of lower proficiency, going through all of the questions and realizing that they can do only a small number of tasks might be discouraging and thus might lower their motivation for their learning process. When the description of the scores are done, it may be helpful to inform the students of their expected scores prior to the test to reduce their anxiety.

For better simulation of the real conversation and for better administration, the implementation of the J-SKIT into a computer program is also highly recommended. Graphics and the quality of digital sounds would provide a better test environment. When graphics are used, the interface should appear as an interlocutor facing the students, and other situational cues should be provided in the background. Digital recording would better preserve the utterances of the students, and would speed up the evaluation process. If the program could contain tasks of different proficiency levels in function separately, the presentation of the sequence of tasks need not be linear as the present version for Part 1. That is, the program would provide the higher level of tasks step by step, but only after a certain numbers of correct answers within the same level. Eliminating the tasks students would not be able to perform would save time for students of lower proficiency.

The current version of the J-SKIT does contain some room for improvement, and more data collection is necessary to examine the validity and effectiveness of the test as a standardized proficiency test. Nevertheless, it is hoped that the improved versions of the J-SKIT serve well for the need in the field -- to evaluate and compare students and programs of Japanese in U.S. high schools and colleges. The two years of preliminary test, the
theoretical framework, and the obtained data would provide a solid foundation for the future development of the J-SKIT.


Szatrowski, Polly. 1986a. "Danwa no bunseki to kyojuhin: kanyū hyōgen o chušihin ni (I)," in Nihongogaku. vol. 5. no. 11.

_____ 1986b. "Danwa no bunseki to kyojuhin: kanyū hyōgen o chušihin ni (II)," in Nihongogaku. vol. 5. no. 12.

_____ 1987. "Danwa no bunseki to kyojuhin: kanyū hyōgen o chušihin ni (III)," in Nihongogaku. vol. 6. no. 1.


APPENDIX A

J-SKIT SAMPLE QUESTIONS

Section 1: Listening Comprehension

Directions: In this section, you will hear several Japanese monologues and dialogues, including announcements, telephone messages, and conversations between two speakers. For each item, you will hear a short description about its context, the Japanese selection, and one question in English. Both the description of the context and the question are also in your booklet. You will hear the Japanese selections only ONCE, and they are given ONLY on the tape. After listening to the selection and the question, choose the most appropriate answer from the four possible answers that are given in your booklet, and mark your answer sheet accordingly.

Example 1: Two friends are talking about a birthday:

(on tape)
A: ねえ、藤井さん、誕生日、いつ？
B: 私？ 七月二十日。

When is Ms. Fujii’s birthday?  
(a) January 2nd.  
(b) April 8th.  
(c) July 20th.  
(d) August 5th.

Example 2: Two friends are discussing means of transportation:

(on tape)
A: バスで行く？それとも電車にする？
B: そうねえ、バスよりは電車のほうが速いんじゃない？

What is the woman’s observation?  
(a) The bus is faster than the train.  
(b) The bus is cheaper than the train.  
(c) The train is faster than the bus.  
(d) The train is cheaper than the bus.
Section 2: Structure

Directions: In this section, there are 15 incomplete utterances in Japanese. Each utterance follows a short description about the context, which is provided in English. For each utterance, four possible completions are provided in Japanese. The contexts, utterances and completions are provided in your booklet, as well as on this tape. After examining the options, choose the one that best completes the utterance to make it grammatical, meaningful, and culturally appropriate. Then mark your answer on the answer sheet.

In the booklet, the utterances and answer choices are presented in three different writing systems: (1) The Japanese writing with furigana on every kanji, (2) Modified Kunreishiki romanization, and (3) Hepburn Romanization. Look at the examples presented in your booklet and choose one of the three.

(1) Japanese Writing with Furigana ちょっと大きいでしょう。

(2) Modified Kunreishiki Romanization Tyotto ookii desyoo.

(3) Hepburn Romanization Chotto okii desho.

Now having selected the writing system with which you are the most familiar, go on to the designated page.

If you selected the Japanese writing with furigana, go to page 6.
If you selected the modified Kunreishiki romanization, go to page 11.
If you selected the Hepburn Romanization, go to page 17.

Example 1: You are wondering about the location of a new restaurant that you have just heard about.

どこ_____ありますか。

(a) に
(b) へ
(c) で
(d) が

Example 2: Your Japanese teacher is having trouble writing down something in English.

あの、私がお書き_____か。

(a) でしましょう
(b) いましょう
(c) しましょう
(d) になりましょう
Section 3: Reading (15 minutes)

Directions: In this section you will read a selection of various types of text in Japanese, including memos, signs, and short letters and articles. You are provided with some information about each text in English. Every text is provided in authentic Japanese, and followed by question(s) in English. Read each text carefully and answer the questions for each passage on the basis of the text’s content by choosing the most appropriate choice from the four possible answers. Then mark your answer on your answer sheet appropriately.

Example 1: This note is attached to a menu at a coffee shop.

*ランチタイム（10時〜2時）
ランチメニューには
スープまたはサラダを
¥100でサービス致します。

Example 1: What kind of special service is offered from 10:00 to 2:00?

(a) Soup or salad for 100 yen when lunch is ordered.
(b) 100 yen off on soup and salads.
(c) All you can eat salad comes with lunch.
(d) 100 yen off on certain meals on the menu.
Section 1: Speaking

Part A: Speech Samples

Directions: In this section, you will engage in an interaction by responding to another speaker, or by initiating a conversation. For each item, the situation of interaction and your task are described in English, both in the booklet and on the tape. This is followed by a Japanese speech, to which you respond, or by a chime, which signals you to begin the conversation. Say in Japanese what is most appropriate in the given situation to perform the task. You have approximately 20 seconds to record your speech samples for each item.

If you are to respond to another speaker, there will be a beep to signal when time is up. If you are to initiate a conversation, you will hear a response by the other speaker when your time is up. When you hear the beep or the response in Japanese, stop speaking even if you have not finished with your speech, and go on to the next item.

Pay special attention to speech style, and utilize Japanese hesitation noises or any other techniques to make your utterance realistic and natural.

Example 1.
You are approaching a stranger on the street.
Get her attention properly and ask whether there is a bank nearby. Initiate a conversation.

(simulated response on tape: あ、あそこです。)

Example 2.
Your boss has just come back from lunch.
When you hear his return, report to him that his wife called and said she would call back later.

(simulated stimulus on tape: ただいま。)

Example 3.
You and your friend have missed the train.
Following your friend’s speech, express your regret that you did not leave home earlier.

(simulated stimulus on tape: あ〜あ、間に合わなかったねえ。)
Part B: Task performance

In this section, you are to use slightly longer speeches to perform the given tasks.

1. Self Introduction

This task primarily measures your fluency in Japanese. Try to say as much as you can.

Suppose you are now in Japan as an exchange student. Today is your first day of classes. Introduce yourself when the teacher requests you to do so. You have approximately 1 minute to finish your task. Again, stop speaking when you hear a beep.

Include the following information in your self introduction:

Your name, and where you come from;
Your experience in studying Japanese -- where, how long, etc.;
Why you are studying Japanese; and
What you would like to do during your stay in Japan.

Start your speech when you hear the teacher’s request:

Section 2: Writing (15 minutes)

1. Completing an Application Form

Directions: Your friend needs your help with completing the Japanese application form for a study abroad program. Based on what your friend wrote down in the English form, complete the form in Japanese. Write clearly, and use hiragana, katakana, Arabic numerals, and kanji appropriately. Leave sections blank if you don’t know how to fill them.

A form your friend wrote:
<table>
<thead>
<tr>
<th>name</th>
<th>Bill Paulman</th>
<th>nationality</th>
<th>U.S.A.</th>
</tr>
</thead>
<tbody>
<tr>
<td>date of birth</td>
<td>June 27, 1974</td>
<td>age</td>
<td>22</td>
</tr>
<tr>
<td>sex</td>
<td>male</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**school currently attending**

**Fill in the application form below in Japanese:**

<table>
<thead>
<tr>
<th>名前</th>
<th>国籍</th>
</tr>
</thead>
<tbody>
<tr>
<td>生年月日</td>
<td>年齢</td>
</tr>
<tr>
<td>現在通っている学校</td>
<td>性別</td>
</tr>
</tbody>
</table>

(The forms presented above are not complete for the security purposes.)
APPENDIX B

J-SKIT VOCABULARY AND KANJI SPECIFICATIONS

Note: The vocabulary and kanji characters listed below serve as a tentative J-SKIT Vocabulary and Kanji Specifications. Approximately, the following 1,300 words are taken from Basic Vocabulary 2,000 in Nihongo kyōiku no tame no kihon goi chōsa by Kokuritsu Kokugo Kenkyūjo (1984) which overlap with at least five of the following documents: Nihongo kihon goi by Okamoto (1944), "Nihongo kyōiku ni okeru kiso gakushūgo" by Kato (1963, 1964), Practical Japanese-English Dictionary by Tamamura (1970, 1978), Ryūgakusei kyōiku no tame no kihon goi hyō by Kabashima and Yoshiida (1971), Gaikokuujin no tame no kihon goi hyō by Bunkacho Kokugoka (1971, 75), A Classified List of Basic Japanese Vocabulary by Neustupny (1977), Nihongo kyōiku goi shiryō (1) (2) -- teigakunen syokyu 500-go by Kokuritsu Kokugo Kenkyūjo (1979), and Nihonjin no chishiki kaisō ni okeru hanashi kotoba no jittai--goi hyō by Shibe (1980). This original 1,257 words are then evaluated and a few changes have been made. The final vocabulary list contains approximately 1,300 words (Source: Kokuritsu Kokugo Kenkyūjo, 1982 and 1984).

() -- kanji appeared between the parentheses are not included in the Kanji Specification
* -- vocabulary and kanji eliminated from the original 1,257 word list
/ -- indicates homonyms
underlined words -- added to the original 1,257 word list

In addition to the list, numbers (both the Japanese hito, hu, mi, ..., and the Chinese ichi, ni, san, ...), days of the month, and months of the day are included in the list.

あいさつ（挨拶） 間 合う／会う／あう（違う） 青い 赤い 上がる 明るい 秋 開く 開ける／明ける 上がる／あがる（揚がる、挙がる） 朝 浅い
朝ご飯 あさって（明後日） 明日 預ける あそこ 遊ぶ 暖かい 頭 新しい 当たる あちら 暑い/厚い/熱い 集まる 集める 後 あなた 兄姉 あの 危ない 甘い あまり/余り 雨 洗う 表す/現す 現れる/表れる ありがとう ある （或） 歩く あれ 安心 あんな

いい/よい いいえ 言う 家 下 いかが 生きる 行く いくつ いくら
いけない 意見 石 医者 上 います（椅子） 忙しい 急ぐ 痛い いた
だく 一日 一年 一年間 一番 いつ 一緒 いっしょうけんめい（一生懸命）いっぱい 一般 いつも 糸 を選ぶ なかなか 田舎 犬 今 意味
妹 いや/嫌な いらっしゃる 要る いる 入れる 色 色々

上 受ける 動かす 動く *牛 後ろ うすい（薄い） うそ（嘘） 歌 歌う うち/家/内 美しい 移る *馬 うまい（旨い、美味い） 生まれる
海 裏 売る うるさい うれしい（嬉しい） 上着 運転 運動

絵 映画 映画館 影響 英和辞典 英語 駅 選ぶ 円 えんぴつ（鉛筆）
えんりょ（遠慮）

御一 おいしい 追う 多い 大きい/大きさ 下母さん/母さん お陰様で
おかれ おきる お気をつけ 置く 奥さん/奥様 送る 退れる お元気
ですか 始ます 起こる おこる（怒る） おじいさん 教える お辞儀 押す
退い おそろしい（恐ろしい） 御大事に 落ちる おっしゃる 夫 音 お
父さん/父さん 弟 男 おととい 大人 *おとらない *踊る おどろく
（驚く） お腹 同じ おはよう 覚える お巡りさん 重い 思う おめでと
う 思い出す おもしろい（面白い）*おもちゃ 表 親 おやすみ

一階 一會 一回 海岸 外国 外国語 外国人 会社 階段 買い物 会話
買う 返す/帰す 顔 科学 かがみ（鏡） 鍵 書く 学生 かける（掛け
る） 過去 傘 重ねる 飾る かし（葉子） 火事 かしこまりました 貸す
ガス 風 風邪（風邪） 数える 家族 方 肩 固い かたか ら 形 片付
ける 勝つ 一月 がっかり 学校 家庭 角 仮名（仮名） 悲しい 必ず
金 かばん（靴） かぶる（被る） かべ（壁） 神/神様 紙 かみ（髪）
火曜 通う からい（辛い） ガラス 体 借りる 軽い 彼川 一侧 かわ
い かわく（乾く） 代わり 変わる 考える 関係 漢字 かんしゃ（感謝）
感じる 完全 簡単 がんばる（頑張る）
木 気 黄色／黄色い 消える 機械 聞く 危険 気候 聞こえる ＊汽船
技術 傷 季節 規則 北 汚い 切手 きっと 切符 昨日 気の毒 きびしい
（厳しい） 気分 決まる 義務 決める 気持ち 着物 客 急／急ぐ 九州 牛乳 今日 教育 教会 教科書 教室 きょうそう（競争） 兄弟 きょううみ（興味） 去年 きより（距離） きれい 切る 着る きれない 金 銀行
近所 金曜

具合 空気 草 菜 くださら 果物 日 くつ（靴） くつした（靴下） 国
首組 雲 畜 畜る 暗い 一ぐらい 暮らす 比べる 来る 苦しい
車 くれる くろい 加える 詳しい 一君

毛 計画 けいけん（経験） 経済 けいさつ（警察） 芸 今朝 景色 消す
結果 結局 結構 結婚 月曜 けれども 原因 げんかん（重関） 元気
研究 健康 現在 権利

子 一個 御一 語 こう公園 工業 高校 工場 交通 声 コーヒー 水
国際 ここ 午後 心 ございます 腰 故障 午前 答える ごちそうさま
でした こちら こっち コップ 事 今年 言葉 子供 この ご飯 細かい
困る 米 ごめんください ごめんなさい これ これから 一頃 ＊殺す
こわい（怖い） こわす（壊す）こわれる（壊れる）今月 今週 今度 こんな
こんにちは 今晚 こんばんは 今夜

一奨 最近 最後 最初 財布 材料 探す 魚 下がる 先 ＊咲く 桜 下
げる さしあげる ＊指す 誘う 一冊 さっし 雑誌 砂糖 さびしい －さ
ん／様 寒い さようなら／さよなら 皿 ＊猿 ＊産業 残念 散歩

字 一時 ＊試合 塩 しかし 仕方 叱る 時間 仕事 辞書 静か 自然
下舌 ＊時代 しっかりと 実際 質問 失礼 自転車 自動車 死ぬ しばらく
自分 島 しまう 閉まる 閉める 社会 写真 シャツ じゃま（邪魔）
週 自由 週間 宗教 十分／充分 授業 出発 ＊種類 準備 紹介 正月
＊商業 正午 上手 ＊状態 丈夫 しょうゆ（醤油） 将来 職業 食事
＊植物 女性 知らせる 調べる 知る 白い 一人 神社 親切 心配 新聞
＊進歩

図 水道 ずいぶん 水曜 吸う 好き 過ぎる 空く すぐ 少ない 少し

117
涼しい 進む 一ずつ すっかり ずっと 捨てる 砂 スポット すみません
住む 汚む する すわる（坐る）

生活 政治 性質 生徒 ＊制度 政府 西洋 世界 席 席位 せっかく ＊
石倉 絶対 説明 背中 ぜひ（是非） 狭い 世話 ＊選挙 ＊選手 先生
全然 戦争 全体 せんだたく（洗濯） 全部

そう 掃除 そうして／そして 相談 底 そこ そこで そちら 卒業 そっ
ち 外 その 側 空 それ それから それとも そんな／そんなに

一たい／一したい 一台 第一 大学 大丈夫 大切 ＊体操 大体 たいてい
（大抵） 台所 大変 太陽 倒れる 高い だから／ですから 一たがる た
くさん 一だけ 確か 出す 助ける ただ だいたい 正しい たたみ（畳）
一達 立つ／建つ 建物 立てる／建てる 例えば 棚 楽しい 楽しむ 頼
む たばこ（煙草） 多分 食べ物 食べる＊玉 卵 ため（為） だめ（駄
目） 足りる 誰 男性 だんだん

血 小さい／小さなもの 近い 遠い 近く 地下鉄 力 ＊地球 知識 地図 父
地方 茶 茶色 ちょうど（茶碗） ちゃんと 中 注意 ＊中央 中国 中
心 注文 調子 ちょっと ちょっと

一日 使う ＊捕まえる 疲れる 月[Moon, month] 次 付く／着く 机
作る 付ける／着ける 都合 伝える 土 続く 続ける 包む 勤める 妻
つまり つまらない 冷たい つより 梅雨 強い 連れる

手 ていねい（丁寧） テーブル 出かける 手紙 適当 出来る 出口 出来
る 鉄 手伝う では ではまた 手袋 でも 寺 出る テレビ 点 天気
電気 電車 天井 電報 電話

戸 一度 ドア どう どうか どうして どうぞ どうぞよろしく 動物 ど
うも 遠い 通る 都会 時時々 特別 時計 ＊溶ける どこ ところ と
ころが ところで 年 図書館 土地 どちら どっち とても どんな 隣
とにかく どの 飛ぶ 止まる／留まる 池まる 止める／停める／泊める 友
達 土曜 鳥 努力 取る どれ どんどん どんな

ない 内容 直す／治す 中 長い なかなか 流れる 泣く なくす なくな

118
る なおる なぜ 夏 何 何か 名前 * 波 言う 並ぶ なる なるべく
なるほど 慣れる

兄さん／お兄さん 勤い 苦い 肉 逃げる 西 一日 日曜 日本 荷物 似
る 庭 * 鶏 一人 * 人形 人間 人数

脱ぐ 盗む

* 根 姉さん／お姉さん 願う ネクタイ 猫 値段 熱 眠い 眠る 寝る
一年

の * 農業 ノート 残る のど（喉） 伸びる／延びる 登る／上る 飲み物
飲む 乗り換えの 乗る

葉 歯 場合 はい 一杯 入る 一ばかり はく（履く） 箱 運ぶ はさみ
（鉄） 喫 端 橋 始まる 始めに 始めて／初めて はじめまして 始める
場所 走る はむ バス 恥ずかしい 畑 三十歳 働く はっきり ＊ 発見
＊ 発達 ＊ 発明 花 鼻 話し 話す 離す／放す 離れる／放れる 母 幅
早い／遅い 林 払う 春 ＊ 貼る 晴れ
晴れる 晚 番 パン ハンカチ 番号 反対 半分

日 火 東 一匹 引く 低い ひげ（髭） 飛行機 久しぶり 非常に／非常に
左 ぴっくり 必要 人 一つ 一人 ひま（暇） ＊ 紐 直 冷やす 表
病院 病気 ひらがな 星 広い ＊ 広がる びんぼう（貧乏）

封筒 夫婦 増える 深い 吹く 拭く 複雑 豚 二つ 二人 普通 二日
太い 太る 船 部分 不便 ＊ 踏む 冬 降る 古い 風呂 一分 文 文化
文学

平和 ページ 下手 別／別に 部屋 減る 変 収 収ペン 変化 勉強 返事
便所 円当 便利

方 ほうえき（貿易） 方向 帽子 方法 ＊ 訪問 法律 外／他 僕 ポケット
ト ＊ 埃 星 欲しい 細い 北海道 程 ほとんど 骨 ＊ 被める 本 一本
本当

119
一枚 毎朝 毎週 毎年 毎日 毎晚 参る 前 曲がる ＊負ける まじめ
（真面目）まず まずい 貧しい すますす また まだ または 町／街 間
違う 間違える 待つ まっすぐ まったく －まで 喫 ＊繰める 間に合う
－まま 守る 迴る／回る －方

見える ＊磨く 右 短い 水 湖 店 見せる －みたいだ 道 三日 見つ
ける 三つ 緑 皆／みんな 南 耳 みやげ（土産） 見る

六日 迎える 昔 向こう 虫 難しい 息子／息子さん 結ぶ 娘／娘さん
むだ（無駄） 六つ 胸 村 ＊紫

目 めがね（眼鏡） めずらしい（珍しい）

もう 申す 燃える 目の 木曜 もし もちろん 持つ もっと 最も もの
／物／者 もらう 問題

一屋／一家 ＊八百屋 野球 焼く 約 約束 野菜 易しい やさしい（優しい） 安い やすい（易い） 休み 休む ＊疲せる やっと 屋根 やはり／やっぱり 山 止める／辞める やる やわらかい（柔らかい）

湯 夕方 郵便局 ゆうべ（昨晚） 有名 雪 ＊輸出 ゆっくり／ゆっくりと
＊輸入 指 ゆめ（夢） 許す ゆれる（揺れる）

様 八日 用事 様子 洋服 よやく よく 横 四日 四つ 予定 呼ぶ
読む 夜 よる（因る／依る／掻る） 喜ぶ よろしい よろしく よろしくお
伝えください 弱い

来週 来年 楽 －らしい ラジオ

理解 立派 理由 利用 両親 両方 料理 旅館 旅行 りんご

るす（留守）

礼 例 歴史 練習 連絡

ろうか（廃下）
若い 分かる 分かる／別れる 訳 分ける わざわざ 忘れ物 忘れる 私
渡す 渡る 笑う 割合／割合に 割る 悪い