

**A STUDY OF THE SECOND - LANGUAGE SOCIALIZATION OF
UNIVERSITY - LEVEL STUDENTS :**

A DEVELOPMENTAL PRAGMATICS PERSPECTIVE

by

SHOICHI MATSUMURA

Bachelor of Economics, Kobe University of Commerce, 1993
Master of Education, Kobe University, 1995

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE STUDIES
(Department of Language and Literacy Education)

We accept this thesis as conforming
to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

August 2000

© Shoichi Matsumura, 2000

In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of Language and Literacy Education
The University of British Columbia
Vancouver, Canada

Date August 4, 2000

ABSTRACT

The present study focused on changes over time in university-level Japanese students' sociocultural perceptions of social status during their year abroad in Canada, and the impact of such altered perceptions on their perceptions at subsequent time points. The sociocultural perception to be examined was perceived "social status" which Brown and Levinson (1987) discussed as a contributory factor in the perception of social asymmetry, power and authority. The study attempted to examine (1) whether (and to what extent) Japanese students, before they came to study in Canada, had recognized English native speakers' understanding of social status and had learned how to offer advice appropriately in English to individuals of various social statuses, (2) what proportion of differential pragmatic development among Japanese students in Canada was accounted for by their *English proficiency* and *amount of exposure to English*, and (3) whether (and to what extent) living and studying in Canada facilitated Japanese students' pragmatic development, which was assessed by the degree of approximation to native speech act behavior in various advice-giving situations repeated during the course of an academic year. To this end, the study compared the development of Japanese exchange students' pragmatic competence during their year abroad in Canada with peers in Japan who did not undertake a year abroad.

TABLE OF CONTENTS

ABSTRACT	ii
TABLE OF CONTENTS	iii
LIST OF TABLES	vi
LIST OF FIGURES	viii
ACKNOWLEDGMENTS	ix
CHAPTER I: INTRODUCTION	
1.0 Overview	1
1.1 Problem	1
1.2 Preliminary Study of Japanese Students' Pragmatic Competence	3
1.2.1 Site and Participants	3
1.2.2 Purpose	3
1.2.3 Results	4
1.2.4 Implications for the Present Study	5
1.3 Purpose of the Present Study	6
1.4 Definitions	8
1.5 Summary	10
CHAPTER II: THEORETICAL BACKGROUND TO THE PRESENT STUDY (LITERATURE REVIEW)	
2.0 Overview	11
2.1 Research on L2 Socialization	11
2.2 Proficiency Effects and Pragmatic Competence	16
2.3 Exposure to L2 and Pragmatic Competence	18
2.4 Implications in Constructing Instruments	20
2.5 Summary	23
CHAPTER III: HYPOTHESES AND RESEARCH QUESTIONS	
3.0 Overview	25
3.1 Hypotheses Tested in Study 1	25
3.2 Hypotheses Tested and Research Questions Addressed in Study 2	27
3.3 Summary	29
CHAPTER IV: METHOD	
4.0 Overview	30
4.1 Research Sites	30
4.2 Subjects	31
4.3 Data Collection Procedure	32
4.4 Instruments	34
4.4.1 A Questionnaire on Personal Information	34

4.4.2 A Questionnaire on Current Uses of English	35
4.4.2.1 Items Measuring Amount of Exposure to English	35
4.4.2.2 An Item Measuring English Proficiency	36
4.4.3 A Multiple-Choice Questionnaire to Assess Perception of Social Status	38
4.4.3.1 Constructing a Multiple-Choice Questionnaire	39
4.4.3.2 Evaluating the Stability of Native Speakers' Preferred Choices in the Multiple-Choice Questionnaire	43
4.5 Data Analysis Techniques	47
4.5.1 The Rationale for the Use of Structural Equation Modeling (SEM) with Latent Variables	47
4.5.2 Advantages of the Use of SEM with Latent Variables in Analyzing Longitudinal Data	50
4.6 Summary	54
CHAPTER V: DESCRIPTIVE STATISTICS	
5.0 Overview	55
5.1 Descriptive Statistics for the Raw Data from UBC-Rits Group	55
5.2 Descriptive Statistics for the Raw Data from Kyoto-Rits Group	57
5.3 Summary	57
CHAPTER VI: VALIDITY AND RELIABILITY	
6.0 Overview	58
6.1 Validity	58
6.1.1 A Brief Review of Classical Validity Techniques	58
6.1.2 Unstandardized Validity Coefficients	60
6.1.3 Standardized Validity Coefficients	63
6.2 Reliability	65
6.2.1 A Brief Review of Classical Reliability Techniques	65
6.2.2 Squared Multiple Correlations	66
6.3 Summary	67
CHAPTER VII: STUDY 1: MODELING THE RELATIONSHIPS AMONG PERCEPTION OF SOCIAL STATUS, ENGLISH PROFICIENCY, AND EXPOSURE TO ENGLISH	
7.0 Overview	69
7.1 Restatement of Hypotheses	69
7.2 Analyses	71
7.2.1 Evaluation of the Measurement Model	71
7.2.1.1 Selecting the Measurement Model	72
7.2.1.2 Assessing the Fit of the Selected Measurement Model	75
7.2.2 Comparison of Structural Models	77
7.3 Results	79
7.4 Summary	83

CHAPTER VIII: STUDY 2: TESTING FOR INVARIANT LATENT MEAN STRUCTURES

8.0 Overview	84
8.1 Basic Concepts Underlying Tests of Latent Means	84
8.2 Evaluating the Baseline Model	85
8.3 The Logic of the Structured Latent Means Model	88
8.4 Evaluating the Structured Latent Means Model	90
8.5 Assessing the Latent Means	94
8.6 Summary	97

CHAPTER IX: CONCLUSION

9.0 Overview	99
9.1 Summary	99
9.1.1 Purpose	99
9.1.2 Theoretical Background	100
9.1.3 Methodology	101
9.1.3.1 Subjects	101
9.1.3.2 Data Collection	101
9.1.4 Analyses	102
9.1.5 Results	103
9.1.5.1 Results of Study 1	103
9.1.5.2 Results of Study 2	103
9.2 Interpreting the Results from L2 Socialization Perspectives	104
9.3 Limitations	109

BIBLIOGRAPHY	111
--------------------	-----

APPENDIX A: Questionnaire on Japanese Students' Speech Act Behavior	123
---	-----

APPENDIX B: Questionnaire on Your Background	124
--	-----

APPENDIX C: Questionnaire on Current Uses of English	125
--	-----

APPENDIX D: A Multiple-Choice Questionnaire (Japanese Version)	127
--	-----

APPENDIX E: A Multiple-Choice Questionnaire (English Version)	132
---	-----

APPENDIX F: LISREL Input File 1	136
---------------------------------------	-----

APPENDIX G: LISREL Input File 2	138
---------------------------------------	-----

LIST OF TABLES

Table 4.1:	<i>Summary of Data Collection Procedure</i>	34
Table 4.2:	<i>Native Speakers' Preferences for Advice Type in Frequencies</i>	44
Table 4.3:	<i>The Coefficients of Stability in Native Speakers' Preferred Choices in the Multiple-Choice Questionnaire</i>	45
Table 5.1:	<i>Summary of Descriptive Statistics for the Raw Data from UBC-Rits Group</i>	55
Table 5.2:	<i>Summary of Descriptive Statistics for the Raw Data from Kyoto-Rits Group</i>	57
Table 6.1:	<i>Estimates of Unstandardized Validity Coefficients for the Measurement Model at Each Time Point</i>	61
Table 6.2:	<i>The t-value for Each Estimated Parameter in the Model</i>	62
Table 6.3:	<i>Estimates of Standardized Validity Coefficients for the Measurement Model</i>	64
Table 6.4:	<i>Estimates of the Squared Multiple Correlations for the Measurement Model</i>	67
Table 7.1:	<i>Summary of Tests for Invariance of Factor Loadings</i>	72
Table 7.2:	<i>Summary of Tests for Invariance of Measurement Errors</i>	74
Table 7.3:	<i>Summary of Selected Goodness-of-Fit Indices for Model 2</i>	75
Table 8.1:	<i>Summary of Selected Goodness-of-Fit indices for the Hypothesized Model</i>	87
Table 8.2:	<i>Summary of Selected Goodness-of-Fit Indices for the Structured Means Model</i>	90
Table 8.3:	<i>Summary of Selected Goodness-of-Fit Indices for the Less Constrained Structured Means Model</i>	91
Table 8.4:	<i>Summary of Selected Goodness-of-Fit Indices for the Structured Means Model Involving the Variant Factor Loadings of CJ's and XL's</i>	92

Table 8.5:	<i>Summary of Selected Goodness-of-Fit Indices for the Structured Means Model Involving the Invariant Factor Loadings of CJ's and XL's at Time 1</i>	92
Table 8.6:	<i>Summary of Selected Goodness-of-Fit Indices for the Structured Means Model Involving the Invariant Factor Loadings of CJ's and XL's at Time 1 and Time 2</i>	93
Table 8.7:	<i>Summary of Estimates of the Kappa Values in the Final Structured Latent Means Model</i>	95
Table 8.8:	<i>Parameter Estimates for UBC-Rits and Kyoto-Rtis Groups</i>	96

LIST OF FIGURES

<i>Figure 3.1:</i>	Theoretical model of the relationships among perception of social status (POSS), English proficiency (PROF), and exposure to English (EXPO).....	26
<i>Figure 3.2:</i>	Theoretical model of change in perception of social status (POSS) based on a four-wave longitudinal design.....	28
<i>Figure 4.1:</i>	A theoretical construct, EXPO (exposure to English) and its two measures (amount of exposure through productive and receptive uses of English).....	36
<i>Figure 4.2:</i>	A theoretical construct, PROF (English proficiency) and its three measures (scores in sections 1, 2, and 3).....	38
<i>Figure 4.3:</i>	A theoretical construct, POSS (perception of social status) and its three measures (sum of scores in scenarios for higher, equal and lower statuses).....	46
<i>Figure 4.4:</i>	Example of a path model of POSS (perception of social status), PROF (English proficiency), and EXPO (exposure to English) on a four-wave longitudinal design.....	49
<i>Figure 4.5:</i>	A complete latent variable model based on a four-wave longitudinal design.....	51
<i>Figure 7.1:</i>	Standardized parameters representing the cross-time relationships among the latent variables in the final model.....	80
<i>Figure 8.1:</i>	Baseline model of change over time in POSS (perception of social status).....	87
<i>Figure 8.2:</i>	Latent means structural model.....	89

ACKNOWLEDGMENTS

I have had the good fortune during my graduate school life in Canada to have met several teachers and friends who have inspired me, whose influence on me has been incalculable. I am very grateful to:

Dr. Lee Gunderson, my supervisor for his sincere and continued support and encouragement since the initial stage of my graduate study;

Dr. Nand Kishor for helping me think harder about methodology;

Dr. Richard Berwick for his help with the design of this study in its earliest stage;

Dr. Elizabeth Lee for her valuable comments and suggestions;

Dr. Robert Conry who did his best to give me a comprehensive exam in spite of his serious illness;

Dr. Katsuhiro Ohashi who helped me arrange the Japanese portion of this research;

instructors and staffs at the UBC-Ritsumeikan academic exchange program, Bill McMichael, Kathy Bell, Lynne McGivern, George Hann, Jean Hamilton, Sheri Wenman, and Joe Greenholtz for giving me the opportunity to conduct the research;

Ritsumeikan University students who agreed to participate as subjects;

Mitsunori Takakuwa who has helped me in many respects during the entire period of my stay in Canada and who has given me a painstaking critique from which this dissertation greatly profited;

and last, but not least, to my parents and wife for putting up with my excuses for too long.

To all these folks, I would like to offer my deepest gratitude.

CHAPTER I

INTRODUCTION

1.0 Overview

The present study is described in this chapter and several important concepts used throughout this dissertation are developed. This chapter begins with the statement of the problem. Next, the purpose of the present study is discussed. Finally, several key theoretical terms used in the study are defined.

1.1 Problem

Since 1993, the year when *The Course of Study* (The National Guideline for Education in Japan) was revised and new textbooks supporting new national foreign language guidelines were initiated for use in secondary schools, the goals of English language teaching in Japan have been to enhance students' pragmatic competence, that is, the ability to interpret and use language appropriately in social contexts. Classes that focused exclusively on developing students' grammatical knowledge were replaced by communication-based classes to which a native speaker of English, an assistant language teacher (ALT),¹ was assigned. The purpose was for students to learn English communicative functions, including, for example, how to make requests in English in particular social contexts. The Ministry of Foreign Affairs of Japan (1998) described the

¹ ALTs are native speakers of English appointed to schools in Japan as part of the Japan Exchange and Teaching (JET) Program by the Council of Local Authorities for International Relations.

reasons for the introduction of such communication-based classes to secondary schools as follows:

Today, there is a great amount of international exchange among people, things and information in all sorts of fields. The importance of understanding each other through direct communication is growing enormously. Therefore, ALTs are expected to play a significant role in promoting communicative teaching and introducing foreign culture in the classroom, thereby helping to develop an educational programme in Japan based upon international understanding. [on-line: available at http://www.mofa.go.jp/j_info/visit/jet/experience.html]

These comments suggest that communication-based classes provide students with the opportunity not only to interact directly with native speakers of English, but to learn the target sociocultural rules necessary to acquire pragmatic competence. However, the problem is that there are few empirical studies that have investigated whether (and to what extent) students learn to use pragmatic knowledge efficiently, and whether the pragmatic competence they acquire in school reaches the level to allow them to function competently with members of a target speech community. Moreover, few studies have investigated whether their pragmatic competence continues to develop or diminishes over time after students graduate from secondary schools.

Indeed, several questions are posed here. What learning environment is necessary to maintain the level of pragmatic competence that students acquire in school? Is living and studying in an English-as-a-second-language (ESL) environment, that is, a target speech community, more effective in maintaining or extending the level of pragmatic competence than studying in an English-as-a-foreign-language (EFL) environment like

Japan? Communication-based classes have been implemented in Japan for about six years without consideration of these issues.

1.2 Preliminary Study of Japanese Students' Pragmatic Competence

1.2.1 Site and Participants

In order to explore these issues, Japanese students who had learned English in communication-based classes led by native speakers of English in secondary schools, and who had the opportunity to stay in the target speech community were selected. They were involved in an eight-month academic exchange program at the University of British Columbia (UBC) in Canada (the UBC-Ritsumeikan Academic Exchange Program). In this program about 100 university-level Japanese students come to study in the target speech community, Canada, every year, and live with their English-speaking roommates at an on-campus facility called UBC-Rits House. Fifteen instructors of the exchange program and 32 English-speaking roommates of Japanese students volunteered to participate in the study.

1.2.2 Purpose

The purpose of the preliminary study conducted in April, 1997 was to gather information on characteristics of Japanese students' pragmatic uses of English in interactions with members of the target speech community. As Richards, Platt and Platt (1992) note, pragmatics include the study of how the interpretation and use of utterances depends on knowledge of the real world, how speakers use and understand speech acts, and how the structure of sentences is influenced by the relationship between the speaker

and the hearer. Therefore, a questionnaire was designed in an open-ended format to collect information on these issues (see Appendix A).

1.2.3 Results

The results showed that many instructors thought that Japanese students frequently used direct speech acts in giving advice and suggestions at the end of the eight-month program--even though indirect speech would have been more appropriate in specific speech settings. For example, one of the instructors commented that many students used such direct speech acts as "You must Verb Phrase (hereafter, VP)" and "You should VP" in response to an instructor's questions such as "Please tell me what I could do in order to make this class more interesting to you all." The remarks of the instructors showed that they frequently considered such speech act behavior by the Japanese students to be impolite.

Some academic exchange program instructors commented in the questionnaire that Japanese students often did not know English polite expressions. Others stated that they did not notice that Japanese polite expressions did not necessarily convey the same degree of politeness when literally translated into English. Moreover, there was a comment that the Japanese students' eight-month residence in Canada was so short that they could not learn how to use polite expressions in context in the same way as native English speakers. Eventually, two variables were identified that seemed to explain why some Japanese students cannot give advice in a native-like manner even after an eight-month stay in the target speech community. They developed following a review of

comments from the instructors: that is, *English proficiency* and *amount of exposure to English*.

1.2.4 Implications for the Present Study

The results of the preliminary study appeared to contradict a widely-accepted notion in the field of first language (Japanese) acquisition--that Japanese communicative style is indirect (e.g., Barnlund, 1975; Clancy, 1986; Doi, 1973, 1974). Clancy (1986) noted that an indirect, somewhat depersonalized mode of expression is highly valued in many contexts in Japanese society. Moreover, Clancy (1986) found that Japanese mothers simplify the acquisition of this communicative style by following children's inappropriate direct utterances with more appropriate indirect phrases. Furthermore, several researchers (e.g., Nakane, 1967; Matsumoto, 1988, 1989) have pointed out that politeness strategies are largely influenced by social status variables in Japan's hierarchical society. For example, rarely do Japanese university students use direct expressions like "You must VP" or "You should VP" when offering advice to a higher-status individual such as a professor.

Why then, do many Japanese students in the program frequently use direct speech acts like "You should VP", when it comes to offering advice in English? In other words, why do they have difficulty offering advice in the forms necessary to signal the socially expected degree of politeness? Is their difficulty related to *English proficiency* and *amount of exposure to English* during their eight-month stay as suggested by several instructors? The findings of this preliminary study warranted further investigation of

Japanese students' politeness strategies. The relationship of their politeness strategies with their *English proficiency* and *amount of exposure to English* should be explored.

1.3 Purpose of the Present Study

The present study was prompted by comments of several native speakers of English who were instructors of an academic exchange program at UBC, and the largely unexamined comparison of pragmatic competence acquired in and outside of Japan. The present study proposed to operationalize several tasks on the basis of two theoretical frameworks, namely, language socialization and interlanguage pragmatics.² Second-language (L2) socialization refers to the process by which individuals, whether children or adults, "acquire tacit knowledge of principles of social order and systems of belief" (Schieffelin & Ochs, 1986b, p. 2) through exposure to and participation in L2-mediated interactions. Interlanguage pragmatics refers to "nonnative speakers' use of pragmatic knowledge" (Kasper & Schmidt, 1996, p. 149). The basic concept underlying these two frameworks is that the development of pragmatic competence is a process of social development (Ninio & Snow, 1996). The present study focused on changes over time in university-level Japanese students' sociocultural perceptions and the impact of such altered perceptions on their pragmatic uses of English when giving advice. The sociocultural perception examined in the study was perceived "social status" which Brown and Levinson (1987) discussed as a contributory factor in the perception of social asymmetry, power and authority.

² A review of the literature relevant to language socialization and interlanguage pragmatics is provided in Chapter II.

The first and second tasks operationalized in the study dealt with research into second-language (L2) socialization by learners in English-as-a-foreign-language (EFL) and English-as-a-second-language (ESL) contexts, respectively. For the first task, the study examined whether (and to what extent) Japanese students in an academic exchange program, before they came to study in Canada, had learned target sociocultural rules of offering advice through communication-based classes in school. The second task was to examine whether (and to what extent) living and studying in the target speech community facilitated Japanese students' pragmatic development, which was assessed by the degree of approximation to native speech act behavior in various advice-giving situations repeated during the course of an academic year. In order to examine the impact of living and studying in the target speech community, the pragmatic development of the Japanese students in the target speech community and of those who continued to stay and study in Japan was compared. The third task was to account for differential pragmatic development among Japanese students in the target speech community as functions of their *English proficiency* and *amount of exposure to English*.

Thus the present study attempted to learn how Japanese students' instructional and life experiences supported the development of pragmatic competence in use of English, and to examine in particular the differences in competence that accrued from experience in an English speaking culture. The study attempted to account for students' acquisition of the competence to offer advice and to compare different levels of competence that resulted from study in Japan and in Canada. In so doing, it was hoped that the findings of the study would contribute to clarification of the L2 socialization process from a developmental pragmatics perspective.

1.4 Definitions

Several key theoretical terms used throughout this dissertation are defined here.

Second-language (L2) socialization:

As mentioned earlier, L2 socialization refers to the process in which individuals, whether children or adults, “acquire tacit knowledge of principles of social order and systems of belief” (Schieffelin & Ochs, 1986b, p. 2) through exposure to and participation in L2-mediated interactions. It should be noted that the term L2 socialization is not always interchangeable with the term secondary or adult socialization, because the concept of secondary or adult socialization includes socialization through and to use an L1. However, secondary socialization and adult socialization have been used interchangeably in the literature (see Wentworth, 1980, for a detailed discussion).

Diachronic socialization:

By this it is meant that socialization is a long-term, developmental process of acquisition of language and culture. Diachronic socialization research explores the social past as well as the social present (see Heath, 1982, for a detailed discussion). Specifically, a diachronic dimension in L2 socialization research is not restricted to observations in the target speech community. Rather, it extends to an examination of L2 learners’ native culture that may affect the L2 socialization process in the target speech community (Matsumura & Takakuwa, 1999).

Synchronic socialization:

By this it is meant that the organization of socializing contexts is sometimes temporary but any temporal context is seen as constructing one aspect of diachronic socialization (Matsumura & Takakuwa, 1999). In L2 socializing contexts, L2 learners have as a goal solving specific problems of interaction at hand by interpreting what social activity is going on and acting/reacting in socially and culturally sensitive and appropriate ways through the use of L2.

Synchronic L2 socialization can be examined for how L2 learners are socialized through socially and culturally organized activities into “expected ways of thinking, feeling, and acting” (Becker et al., 1961; Wentworth, 1980, cited in Ochs, 1986, p. 2).

It is important to note here that diachronic and synchronic socialization are not mutually exclusive, nor is one the prerequisite for the other, because human beings, once they are born, are diachronically socialized through countless, various synchronic socialization events and also because the cultural norms and values that they have acquired diachronically at a particular point in time will affect their future synchronic socialization.

Social status:

One’s recognized positions and roles in society and/or in a particular social situation (Brown & Levinson, 1987). When members of a society interact with one another, their linguistic behaviors are influenced by their conceptions of their own and others’ social status. The present study assumes that sociolinguistic

knowledge of social status is related to the pragmatic competence to use direct and indirect speech acts.

Pragmatic competence:

“[A] variety of abilities concerned with the use and interpretation of language in contexts. It includes speakers’ ability to use language for different purposes--to request, to instruct, to effect changes” (Bialystok, 1993, p. 43). The present study examines pragmatic competence to give advice in English to higher-status, the same status, and lower-status persons.

Pragmatic development:

This refers to the approximation over time to native speech act behavior in various social contexts. The present study will examine the acquisition of rules of politeness and culturally determined rules for offering advice in English to higher-status, the same status, and lower-status persons.

1.5 Summary

This chapter has provided an outline of the present study and has introduced two theoretical issues, that is, language socialization and interlanguage pragmatics. Following the identification of the problem, the need and importance of conducting the study were discussed. Next, the purpose of the study, and the definitions of the key theoretical terms were presented. The next chapter is devoted to addressing the theoretical background of the present study while reviewing the literature related to language socialization and interlanguage pragmatics.

CHAPTER II

THEORETICAL BACKGROUND TO THE PRESENT STUDY (LITERATURE REVIEW)

2.0 Overview

The aim of this chapter is to locate the present study in the fields of language socialization and interlanguage pragmatics and to discuss the significance of the study within these fields. This chapter begins by identifying several methodological problems in previous L2 socialization studies, followed by a review of the literature that examines the relationship of *English proficiency* and *amount of exposure to English* with English pragmatic uses and interpretations. Finally, the implications in constructing the instruments in the present study are discussed.

2.1 Research on L2 Socialization

It has been over twenty years since Hymes (1971, 1972a, 1972b) and Campbell and Wales (1970) proposed the view that language learning is a social and contextual process. There have been a considerable number of theories developed to account for the interplay of language and culture in both first language (L1) (Ochs, 1988; Ochs & Schieffelin, 1979; Schieffelin & Ochs, 1986a) and second language (L2) acquisition (e.g., Halliday & Hasan, 1985; Schumann, 1978). Schieffelin and Ochs' language socialization model, among others, has been applied to various English-as-a-second-language (ESL) contexts in recent years, relating the developmental nature of L2 acquisition to sociocultural competence that L2 learners acquire over time in a target speech community

(Atkinson & Ramanathan, 1995; Crago, 1992; Crago, Annahatak, & Ningiuruvik, 1993; Harklau, 1994; Poole, 1992; Willett, 1995).

One central notion of language socialization theory is that children and other novices in society learn to function competently with members of that society by organizing and reorganizing sociocultural information that is conveyed through the form and content of actions of others (Schieffelin & Ochs, 1986a, 1986b). This theoretical framework views the acquisition of linguistic competence and sociocultural competence as interdependent. Schieffelin and Ochs (1986a) state that as children learn to become competent members of their society they also learn to become competent speakers of their language. Acquiring pragmatic competence, that is, the ability to use and interpret language appropriately in contexts, is an essential part of the language socialization process, because without pragmatic competence it is hard to participate in ordinary social life within a variety of social contexts.

The study of L2 socialization from a developmental pragmatics perspective requires different approaches from the one employed to study children undergoing socialization through L1. This is primarily because L2 learners have formed their cultural norms and values through L1 and have acquired linguistic and sociocultural competence in their L1. Indeed, in addition to such L1 competence, they have also obtained some degree of L2 linguistic and sociocultural knowledge through school education and media exposure in their home countries (see Ely & Gleason, 1995 for a summary of this line of inquiry). Unfortunately, however, because of methodological limitations discussed below, few studies have investigated such L2 socialization processes in detail.

A first methodological problem is that previous studies have looked solely at synchronic socialization observed in the target speech community without incorporating a diachronic perspective into its interpretation. Specifically, few studies have been designed to examine the extent to which L2 learners have acquired pragmatic competence before they enter the target speech community. Indeed, there is a need to examine international students' educational and cultural backgrounds, because they are the important components of an explanation for individual differences in L2 socialization in the target speech community. For example, few studies have examined whether the development of pragmatic competence in the target speech community varies according to the amount of prior pragmatic knowledge that L2 learners have already obtained in their home countries. A review of the literature suggests that the typical longitudinal designs that have been employed in L2 socialization studies involve a researcher starting to observe L2 learners some time after their arrival in a target speech community (e.g., Poole, 1992; Schecter & Bayley, 1997; Willett, 1995). Under such circumstances, it is difficult to confirm what L2 sociocultural competence they may have acquired in their home countries before their arrival. Moreover, given that what L2 learners think, feel, and act in the present may be connected to their past experiences, observing L2 socialization in the target speech community alone may not adequately describe why it happens. Thus socialization in the target speech community should be accounted for and corroborated by a careful examination of what they have acquired in their home countries.³

³ The feasibility of this modified longitudinal design may be questionable in some areas of L2 socialization research (e.g., research on immigrant children's L2 socialization), because researchers may not be able to obtain information on who is coming and when.

A second methodological problem with previous L2 socialization studies is the adoption of a taken-for-granted view of culture as the basis for interpretation and explanation of L2 learner's cultures of origin and the target culture. Such comments of research findings as "the video activity represents a typical White middle-class American accommodation context" (Poole, 1992, p. 604) indicate the adoption of such a view. Atkinson and Ramanathan (1995) and Poole (1992) described the adoption as "a necessary convenience" (p. 557) and "a convenient point of comparison" (p. 599) in their own research, respectively. In other disciplines, however, such views have been criticized, because they promote a monolithic, static, and exoticized image of culture (e.g., Kubota, 1999; Raimes & Zamel, 1997; Spack, 1997; Susser, 1998; Zamel, 1997). L2 socialization research cannot escape from this criticism either. Specifically, the adoption of a taken-for-granted view of culture ignores the dynamic link between language and culture--that the system of social and cultural structures changes over time and accordingly, what language practices are socially and culturally appropriate or expected by members of the social group also change over time. If researchers seek to find evidence of L2 socialization drawing on the cultural views articulated by research conducted in the distant past, say, ten years ago, they may misinterpret in what direction L2 learners are socialized.

A third methodological problem is that few L2 socialization studies have employed an adequate number of subjects to examine 'intracultural variance,' that is, the extent to which the subjects under study are typical and atypical of their first cultures, and under the particular influence of their cultural backgrounds. As a result, findings linked to the L2 socialization process are unable to be generalized to a population of individuals

who share the same culture and language. The lack of an examination of intracultural variance causes such problems as cultural stereotypes and unsubstantiated generalizations. It should be noted that any L2 learners under study are not necessarily representative of the cultures that they are supposed to represent. It should also be noted that the cultural norms and values in focus are not necessarily familiar to those who are born and raised in that culture, nor are they shared to the same degree among the people of the culture.

A fourth problem is that few studies have been concerned with L2 socialization that takes place in their home countries. Because of a lack of a comparison group that consists of L2 learners who continue to stay in their home country, previous studies do not answer the question as to how L2 socialization in a target speech community differs from that in the home country. There have been no efforts to examine whether there is a difference in the route and rate of pragmatic development between L2 learners in the target speech community and in their home country.

Because of these methodological problems, previous studies have revealed little about the characteristics of the L2 socialization process in a target speech community and an L2 learner's country of origin. The present study, however, was designed to solve these problems. Specifically, by employing two large groups (one in the target speech community, Canada and the other in the subjects' home country, Japan), the study attempted to clarify the impact of L2 learning environment on pragmatic development, that is, how the route and rate of pragmatic development differed between an ESL group and an EFL group. Moreover, by starting data collection before the first group entered Canada, the study attempted to obtain information on the extent to which participants in the first group had already obtained L2 pragmatic knowledge in Japan. Furthermore, in

order to account for individual differences in route and rate of pragmatic development in the target speech community, the present study examined their increasing approximation to native speech act behavior in various advice-giving situations, as functions of *L2 proficiency* and *amount of exposure to L2 in the target speech community*.

2.2 Proficiency Effects and Pragmatic Competence

A number of interlanguage pragmatics studies have examined the use of speech act realization strategies by learners at different proficiency levels (Blum-Kulka & Olshtain, 1986; Maeshiba, Yoshinaga, Kasper & Ross, 1996; Olshtain & Blum-Kulka, 1985; Robinson, 1992; Takahashi & Beebe, 1987; Takahashi & DuFon, 1989; Trosborg, 1987). However, most of these studies were cross-sectional and, therefore, did not reveal developmental changes in pragmatic competence relating to L2 proficiency. Kasper and Schmidt (1996) noted, "Unlike other areas of second language study, which are primarily concerned with acquisitional patterns of interlanguage knowledge over time, the great majority of studies in interlanguage pragmatics has not been developmental" (p. 150). Nonetheless, findings of previous studies are informative in designing the present study. The remainder of this section is dedicated to reviewing several studies that employed Japanese learners of English.

Takahashi and Beebe (1987) examined in discourse completion tasks the refusal strategies used by Japanese students learning English as a second language (ESL) and English as a foreign language (EFL). They found that high-proficiency learners, that is, ESL learners in their study, often used a typically Japanese formal tone when performing refusals in English, whereas low-proficiency learners, that is, EFL learners, could not do

so due to limited L2 proficiency. Takahashi and Beebe (1993) noted that more proficient learners had enough control over L2 to express their intentions at the pragmatic level and accordingly, they were more likely to transfer L1 sociocultural norms to L2 and made pragmatic errors.

Results contradicting Takahashi and Beebe's (1987, 1993) view were also provided by a number of studies. Takahashi and DuFon (1989) investigated the request strategies by Japanese learners of English in open-ended role play. They found that beginning-level learners displayed preference for indirect speech acts by using Japanese hinting strategies, whereas the advanced learners formulated their speech more efficiently by making more direct and native-like requests. They pointed out that perceptions of request strategies differed between the low- and high-proficiency subjects. Robinson (1992) examined the refusal strategies used by Japanese learners of English in discourse completion tasks, and found that the low- and high-proficiency subjects were both aware of the differences in appropriate American and Japanese refusal behaviors. However, the lower proficiency subjects were more influenced by their L1 refusal style, whereas the higher proficiency learners' strategies were more similar to native speakers'. Moreover, Maeshiba, Yoshinaga, Kasper and Ross (1996) examined the apology strategies used by Japanese learners of English at two proficiency levels, intermediate and advanced, in discourse completion tasks. They found that the intermediate-level learners were more likely to use L1 apology strategies than the advanced learners. In contrast to Takahashi and Beebe's (1987, 1993) view, the results of these studies indicate that with increasing L2 proficiency, the subjects' pragmatic uses of L2 approximated to native speakers'.

Furthermore, Takahashi (1996) examined the relationship of Japanese EFL learners' request strategies and their L2 proficiency by means of a judgment test of speech act behavior. The results of her study showed that both low- and high-proficiency learners relied equally on their L1 request conventions or strategies in L2 request performance. The author concluded, unlike other researchers, that the false projection of L1 form-function mappings onto L2 contexts did not seem to be a function of the learners' L2 proficiency.

Inspection of the results of these studies suggests that proficiency effects on L2 pragmatic competence vary depending on speech act type (e.g., request, apology, refusal) and processing modes (e.g., perception versus production). The present study investigated proficiency effects on the subjects' perception of advice.

2.3 Exposure to L2 and Pragmatic Competence

As stated in the previous section, the developmental change in pragmatic competence has been rarely addressed in previous studies of interlanguage pragmatics. Although there is a study that examined the subjects' developing pragmatic competence as a function of the length of stay in the target speech community (Olshtain & Blum-Kulka, 1985), no studies of interlanguage pragmatics have investigated the development of pragmatic competence as a function of exposure to L2. Perhaps, exposure to L2 might be a better indicator of the L2 learners' pragmatic development than length of stay in the target speech community.

The primary purpose of the Olshtain and Blum-Kulka's (1985) study was to investigate the acculturation of learners to the target speech community by examining the

degree of approximation of their speech act behavior. Olshtain and Blum-Kulka maintained that “speech act behavior serves as a useful indicator of acculturation related to length of stay in the target community” (p. 304). Specifically, they examined the perception of politeness in requests and apologies by nonnative speakers of Hebrew. They developed a judgment test consisting of eight items: four request situations and four apology situations. Focus was given to the receptive rather than productive aspect of pragmatic competence. They found that the response patterns of L2 learners to the judgment test changed over time as a function of the learners’ length of stay in the target speech community. They concluded that irrespective of the level of linguistic competence, learners may reach native-like speech-act acceptability patterns as a function of the length of stay in the target speech community.

Olshtain and Blum-Kulka’s (1985) study appears to have revealed developmental patterns in learners’ acquisition of pragmatic competence. However, the findings of this study should be interpreted with caution due to methodological limitations. The problem is that Olshtain and Blum-Kulka assigned the subjects to several groups depending on their lengths of time in the target speech community and compared the groups with respect to this pragmatic development. Because focus was given to the developmental change in pragmatic competence, the study should have used a longitudinal design in which data are collected from the same subjects on the same instruments on several occasions.

Moreover, Olshtain and Blum-Kulka’s definition of the variable, *length of stay in the target speech community*, itself, is unclear. It was assumed in the study that the longer L2 learners have stayed in a target speech community, the more exposure to L2 they

receive. However, their assumption is not always a reflection of reality. The Japanese students in the exchange program at UBC, for example, vary in their exposure to English in Canada. They have a chance to communicate in Japanese with their friends, read Japanese books and newspapers, watch Japanese community TV, and so on, even when they stay in Canada. One former student in the program commented, "We communicate with each other in Japanese once we step out of the classroom" (personal communication, 1997). On the other hand, several students in the exchange program actively participated in social events, communicated in English even with their Japanese peers, and preferred to read English over Japanese materials, during the entire period of the program. It is highly likely that a similar situation applied to the subjects in Olshtain and Blum-Kulka's (1985) study. Thus, in addition to *L2 proficiency*, the present study included not *length of stay in the target speech community* but *exposure to L2* to account for learners' developing pragmatic competence.

2.4 Implications in Constructing Instruments

This section focuses on two studies, Takahashi (1996) and Rose (1994), in which the low validity of instruments used makes the research findings questionable. McMillan and Schumacher (1993) stated, "Validity is a situation-specific concept: validity is assessed depending on the purpose, population, and environmental characteristics in which measurement takes place" (p. 223). This is an important point to explore in L2 studies.

Takahashi (1996) examined the transferability of Japanese indirect request strategies when Japanese learners of English make English requests in corresponding L2

contexts. Takahashi constructed a questionnaire consisting of two sections. The first section was comprised of four situations described in English, in which the degree of imposition differed (two high and two low imposition situations). The first section aimed to examine subjects' perception of the contextual appropriateness of five Japanese request expressions for each situation. In the second section, the same situations used in the first section were presented with five pairs of Japanese and English request expressions. The second section aimed to examine the equivalence of perception between Japanese request strategies and the corresponding English equivalents in terms of contextual appropriateness. For example, the subjects were asked to rate on a 7-point scale the extent to which the English request expression "I would like you to VP" is equivalent to the Japanese request expression "*V-te itadaki-tai-n-desu-kedo*" under the condition in which the subjects put themselves in a situation in which they requested their Japanese professor to do something on their Japanese university campus.

However, the validity of the instruments used in this study is low for several reasons. First of all, the equivalence of Japanese-English pairs is questionable. By Takahashi's definition, the Japanese request forms "*V-te itadaki-tai-n-desu-kedo*" and "*V-te hoshii-n-desu-kedo*", for example, are equivalent to "I would like you to VP" and "I want you to VP", respectively. However, if the subjects read "*V-te hoshii-n-desu-kedo*" in rising and soft intonation, "I would like you to VP" is more equivalent than "I want you to VP." In addition, the phrase "*desu*" in "*V-te hoshii-n-desu-kedo*" is polite enough to translate the whole expression into "I would like you to VP." Thus presenting the Japanese-English equivalence judgment test only as written forms is problematic because

the subject's judgment depends largely on the intonation and tone of the reader's voice (see Wierzbicka, 1991).

The second problem is related to selection of the subjects. The subjects were 142 university-level Japanese students in Tokyo whose mean length of residence in English-speaking countries was 1.2 months. The chances are high that they did not know expressions such as "Would it be possible (for you/me) to VP?" and rated it as totally inappropriate. Moreover, it is unusual for the students in Japanese universities to make English requests to their Japanese professor on their Japanese university campus. Thus Takahashi's (1996) findings are suspect because of the low validity of the instruments used.

Rose (1994) examined the validity of a discourse completion test and a multiple-choice questionnaire when collecting speech act data in non-Western contexts. Subjects were Japanese university-level students as an experimental group and American university students as a reference group. The two instruments were prepared in both English and Japanese. Each group worked on the two instruments in their first language. Based on a review of the literature relevant to Japanese interactions, Rose hypothesized that Japanese subjects were less direct in making requests than Americans on both instruments. Contrary to expectations, he found that Japanese subjects were more direct in the discourse completion test (DCT) but less direct in multiple-choice questionnaires than Americans. He suspected that the DCT may be inappropriate for collecting data on Japanese subjects.

Rose used translation in coding data collected in Japanese, but he, like Takahashi (1996), ignored the fact that translation from Japanese to English may differ depending

on the intonation and tone of a reader's voice. For example, he translated "*Na oshiete. Onegai*" into "Hey, teach me, please!", but with a gentle tone of voice that Japanese request expression could be translated into "Um, could you teach me, please."

The second problem that makes the findings questionable is his assumption that both instruments provided information about subjects' production of requests in face-to-face interactions. However, the distinction must be made between the two instruments in terms of the types of elicited responses: that is, DCTs are classified as constrained production instruments, whereas multiple-choice questionnaires provide information about subjects' perception of alternative speech act realizations or about the pragmatic meaning subjects assign to offered stimulus material (Kasper & Dahl, 1991). Given this classification of instruments, Rose collected data on two different types of processing modes in the realization of requests. Thus it is not surprising that he did not obtain the same results from the two instruments, including, for example, that the Japanese subjects produced direct speech acts more frequently than indirect requests. Rose's study, like Takahashi's (1996), has methodological problems, but the findings warrant further investigation.

2.5 Summary

This chapter reviewed a number of methodological problems in L2 socialization and interlanguage pragmatics studies. It was argued that unique characteristics of L2 socialization emerge from the fact that L2 learners, at least to some extent, have formed their cultural norms and values through L1 and have already acquired linguistic and sociocultural competence in their L1, and therefore, that the study of L2 socialization

requires different approaches from the one employed to study children undergoing socialization through an L1. It also argued that examining diachronic socialization and intracultural variance is critical, especially when L2 learners are the focus of the study. Specifically, a study has to be designed to examine 1) the extent to which L2 learners have acquired the target sociocultural competence before they enter the target speech community with a special focus on L2 learners' educational and cultural backgrounds, and 2) the extent to which the L2 learners under study are typical and atypical of the culture, and under the particular influence of their cultural background. These are important components of an explanation for the L2 socialization process not only in the target speech community but also in the L2 learners' country of origin.

Next, it was suggested that the present study would make a unique contribution to interlanguage pragmatics because it focused on the developmental aspects of L2 learners' pragmatic competence. The rationale for the inclusion of the two variables, *English proficiency* and *amount of exposure to English* was addressed. Finally, the implications of constructing valid instruments to measure L2 learners' pragmatic competence were discussed. The next chapter addresses several hypotheses tested in the present study.

CHAPTER III

HYPOTHESES AND RESEARCH QUESTIONS

3.0 Overview

In this chapter hypotheses to be tested and research questions addressed in the present study are listed accompanied by a schematic representation of underlying theoretical models. As discussed in earlier chapters, the study attempted to account for change over time in Japanese students' perception of social status when giving advice in English, as functions of their English proficiency and amount of exposure to English. Moreover, the study aimed to compare the different levels of pragmatic development that resulted from study in Japan and Canada. Because different analytic strategies were used, these two tasks are discussed in separate chapters of this dissertation. The former is investigated in Study 1 in Chapter VII where focus was given to Japanese students who came to study in the target speech community, whereas the latter is examined in Study 2 in Chapter VIII where two groups--the Japanese students in the target speech community and those who continued to stay and study in Japan--were compared. Thus hypotheses and research questions are summarized separately for each study.

3.1 Hypotheses Tested in Study 1

First, a theoretical model underlying hypotheses to be tested in Study 1 is presented. Figure 3.1 below represents the hypothesized relationships among three constructs--perception of social status when giving advice in English, English proficiency, and exposure to English (these three constructs are denoted as POSS, PROF, and EXPO

in the figures and tables used throughout the rest of this dissertation). It should be noted that the relationships in Figure 3.1 were developed based on the results of the preliminary research discussed in Chapter I and the findings of previous studies reviewed in Chapter II.

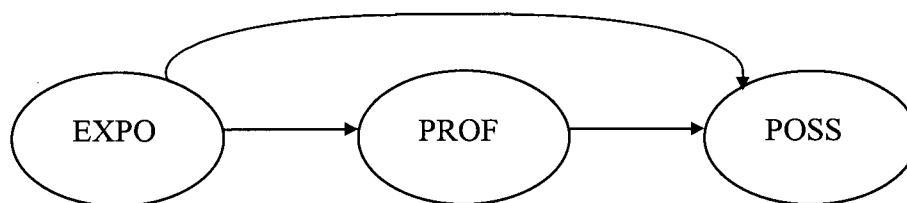


Figure 3.1 Theoretical model of the relationships among perception of social status (POSS), English proficiency (PROF), and exposure to English (EXPO).

The critical feature of Figure 3.1 is that PROF is hypothesized to have direct effects on POSS, whereas EXPO is hypothesized to have direct effects on POSS and indirect effects on POSS through its impact on PROF. In other words, PROF functions as an intervening variable in this hypothesized model. The number of intervening variables plays a critical role in estimating direct and indirect effects in a longitudinal design. Gollob and Reichardt (1991) suggest that when the design is longitudinal, testing a model that includes an indirect effect with k intervening variables requires a model that has at least $k + 2$ time points. As shown in Figure 3.1, PROF is the only intervening variable in Study 1 and therefore, it is necessary to design a minimum of three-wave longitudinal design in which data on all three variables are collected on three occasions from the same sample. As discussed in detail in the next chapter, Study 1 was designed to conduct a

four-wave longitudinal test of the model in Figure 3.1. Hypotheses tested in Study 1 were stated as follows:

Hypotheses 1:

The change over time in Japanese students' perception of social status when giving advice in English is a consequence of the increase of their English proficiency.

Hypothesis 2:

The change over time in Japanese students' perception of social status when giving advice in English is a consequence of the increase of their amount of exposure to English.

Hypothesis 3:

The change over time in Japanese students' perception of social status when giving advice in English is a consequence of the increase of their amount of exposure to English mediated by the increase of English proficiency.

3.2 Hypotheses Tested and Research Questions Addressed in Study 2

Because Study 2 focused on multi-group analyses concerning change over time in Japanese students' perception of social status, the underlying theoretical model can be depicted as in Figure 3.2.

It should be noted that focus is given to factor correlations rather than cause-effect relationships between the same latent variables across time so that factors are linked to each other by double-headed arrows. On the basis of this theoretical model,

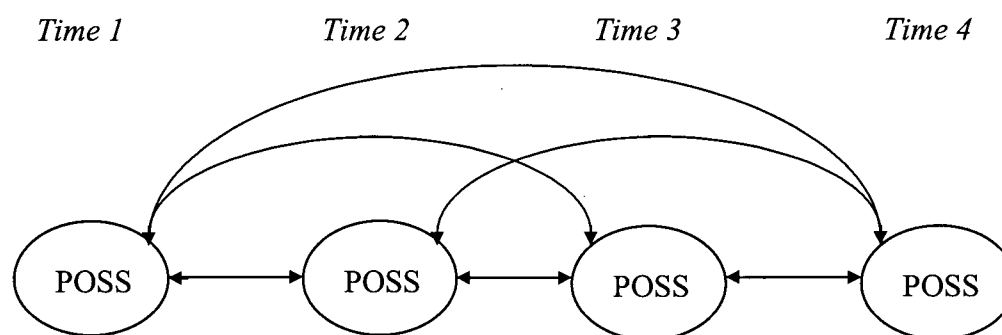


Figure 3.2 Theoretical model of change in perception of social status (POSS) based on a four-wave longitudinal design.

Study 2 attempted to compare the two groups--Japanese students studying in the target speech community and those studying in Japan. To this end, one hypothesis and two research questions were stated as follows:

Hypothesis 4:

The Japanese students studying in the target speech community come to show increasingly and significantly higher levels of pragmatic competence to offer advice in English than those studying in Japan.

Research question 1:

Do the students studying in the target speech community come to show the same preferences for advice type as native speakers of English, depending on the status relationship of the conversational participants?

Research question 2:

Do the students studying in Japan come to show the same preferences for advice type as native speakers of English, depending on the status relationship of the conversational participants?

3.3 Summary

This chapter first described three hypotheses tested in Study 1. As illustrated in Figure 3.1, Hypotheses 1, 2, and 3 stated the cause-effect relationships among three constructs, that is, perception of social status (POSS), English proficiency (PROF), and exposure to English (EXPO). Next, one hypothesis and two research questions were presented for Study 2 in which focus was given to multi-group analyses concerning change over time in perception of social status. As shown in Figure 3.2, the change was assessed on the basis of factor correlations across time. The next chapter discusses in detail the methodology in these two studies.

CHAPTER IV

METHOD

4.0 Overview

This chapter begins with a description of research sites, followed by a description of the recruitment of the subjects. A specification of the instruments used for Study 1 and Study 2 is discussed accompanied by illustrations of the relationships between three theoretical constructs (i.e., perception of social status, English proficiency, and exposure to English) and their respective measures. Finally, statistical techniques employed for Study 1 and Study 2 are presented together with a brief explanation of technical terms.

4.1 Research Sites

Studies 1 and 2 were conducted at Ritsumeikan University in Kyoto, Japan and the University of British Columbia (UBC) in Canada. Ritsumeikan University is a prestigious private university in western Japan. It launched an academic exchange program with UBC in 1991, the purpose of which is to provide Ritsumeikan students with an integrated language and content program in an English immersion environment (personal communication with a head teacher of the program, 1995). About 100 Ritsumeikan students participate in an eight-month program each year, and about 80 percent of them live in suites with about 160 UBC students in an on-campus facility called UBC-Rits House, whereas about 20 percent of them live in several on-campus dormitories. The Canadian portion of the studies was conducted at UBC-Rits House.

The Japanese portion of Study 2 was conducted at Ritsumeikan University. An instructor teaching two English classes (one for the second-year students and the other for the third-year students) and an instructor teaching a class on International Relations to third-year students volunteered to administer questionnaires in their classes. Each class comprised about 50 to 60 students.

4.2 Subjects

The subjects in Study 1 consisted of 101 Ritsumeikan students who came to Canada to study for eight months in the UBC-Rits academic exchange program.⁴ They were second- or third-year students enrolled in various departments at Ritsumeikan University. Their levels of English proficiency as measured by the *Test of English as a Foreign Language* (TOEFL) ranged from 480 to 600 when they were preparing in Japan for studying abroad. Some had lived and studied abroad, and others had never stepped outside of Japan. About 30 percent of the UBC-Rits students came to Canada at the beginning of August and took several ESL courses at UBC until the program started in September, whereas the others came to UBC at the end of August.

In addition to UBC-Rits students, Study 2 employed 132 Ritsumeikan students who did not come to Canada and who continued to study in Japan.⁵ They were also second- or third-year students enrolled in various departments. They were required to take two or three English classes per term, the contents of which were literature, linguistics, or conversation. Like UBC-Rits students, some Kyoto-Rits students had the experience of

⁴ They are called "UBC-Rits students" throughout the rest of the dissertation.

⁵ They are called "Kyoto-Rits students" throughout the rest of the dissertation.

traveling, living, and studying abroad, and others had never been to a foreign country. It is important to note that students in both groups were those who had learned English through communication-based classes for three years in Japanese upper secondary schools.⁶

4.3 Data Collection Procedure

For purposes of Study 1, data were collected from UBC-Rits students both in Japan and Canada every three months starting July, 1998. In July when they were preparing in Japan for studying abroad, the researcher visited Ritsumeikan University to collect data. The researcher visited the Academic Writing Course that all exchange program participants were required to take from the mid April to the end of July, and asked them to work on two questionnaires described in the next section.⁷ Because the questionnaires were constructed on the basis of Wolf's (1988) suggestion--that for educational research, a full questionnaire should require certainly less than 30 minutes to complete, and preferably, less than 15 to 20 minutes, it actually took the UBC-Rits students less than 15 minutes to complete even at the first administration. After they came to UBC in August, the researcher visited all sections of the course titled LANE 206, offered in the academic exchange program, and asked them to work on the same two questionnaires. This in-class data collection at UBC was conducted in October 1998, and January and April, 1999. Because two students returned to Japan in the middle of the

⁶ As discussed in Chapter I, communication-based classes have been actually implemented in Japanese secondary schools since 1993.

⁷ It is important to note here that at the time of data collection, the researcher explained that all the participants in the present study had the right to refuse to participate at any time. It was assured that none of the participants would be put at a disadvantage whether they participated in the present study or not.

academic year and data from two students were incomplete (i.e., they left most questions unanswered), complete data across all four time points were available for a total of 97 UBC-Rits students.

As for data from Kyoto-Rits students, the researcher asked two instructors to administer the questionnaires in their classes, when the researcher visited Ritsumeikan University in July. Data collection from Kyoto-Rits students was conducted four times at approximately the time of data collection from UBC-Rits students. At the first data collection, a total of 132 Kyoto-Rits students volunteered to join this research project. Because several students decided to withdraw from the classes in which the questionnaires were administered, complete data across four time points were available for a total of 102 Kyoto-Rits students. As a result, 97 UBC-Rits students and 102 Kyoto-Rits students were compared in Study 2. A summary of the data collection procedure from these two groups is shown in Table 4.1.

Table 4.1 *Summary of Data Collection Procedure*

	<i>Time 1</i> (July)	<i>Time 2</i> (Oct.)	<i>Time 3</i> (Jan.)	<i>Time 4</i> (Apr.)
UBC-Rits Students N=97	Japan (QPI, QUCE, MCQ)	Canada (QUCE, MCQ)	Canada (QUCE, MCQ)	Canada (QUCE, MCQ)
	← Participating in the exchange program at UBC →			
Kyoto-Rits Students N=102	Japan (QPI, MCQ)	Japan (MCQ)	Japan (MCQ)	Japan (MCQ)

Note. The instruments used for data collection at each time point are enclosed in parentheses. QPI, QUCE, and MCQ denote a questionnaire on personal information, a questionnaire on current uses of English, and a multiple-choice questionnaire, respectively, which are discussed in the next section.

4.4 Instruments

4.4.1 A Questionnaire on Personal Information

The questionnaire on demographic information was administered to the UBC-Rits students and Kyoto-Rits students in July at Ritsumeikan University in Japan. It was constructed to obtain background information about the two participant groups. Specifically, it was composed of items concerning the students' educational backgrounds, parents' first languages, experiences of living in foreign countries, and the like. The information from this questionnaire was important to standardize the information

obtained on student backgrounds and experiences across the two groups. A full copy of this questionnaire is in Appendix B.

4.4.2 A Questionnaire on Current Uses of English

A questionnaire on uses of English (see Appendix C) was administered four times, in July and October 1998, and January and April 1999, to UBC-Rits students (see Table 4.1). For purposes of Study 1, the questionnaire was constructed to obtain information on *amount of exposure to English* and *English proficiency*, by which change in UBC-Rits students' perception of social status when giving advice in English was accounted for.

4.4.2.1 Items Measuring Amount of Exposure to English

The questionnaire on uses of English was constructed to obtain information on the contexts and characteristics of UBC-Rits students' uses of English in daily life, both inside and outside the classroom. It was administered in classes to sample four weeks of English use; students were asked to report uses of English during the week just preceding administration of the questionnaire. The questionnaire was a one-page record organized into several categories, including productive and receptive uses of English. It was designed to obtain information on the day-average amount of exposure (in hours and minutes) to English via TV, movies, books, classes, and the day-average exposure (in hours and minutes) of interactions with roommates in English outside of classrooms.

The amount of exposure to English was assessed using two measures: that is, *exposure through productive uses of English* indicated by the sum of hours and minutes reported in items (a) to (c) and (n) to (q), and *exposure through receptive uses of English*

indicated by the sum of hours and minutes reported in items (d) to (m). A schematic representation of the relationship between the theoretical construct and its two measures is shown in Figure 4.1.

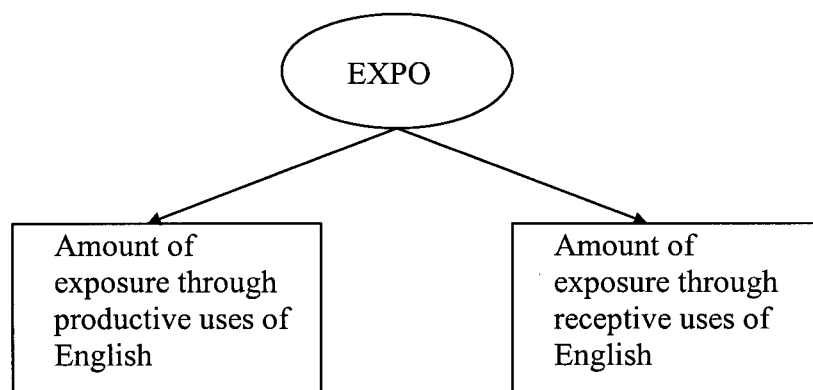


Figure 4.1 A theoretical construct, EXPO (exposure to English) and its two measures (amount of exposure through productive and receptive uses of English).

4.4.2.2 An Item Measuring English Proficiency

The questionnaire on uses of English also included an item regarding the UBC-Rits students' English proficiency. Study 1 focused on their proficiency as measured by the *Test of English as a Foreign Language* (TOEFL). It should be noted, however, that there was no implication that the TOEFL was the best test to measure English proficiency. Indeed, the TOEFL has both strengths and weaknesses.

One major weak point is reliability. Although Gronlund (1985) states that standardized tests such as the TOEFL have been thoroughly tested and their reliability and validity have been carefully investigated and demonstrated for the intended uses of the test, the test score reliability reported by the test publisher does *not* always apply to a

particular group sampled from the population. In other words, because reliability is sample-specific (Pedhazur, 1997), it is necessary to calculate the reliability for the sampled group. Unfortunately, only students' section scores were obtainable in Study 1 so that investigation of the reliability of item scores was impossible. Another difficulty is related to a validity problem. There are many debates as to whether the TOEFL really measures English proficiency (e.g., Duran, Canale, Penfield, Stansfield, & Liskin-Gasparro, 1985; Stansfield, 1986). However, this issue is beyond the scope of the present study.

The major reason for including TOEFL in Study 1 was that "scores from the TOEFL are used by many colleges and universities in North America as a complement to other types of information such as grades, academic achievement tests, and letters of recommendation, for deciding which non-native English-speaking students to accept into academic programs" (Bachman, 1991, p. 58). In fact, the decision concerning whether or not the UBC-Rits students were able to participate in the exchange program was made on the basis of their TOEFL scores. The second reason was that the content of the TOEFL is developed to measure language proficiency, as pointed out by Bachman (1991). Moreover, the test itself has been designated by its makers as a measure of English proficiency (Educational Testing Service, 1996). The TOEFL is the largest mass assessment of proficiency in the world today (about 1,000,000 administrations in 1997) and has been the basis of considerable research employing it as a standard measure of proficiency in English as a second language. The third reason was that the UBC-Rits students were required as part of the in-class activities to take the TOEFL three times during the exchange program. Institutional TOEFL administrations occurred, and scores

reported, just before students were asked to summarize their use of English during the week. For these reasons, Study 1 used TOEFL as a test to measure the *variance* of the UBC-Rits students' English proficiency, not their English proficiency per se. Their proficiency was assessed using scores in three sections of the TOEFL. A schematic representation of the relationship between the theoretical construct and its three measures is shown in Figure 4.2.

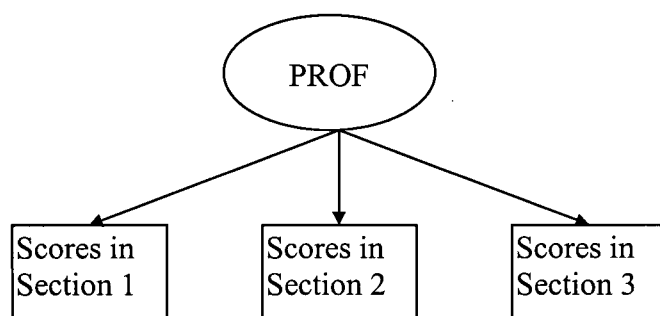


Figure 4.2 A theoretical construct, PROF (English proficiency) and its three measures (scores in sections 1, 2, and 3).

4.4.3 A Multiple-Choice Questionnaire to Assess Perception of Social Status

Since both Study 1 and Study 2 aimed to observe UBC-Rits and Kyoto-Rits students' change over time in their preferences for a particular speech act type, the multiple-choice questionnaire described below was administered repeatedly to both groups. Data were collected both in Japan and Canada at three-month intervals starting in July (see Table 4.1 above).

4.4.3.1 Constructing a Multiple-Choice Questionnaire

The multiple-choice questionnaire consisted of 12 scenarios and four response choices for each scenario. Triandis, Chen and Chan (1998) state, “The scenario approach has an advantage because it samples situations that are close to those that occur in everyday university student life” (p. 277). In the multiple-choice questionnaire administered to UBC-Rits and Kyoto-Rits students, all scenarios were written in both English and Japanese to avoid misunderstanding of scenarios caused by their varying levels of English reading comprehension, but all choices were offered only in English because pairs of English-Japanese equivalents may affect participants’ decision-making process as described in Chapter II and may cause a validity problem for the instrument to the extent that participants would make their decisions on the basis of the Japanese translations when in fact what was required was a decision based exclusively on the English alternatives.

Each scenario represented one of three social status variables: *higher status* (a supervisor), *status equal* (a classmate), and *lower status* (a first-year university student). Subjects were asked to play a role as addressers to these three types of people. The reason that a first-year university student falls into the lower status category is that as part of an existing Japanese hierarchical system, second- and third-year students are considered to be “*senpai*” that is, be in a higher status than first-year students, and according to this hierarchy, first-year students normally use polite expressions when talking to “*senpai*.” Conversely, it is rare that second- and third-year students use polite expressions when talking to first-year students. Initially at least, choices may be expected to reflect more of

the Japanese understanding of appropriate uses than choices made later in the year. The imaginary supervisor, classmate and first-year student are described as follows:

Supervisor:

P.D. is your supervisor. You have been taking P.D.'s seminar for three months.

You and P.D., together with other students, have gone out for dinner several times after the seminar. You have visited P.D.'s office several times to talk about the topic you would present in the seminar.

Classmate:

C.J. is your classmate. You and C.J. often go out for lunch together after the class.

You have borrowed C.J.'s notebook several times before. You regard C.J. as a good friend.

First-year university student:

X.L. is a first-year student. You and X.L. belong to the same club. You and X.L. often go out for dinner together after the club activity. You regard X.L. as a good friend.

It should be noted that in this questionnaire respondents were asked to imagine that all scenarios happened in Canada and all imaginary characters were native speakers of English. These instructions made their responses represent their understanding of the sociocultural rules in the target speech community that link the use of language with the perception of social status. Moreover, as in Hinkel (1997), all references to personal names and gender markers were avoided in all scenarios and response choices so as not to obscure the social status variable.

There were 12 scenarios on the multiple-choice questionnaire. Four scenarios were provided for each social status value. The brief descriptions of four scenarios with the *supervisor* (higher status) are shown below:

1. *Restaurant*: The supervisor is about to make a bad menu choice.
2. *Illness*: The supervisor works in the office late at night and looks pale.
3. *Bookstore*: The supervisor is considering buying an expensive book without knowing that another bookstore sells it at a 20 percent discount.
4. *Repairing*: The supervisor is considering a trip to Banff from Vancouver in a car which breaks down frequently.

The brief descriptions of four scenarios with the *classmate* (equal status as students) are shown below:

1. *Class*: C.J. considers skipping today's afternoon class.
2. *Computer lab*: C.J. works on the assignment late at night and is visibly tired.
3. *Broken vending machine*: C.J. couldn't get a pop nor get the money back from a broken vending machine.
4. *Tipping*: C.J. has forgotten to leave a tip when leaving a restaurant.

The brief descriptions of four scenarios with the *first-year university student* (lower status) are shown below:

1. *Academic course*: X.L. considers taking a difficult academic course.
2. *Library*: X.L. studies in the library late at night and looks pale.
3. *Cafeteria*: X.L. didn't get the exact amount of change at the cashier of the cafeteria.
4. *Repair Shop*: X.L. is thinking of taking a car to a notorious repair shop.

The four response choices in each scenario represented one of four speech act realizations in advice-giving situations: that is, direct advice, hedged advice, indirect comment, and not giving advice. In keeping with earlier research (Hinkel, 1997; Rose, 1994), all response choices in the multiple-choice questionnaire were constructed on the basis of responses to the discourse completion tests administered as part of the pilot study in which participants were 91 Japanese students participating in the exchange program in the year preceding this research project. The direct advice items selected from the pool of choices included the use of 'should' without hedging. Hedged advice options were constructed to include lexical hedging (maybe, I think) that Japanese learners of English putatively use frequently in conversation. Indirect comments with no advice were also included as one of the four response options in each scenario, and they were selected such that the speaker's intentions were not made explicit (Brown & Levinson, 1987; Levinson, 1983). As in Hinkel's (1997) study, the fourth selection was an explicit choice for opting out that remained constant for all scenarios. Examples of direct and hedged advice, indirect comments, and opting out are shown in 1 to 4, respectively.

1. You should go home. You look like you don't feel well.
2. Maybe it's better to go home. You look like you don't feel well.
3. You look like you don't feel well.
4. Nothing

Each scenario presented as choices direct and hedged advice, and indirect comment in random order. The opting-out option was always placed in the fourth choice. It should be noted that in order to reduce the memory carry-over effect caused by using the same material four times at three-month intervals to the same subjects, the 12

scenarios and four choices in the questionnaire were randomly re-ordered in each administration. A full copy of the multiple-choice questionnaire is in Appendix D.

4.4.3.2 Evaluating the Stability of Native Speakers' Preferred Choices in the Multiple-Choice Questionnaire

As discussed in Chapter I, Japanese students' pragmatic development was assessed by the approximation of their preferences for advice type to native speakers'. It was thus necessary to determine which response choice in each scenario was preferable to the other choices in native speakers' eyes, and to evaluate the stability in native speakers' preferred choices. Since there were no right or wrong answers in the multiple-choice questionnaire designed to examine preference, native speakers' responses were expected to vary, to some extent, in each scenario, but to be consistent at different points in time. The degree of stability in their preferred choices was estimated by the test-retest method.

The questionnaire from which Japanese translations were deleted (see Appendix E) was distributed to over 100 native speakers who were the UBC-Rits students' roommates or floor mates living at on-campus residences at UBC. At the first administration, a total of 82 native speakers responded to the questionnaire completely. Because critical factors in evaluating the magnitude of a stability estimate must include the elapsed time between testings (Crocker & Algina, 1986), the second data collection was conducted approximately five months after the first one.⁸ The randomly re-ordered

⁸ Crocker and Algina (1986) state that there is no single answer to how much time should elapse between testings, and that the time period should be long enough to allow effects of memory or practice to fade but not so long as to allow maturational or historical changes to occur in the examinees' true scores.

multiple-choice questionnaire was distributed again to the 82 respondents who had volunteered to participate in the first data collection. Complete data across the two time points were available for a total of 71 native speakers. Table 4.2 shows their preferred choices in each scenario at the first and second data collections.

Table 4.2 *Native Speakers' Preferences for Advice Type in Frequencies*

Scenario #	<u>Higher Status</u>				<u>Status Equal</u>				<u>Lower Status</u>			
	1	4	7	10	2	5	8	11	3	6	9	12
Direct (1)	6	11	11	10	11	46	34	5	10	14	35	31
(2)	2	10	8	9	13	48	37	2	8	10	30	27
Hedged (1)	37	21	39	33	19	12	12	36	14	19	12	19
(2)	42	25	40	31	20	11	10	38	13	24	15	21
Indirect (1)	20	31	20	22	37	10	17	10	41	37	23	16
(2)	21	33	16	24	33	8	20	8	43	35	20	14
Not Giving (1)	8	8	1	6	4	3	8	20	6	1	1	5
(2)	6	3	5	7	5	4	4	23	7	2	6	9

Note. N = 71 in each scenario. The first and second rows in each advice type show the frequencies in the first and second data collections, respectively.

The results in Table 4.2 indicated that at both administrations, there were no scenarios in which two response choices were chosen by an equal number of native speakers, and that the order of the most to least preferred choices in each scenario did not change across the two time points. The results also indicated that the native speakers' preferred solutions to the scenarios were within a range that included adjacent forms of advice (e.g., "hedged advice" and "indirect advice" for scenario 1) along the continuum of directness from "direct advice" to "not giving advice." In some cases, the native

speaker's solution was a choice of non-adjacent patterns (e.g., "hedged advice" and "not giving advice" for scenario 11).

Table 4.3 exhibits the degree of stability in native speakers' preferred choices across the two time points estimated by the test-retest method. The results in Table 4.3 indicated that the test-retest coefficients ranged from the high .70s to low .90s, suggesting that their preferred choices were quite stable across the two time points. Taken altogether, the results in Tables 4.2 and 4.3 suggest that their preferred choices were stable enough to function as the baseline against which to assess UBC-Rits and Kyoto-Rits students' perception of social status when offering advice in English.

Table 4.3 *The Coefficients of Stability in Native Speakers' Preferred Choices in the Multiple-Choice Questionnaire*

<u>Scenarios</u>	<u>Coefficients</u>
<u>Higher Status</u>	
1	.797
4	.780
7	.917
10	.845
<u>Status Equal</u>	
2	.905
5	.894
8	.914
11	.933
<u>Lower Status</u>	
3	.908
6	.844
9	.771
12	.858

Note. N = 71 in each scenario.

Scoring the UBC-Rits and Kyoto-Rits students' choice in each scenario was straightforward on the basis of the native speakers' preferred choice shown in Table 4.2. Specifically, when the response from a Japanese student was the one that the native speakers thought of as most appropriate, the student received four points. Because four scenarios were included in each status relationship, scores from a Japanese student varied from 4 to 16 for each status relationship unless he/she left some questions unanswered. Thus Japanese students' perception of social status was assessed using three measures, that is, scores in scenarios for higher status, status equal, and lower status. A schematic representation of the relationship between the theoretical construct and its three measures is shown in Figure 4.3.

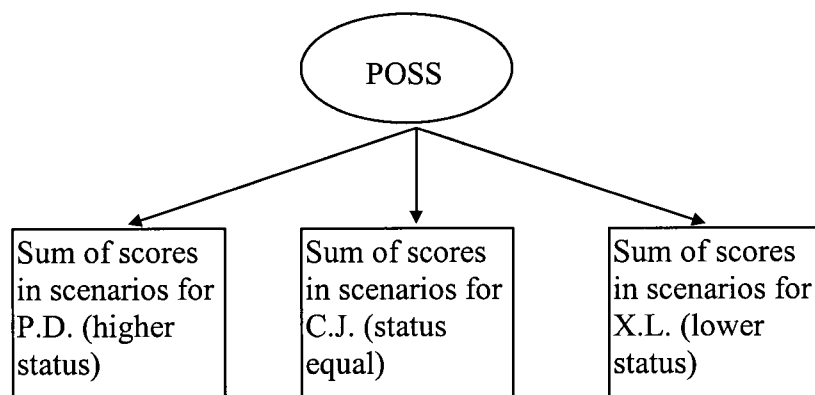


Figure 4.3 A theoretical construct, POSS (perception of social status) and its three measures (sum of scores in scenarios for higher, equal and lower statuses).

4.5 Data Analysis Techniques

4.5.1 The Rationale for the Use of Structural Equation Modeling (SEM) with Latent Variables

The hypotheses and research questions shown in Chapter III require a statistical method that has the ability to analyze longitudinal data. Several methods can be used to analyze longitudinal data; for example, multiple regression analysis, repeated-measures analysis of covariance (ANCOVA), path analysis, and structural equation modeling (SEM) with latent variables. There are both advantages and disadvantages to these statistical techniques when applying them to longitudinal data in Studies 1 and 2. First, multiple regression analysis could be performed by treating perception of social status, English proficiency, and amount of exposure to English at Time 1 as independent variables, and perception of social status at Time 2 as a dependent variable. The overall R^2 would represent the proportion of variance in perception of social status at Time 2 predicted by the three variables at Time 1. As some have pointed out (e.g., Pedhazur, 1997), however, the estimation of R^2 and the standardized regression coefficient (β) in the context of multiple regression is sensitive to measurement errors. Specifically, Pedhazur (1997) points out that measurement errors in the dependent variable lead to a downward bias in the estimation of the β 's and R^2 . Those in the independent variables lead to a downward bias in the estimation of the β 's, and to either upward or downward bias in the estimation of the regression coefficient (B). Thus it is no exaggeration on the part of Fleiss and Shrout (1997) when they state that "effects of measurement errors can become devastating" (p. 1190).

Second, repeated-measures ANCOVA could be used with scores at Time 1 as a covariate and scores at Times 2, 3, and 4 as dependent variables. This statistical technique seems useful, especially when the subjects are not drawn at random as in Studies 1 and 2. It should be noted, however, that this technique is valid only when all stringent assumptions are met (see Cohen & Cohen, 1983; Pedhazur, 1997 for a review). It should also be noted that ANCOVA can reduce, but not entirely eliminate, selection threats to the internal validity of quasi-experimental studies. Glass and Hopkins (1996) state, "In reality, ANCOVA is never able to make the results of a quasi-experiment as definitive as those of randomized experiments" (p. 593).

A third method, path analysis based on multiple regression analysis, also has several disadvantages. Figure 4.4 represents an illustrative application of path analysis to the hypothesized relationships among the three constructs shown in Figure 3.1 based on a four-wave longitudinal design.

Path analysis is based on a set of restrictive assumptions (see Pedhazur, 1997 for a detailed discussion). Both Studies 1 and 2 violated at least two assumptions--that variables are measured without errors, and that residuals are not correlated. The first assumption is rarely met in practice. Consequently, the presence of measurement errors may be very damaging to results of path analysis as well as multiple regression analysis. The second assumption is unreasonable in a longitudinal study in which subjects are measured at several points in time on the same variables. One example of the violation of this assumption is the memory carry-over effects caused by the repeated administration of the same instrument.

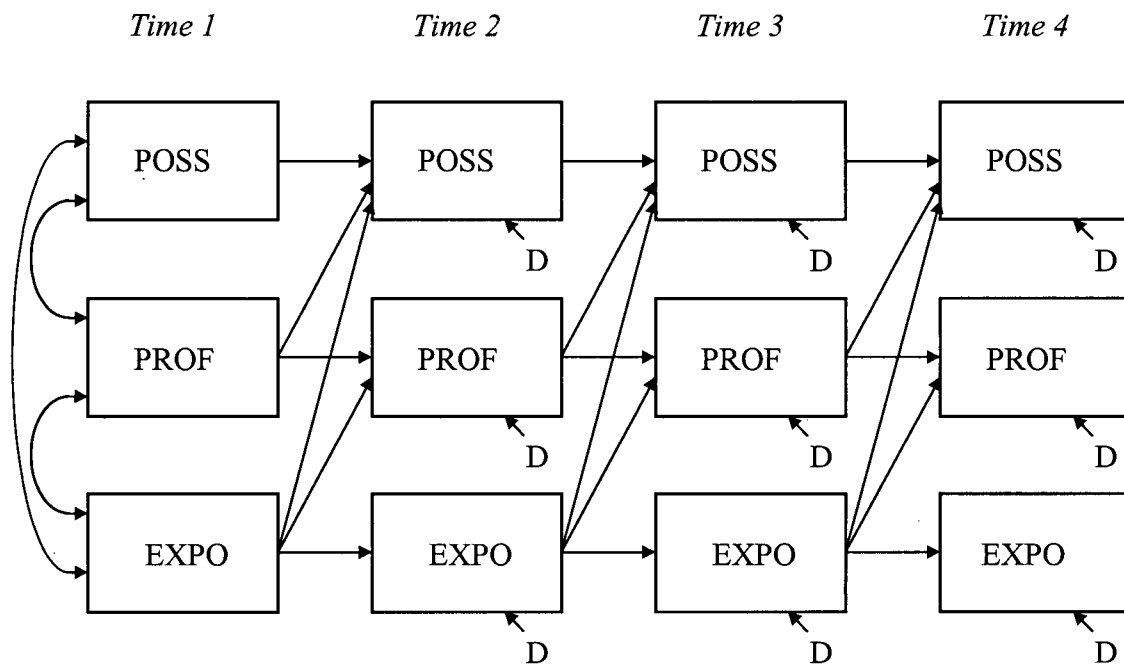


Figure 4.4 Example of a path model of POSS (perception of social status), PROF (English proficiency), and EXPO (exposure to English) on a four-wave longitudinal design.

Note. D's denote random disturbance (see the next section for detail).

In sum, it is difficult or impossible for these statistical techniques to evaluate adequately the hypotheses and research questions shown in Chapter III. Thus Studies 1 and 2 employ the last option, structural equation modeling (SEM) with latent variables, which is “a comprehensive, flexible approach to modeling relations among variables” (Hoyle & Smith, 1994).

4.5.2 Advantages of the Use of SEM with Latent Variables in Analyzing

Longitudinal Data

The flexibility of SEM with latent variables has been widely acknowledged only during the past decade (see Bentler, 1986, for a review). First, it enables researchers to translate questions regarding theoretical constructs into precise and testable hypotheses and to compare alternative models of cause-effect relationships (Connell, 1987). Second, SEM has the advantage of calculating all of the parameters in the model simultaneously and providing a test of overall fit of the model to the data (Farrell, 1994). Third, it allows measurement errors based on the notion that the measures often contain both random and nonrandom errors (Bollen, 1989).⁹ Fourth, it enables researchers to examine the consistency of a model over time across different groups of subjects, and the equality of estimates of particular parameters over time in the different groups (Byrne, 1998; Farrell, 1994; Hoyle & Smith, 1994). As discussed in detail in subsequent chapters, the present study took such advantages of SEM with latent variables, not only when evaluating the hypotheses and research questions (Chapters VII and VIII), but also when assessing the validity and reliability of the instruments used (Chapter VI). Figure 4.5 on the next page illustrates in SEM terms the hypothesized relationships among perception of social status, English proficiency, and amount of exposure to English based on a four-wave longitudinal design.

Notice that Figure 4.5 is a synthesized model of Figures 3.1, 3.2, 4.1, 4.2, and 4.3. Several characteristics of the model in Figure 4.5 need to be addressed here. First of all, latent (unobserved) variables, that is, theoretical constructs or factors are enclosed in

⁹ Bollen (1989) compared the results with and without measurement error, thus allowing an assessment of the differences that the error makes.

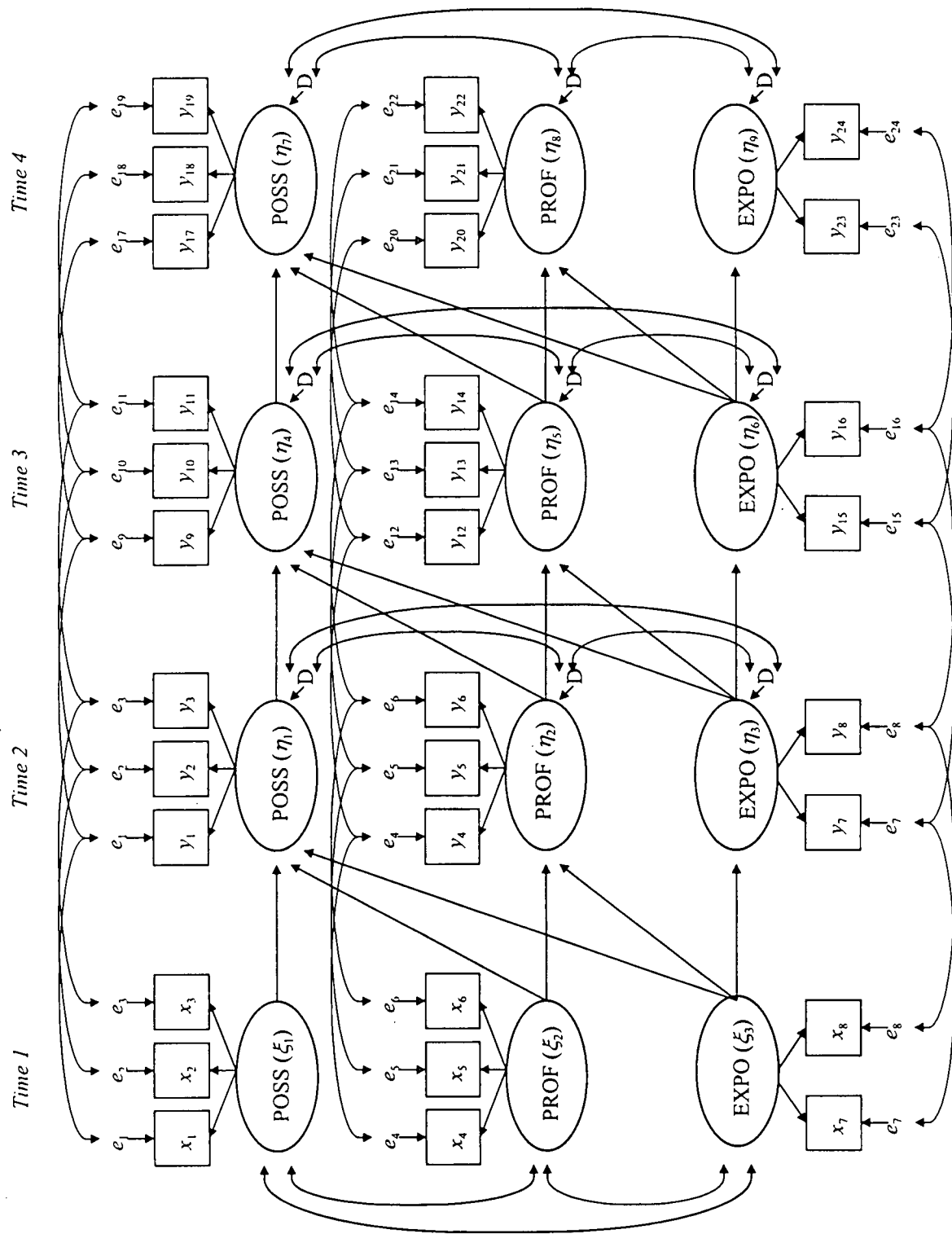


Figure 4.5 A complete latent variable model based on a four-wave longitudinal design

circles (ovals), whereas observed variables, that is, measures of the latent variables denoted as x_i 's ($i = 1-8$) and y_i 's ($i = 1-24$) are enclosed in squares (rectangles). Second, the latent and observed variables in the model can be categorized into exogenous (independent) and endogenous (dependent) variables. As represented by x 's and y 's, the variables at Time 1 are exogenous ones, whereas those at Times 2, 3, and 4 are endogenous ones in the model. The latent exogenous and endogenous variables are represented by ξ_j (ksi) ($j = 1-3$) and η_j (eta) ($j = 1-9$), respectively. Third, the latent endogenous variables are only partially accounted for by the model. The unexplained component is represented by D (the random disturbance) in the model. Fourth, associated with each observed variable is an error term denoted as e_i ($i = 1-8$ for x 's; $i = 1-24$ for y 's). Let us turn next to *structural parameters* in the model. Bollen (1989) explains structural parameters as follows:

The structural parameters are invariant constants that provide the "causal" relation between variables. The structural parameters may describe the causal link between unobserved variables, between observed variables, or between unobserved and observed variables. (p. 11)

A first structural parameter addressed here is the β (beta) coefficient that links the latent endogenous variables. A second parameter is the γ (gamma) regression coefficient that links the latent exogenous and endogenous variables. Third, the unidirectional arrows leading from the latent variable to each of the observed variables indicate that these score values are influenced by the latent variable. These coefficients are represented by λ_{ij} (lambda) ($i = 1-8, j=1-3$ for x 's; $i = 1-24, j=1-9$ for y 's). Finally, curved two-way arrows represent covariances or correlations between pairs of variables.

Let us move now on to two parts of the SEM model, namely the measurement model and structural (equation) model. Jöreskog and Sörbom (1996) define them as follows:

The **measurement model** specifies how latent variables or hypothesized constructs depend upon or are indicated by the observed variables. It describes the measurement properties (reliabilities and validities) of the observed variables. (p. 1, bold type in original.)

The **structural equation model** specifies the causal relationships among the latent variables, describes the causal effects, and assigns the explained and unexplained variance. (p. 1, bold type in original.)

Notice that Figures 4.1, 4.2, and 4.3 shown earlier in this chapter are examples of the measurement model, whereas Figures 3.1 and 3.2 in Chapter III are examples of the structural model. Thus Figure 4.5 can be said to be a general structural equation model, or structural equation model with latent variables (Hoyle & Smith, 1994).¹⁰ Studies 1 and 2 use this statistical technique in analyzing the data collected.

There are several characteristics of the model shown in Figure 4.5 that merit attention. First, three latent variables were represented by their respective observed measures at each time point. In other words, each observed variable was linked to a single latent variable within each of the four time points. Hoyle and Smith (1994) noted that a desirable measurement model is one in which each latent variable or facet of a construct is uniquely and adequately represented by three or more indicators. Bentler and Chou (1987) stated that in general, a minimum of three indicators per latent variable is

¹⁰ For an exploration of SEM terms, see Bollen (1989), Byrne (1998), Hayduk (1987), and Jöreskog and Sörbom (1996).

recommended unless another latent variable may serve as an indicator of the latent variable. As can be seen in Figure 4.5, the model met this condition. Second, the model included correlations among all the latent variables denoted as D 's (disturbance terms) and serial correlations among measurement errors. As discussed above, repeated measurement using the same instrument often results in correlated measurement errors (e.g., Judd & Milburn, 1980; Kessler & Greenberg, 1981). Memory carry-over effects were one possible systematic measurement error because students' memories of responses in the first administration of the multiple-choice questionnaire could influence their responses in a subsequent administration. With this hypothesized model in hand, a confirmatory factor analysis (technically, a special case of SEM with latent variables) was performed in Studies 1 and 2.

4.6 Summary

This chapter has provided a detailed description of research sites in Japan and Canada, two groups of subjects (UBC-Rits and Kyoto-Rits groups), and data collection procedure, followed by a description of the instruments used in Studies 1 and 2. Next, the advantages of SEM with latent variables in analyzing longitudinal data were illustrated by comparing this approach with several other statistical methods including multiple regression analysis, repeated-measures ANCOVA, and path analysis. Finally, the overall model on which Studies 1 and 2 placed their analytic bases was presented together with a brief explanation of technical terms used in subsequent chapters. The next chapter shows the results of descriptive statistics for the data collected from the subjects in the studies.

CHAPTER V

DESCRIPTIVE STATISTICS

5.0 Overview

The aim of this chapter is to report the results of descriptive statistics for the data collected from the two groups (UBC-Rits and Kyoto-Rits groups). Means, standard deviations, skewness and kurtosis of the raw data are summarized separately for each group.

5.1 Descriptive Statistics for the Raw Data from UBC-Rits Group

Table 5.1 shows the descriptive statistics for the raw data collected through the multiple-choice questionnaire (MCQ), the TOEFL, and the questionnaire on uses of English (QCUE) at four points in time.

Table 5.1 *Summary of Descriptive Statistics for the Raw Data from UBC-Rits Group*

Variables	Mean	SD	Skewness	Kurtosis
<u>MCQ</u>				
PD1	12.13	1.91	-.501	-.825
PD2	12.82	1.64	-.894	.243
PD3	12.99	1.82	-1.048	.083
PD4	13.12	1.54	-1.227	.702
CJ1	10.78	2.08	.023	-.124
CJ2	11.43	1.78	-.056	-.710
CJ3	12.98	1.71	-.074	-.129
CJ4	13.26	1.57	-.107	-.346

(To be continued on the next page)

(Continued)

XL1	11.04	1.82	-.045	-.434
XL2	11.74	1.55	-.121	-.150
XL3	12.75	1.55	-.172	.158
XL4	13.00	1.49	-.265	.010

TOEFL

L1	51.42	3.37	.518	.258
L2	51.44	3.59	.294	.240
L3	52.47	3.35	-.073	.082
L4	53.46	3.67	.348	-.306

G1	52.23	3.12	.436	.029
G2	51.34	3.49	.346	1.028
G3	51.81	3.10	-.112	.573
G4	52.79	3.88	.706	.428

R1	51.65	2.83	1.025	2.468
R2	50.55	3.18	.165	1.177
R3	51.16	3.16	.171	.715
R4	52.04	3.25	.492	1.083

QCUE

PRO1	37.00	40.30	1.377	2.655
PRO2	246.34	126.76	.458	-.480
PRO3	273.05	130.67	.529	-.218
PRO4	299.14	133.78	.526	-.288

REC1	82.68	71.35	.812	-.143
REC2	200.92	106.03	.760	.936
REC3	285.58	123.05	.056	-.344
REC4	275.08	113.23	.068	.092

Note. N = 97. PD, CJ, XL in the MCQ denote the observed variables represented by the sum of scores in scenarios for higher status, status equal, and lower status, respectively. L, G, and R in the TOEFL denote the observed variables represented by the scores in sections 1, 2, and 3, respectively. PRO and REC denote the observed variables represented by the amount of exposure (in minutes) through productive and receptive uses of English, respectively. The number attached to each observed variable indicates the data collection point (1 = Time 1; 2 = Time 2; 3 = Time 3; and 4 = Time 4).

5.2 Descriptive Statistics for the Raw Data from Kyoto-Rits Group

As discussed in the previous chapter, only the multiple-choice questionnaire (MCQ) was administered to Kyoto-Rits group. Hence Table 5.2 illustrates the descriptive statistics for the raw data obtained by means of the MCQ only.

Table 5.2 *Summary of Descriptive Statistics for the Raw Data from Kyoto-Rits Group*

Variables	Mean	SD	Skewness	Kurtosis
<u>MCQ</u>				
PD1	12.66	1.63	-.512	.039
PD2	12.66	1.82	-.476	-.171
PD3	12.63	1.85	-.284	-.525
PD4	12.66	2.15	-.314	-.408
CJ1	11.53	1.57	-.066	.147
CJ2	11.52	1.73	-.144	-.451
CJ3	11.56	1.85	.119	-.133
CJ4	11.56	2.16	-.167	-.329
XL1	11.48	1.55	-.172	.158
XL2	11.44	1.67	-.115	-.036
XL3	11.57	1.91	-.274	-.003
XL4	11.61	2.12	-.184	-.179

Note. N = 102

5.3 Summary

The results of descriptive statistics for the raw data collected through three instruments were summarized separately for UBC-Rits and Kyoto-Rits groups. Means reported in this chapter are used in Study 2 discussed in Chapter VIII. In the next chapter the validity and reliability of the scores obtained through the three instruments are assessed using SEM with latent variables.

CHAPTER VI

VALIDITY AND RELIABILITY

6.0 Overview

This chapter is concerned with the validity and reliability of scores in the three instruments introduced in Chapter IV. The rationale for the use of structural equations approach to validity and reliability is discussed accompanied by a brief review of traditional techniques. Finally, estimates of validity and reliability of the measures in the overall model as shown in Figure 4.5 are illustrated in summary tables.

6.1 Validity

6.1.1 A Brief Review of Classical Validity Techniques

Validity is concerned with whether a variable measures what it is supposed to measure. Content validity, criterion validity, construct validity, and convergent-discriminant validity are four traditional validity types, and they are all popular in the research validation process in behavioral sciences (Bollen, 1989). Content validity concerns whether the items adequately represent a performance domain or construct of specific interest. Crocker and Algina (1986) state that content validation continues until a theoretical definition of a construct is agreed upon by many researchers, and selected indicators fully cover the domain of the construct. Because the questionnaires used in Studies 1 and 2 were constructed based on the findings of and implications from a considerable body of literature and several preliminary studies, content validity was ensured theoretically to a certain extent. Put another way, it is unlikely that the questionnaires consisted of items which were totally irrelevant to the theoretical

constructs under investigation. It should be noted, however, that because the questionnaires had never been administered in their present forms prior to this research project, re-administration is a must to evaluate content validity thoroughly.

As for criterion validity, construct validity, and convergent-discriminant validity, psychologists have explored the weakness of these classical validity techniques in recent years, suggesting that they rely on correlations rather than structural coefficients to test validity (Bollen, 1989). Criterion validity requires the correlation of the criterion and the observed measure, whereas construct validity and convergent-discriminant validity need the correlation of measures of the same and different constructs. The problem is that these correlations may have little to do with the validity of a measure. That is, the three techniques use only observed measures rather than incorporating latent variables into the analysis. This assumes implicitly that each measure depends only on one latent variable and that the correlation of two observed variables accurately mirrors an association involving latent variables (see Bollen, 1989, for a detailed discussion concerning this issue).

Considering these disadvantages of classical validity techniques, the present study employed alternative approaches proposed by Bollen (1989). They are “several measures of validity that correspond to structural equations while also being related to the traditional measures” (Bollen, 1989, p. 206). Specifically, the alternatives used were *unstandardized validity coefficient* and *standardized validity coefficient*, wherein “the validity of a measure (x_i) of a latent variable (ξ_j) is the magnitude of the direct structural relation between ξ_j and x_i ” (Bollen, 1989, p. 197).

6.1.2 Unstandardized Validity Coefficients (λ_{ij})

The first gauge of the extent of the direct structural relationship between ξ_j and x_i is the unstandardized coefficient linking x_i and ξ_j , which is denoted as λ_{ij} (see section 4.5.2 in Chapter IV for a review of SEM terms). Because three observed measures in the multiple-choice questionnaire ($x_1 - x_3$) described in Figure 4.5 in Chapter IV are connected to a latent variable labeled POSS (ξ_1) (perception of social status) only, λ_{ij} is analogous to a regression coefficient from the simple regression of x_i on ξ_1 . The same story holds for the relationship between scores in three sections of the TOEFL ($x_4 - x_6$) and a latent variable labeled PROF (ξ_2) (English proficiency) and the relationship between amount of exposure through productive and receptive uses of English ($x_7 - x_8$) and a latent variable labeled EXPO (ξ_3) (exposure to English). Thus the formula for the measurement model at Time 1 can be shown as follows:

$$x_i = \lambda_{ij} \xi_j + \delta_i$$

where δ_i is a measurement error for x_i (i.e., e_i in Figure 4.5) with $E(\delta_i) = 0$, $\text{COV}(\xi_j, \delta_i) = 0$, $i = 1-8$, $j = 1-3$.

It is important to note that unstandardized λ_{ij} coefficient does not function as a validity coefficient if observed variables depending on the same latent variable are measured on very different scales, say, one variable in kilograms and the other in pounds to measure weight. The reason for that is described as follows:

To proceed with estimating a model, the latent variable must be assigned a scale.

A frequent means of doing this is to set one of the λ_{ij} coefficients leading from the

same latent variable to one. This sets the latent variable's scale to that of the observed variable, with its λ equal to one. The other λ_{ij} 's leading from the same latent variable are interpretable relative to the unit of the observed variable with a λ of one (Bollen, 1989, p. 198).

Because observed variables leading from the same latent variable were measured on the same scale in the model shown in Figure 4.5, the unstandardized λ coefficient can function as a validity measure. Table 6.1 exhibits estimates of the unstandardized validity coefficients in the measurement model at each time point.

Table 6.1 *Estimates of Unstandardized Validity Coefficients for the Measurement Model at Each Time Point*

<u>Latent variables</u>	<u>Time 1</u>	<u>Time 2</u>	<u>Time 3</u>	<u>Time 4</u>
POSS				
λ_{11}	1.00	1.00	1.00	1.00
λ_{21}	1.54	1.22	1.16	1.03
λ_{31}	1.30	.99	1.00	1.11
PROF				
λ_{42}	1.00	1.00	1.00	1.00
λ_{52}	.87	1.08	.94	1.06
λ_{62}	1.28	.99	.97	.92
EXPO				
λ_{73}	1.00	1.00	1.00	1.00
λ_{83}	1.64	.68	.61	.75

Note. For purposes of statistical identification the first validity coefficient of each latent variable is set to one. N = 97.

Notice that at Time 1 the value of λ_{83} (exposure through receptive uses of English) was distinctively higher than that of λ_{73} (exposure through productive uses of English). In contrast, at Times 2, 3, and 4, the values of λ_{73} 's were distinctively higher than those of

the other λ_{83} 's. These results suggest that λ_{83} is more valid and responsive to ξ_3 at Time 1, and so is λ_{73} at Times 2, 3, and 4. It is worthy of noting that although the values of λ_{83} 's at Times 2, 3, and 4, and the λ_{73} at Time 1 were lower than the value of the other measure at the same time point, they were all statistically significant parameters in the model. Specifically, in the output file of LISREL 8 and later versions, each estimated parameter is presented along with its related standard error and t -value. Byrne (1998) explains statistical significance of parameter estimates as follows:

The test statistic here is the t -statistic which represents the parameter estimate divided by its standard error; as such, it operates as a z -statistic in testing that the estimate is statistically different from zero. (p. 104)

Table 6.2 *The t-value for Each Estimated Parameter in the Model*

<u>Latent variables</u>	<u>Time 1</u>	<u>Time 2</u>	<u>Time 3</u>	<u>Time 4</u>
POSS				
λ_{21}	6.04	7.61	7.21	7.04
λ_{31}	6.00	7.07	6.99	8.04
PROF				
λ_{52}	4.92	7.88	8.87	8.68
λ_{62}	6.02	7.92	8.97	9.00
EXPO				
λ_{83}	2.71	5.10	4.17	4.33

Note. The t -values of λ_{11} , λ_{42} , and λ_{73} are not provided because these parameters are fixed for purposes of statistical identification.

Table 6.2 exhibits the t -value for each estimated parameter (λ_{ij}) in the model. Results of the test statistics reported in Table 6.2 reveal that all parameters in the model are $> \pm 1.96$ based on a level of .05, thereby suggesting that the hypothesis (that the estimate = 0.00) can be rejected. Thus interpretation of estimates of λ_{ij} and its t -values

indicates that the two measures of the latent variable EXPO (exposure to English) exhibit different degrees of validity but they are both important to the model. This holds for the measures of POSS (perception of social status) and PROF (English proficiency), although inspection of Table 6.1 reveals that each measure is responsive to its respective latent variable to the relatively same degree across four time points.

6.1.2 Standardized Validity Coefficient (λ_{ij}^s)

The second technique to assess validity is the standardized validity coefficient, which is “ λ_{ij} times the ratio of the standard deviations for the latent variable, ξ_j , and the observed variable, x_i that depends on it” (Bollen, 1989, p. 199). It gives the expected number of standard deviation units x_i changes for a one standard deviation change in ξ_j . The formula for λ_{ij}^s is,

$$\lambda_{ij}^s = \lambda_{ij} [\phi_{jj} / \text{VAR}(x_i)]^{1/2}$$

where ϕ_{jj} is the variance of ξ_j (that is, the covariance of ξ_j with itself) and $\text{VAR}(x_i)$ is the variance of x_i .

The standardized validity coefficient is preferable to the unstandardized validity coefficient in at least two cases. One is that focus is given to the relative validity of observed variables scaled in different ways; and the other is that one observed variable depends on two or more latent variables and the relative influence of the latent variables needs to be compared. Although the measurement model shown in Figure 4.5 does not fit into either case, the standardized validity coefficient is worthy of report. This is because

unlike the unstandardized validity coefficient, λ_{ij}^s has an upper limit on its varying range with values closer to one indicating higher validity and therefore, it is easier to interpret than λ_{ij} . Thus the standardized λ_{ij}^s coefficients in the model at each time point are presented in Table 6.3 below.

Table 6.3 *Estimates of Standardized Validity Coefficients for the Measurement Model*

<u>Latent variables</u>	<u>Time 1</u>	<u>Time 2</u>	<u>Time 3</u>	<u>Time 4</u>
POSS				
λ_{11}^s	.61	.70	.67	.73
λ_{21}^s	.85	.79	.81	.74
λ_{31}^s	.83	.74	.78	.84
PROF				
λ_{42}^s	.62	.73	.80	.81
λ_{52}^s	.58	.81	.81	.81
λ_{62}^s	.94	.82	.82	.83
EXPO				
λ_{73}^s	.58	.71	.71	.59
λ_{83}^s	.70	.72	.52	.62

Note. N = 97.

Examinations of the λ_{ij}^s values reported in Table 6.3 reveal moderately strong measures of all three latent variables, with the strongest indicator being the measure of λ_{62}^s (= .94) at Time 1 and the weakest indicator being the measures of λ_{52}^s and λ_{73}^s (= .58) at Time 1.

6.2 Reliability

6.2.1 A Brief Review of Classical Reliability Techniques

Reliability is the consistency of measurement. Much of the applied linguistics literature on reliability originates in classical measurement theory from psychology. The test-retest method, alternative forms, split-halves, and Cronbach's alpha are the four most popular techniques to estimate the reliability of measures. Unfortunately, however, none of these four techniques are appropriate to assess reliability of the measures shown in Figure 4.5, because several underlying assumptions are potentially violated. Specifically, it was hypothesized in the multiple-choice questionnaire that the true scores obtained from UBC-Rits and Kyoto-Rits students may change over time, which violates the assumption of the test-retest method that the true scores at two points in time are equal. Moreover, because of the fairly short format of the questionnaires with 12 items in the multiple-choice questionnaire, and with 20 items in the questionnaire on current uses of English, memory carry-over effects are likely to exist. Such effects counter what is assumed with the test-retest method, that is, uncorrelated measurement errors [$\text{COV}(e_t, e_{t+1}) = 0$, where e_t and e_{t+1} refer to the measurement errors at time t and $t + 1$, respectively].

The second technique to estimate reliability, alternate forms, is not operating at all because the same measures were used across four time points in data collection. The third technique, split-halves, has been criticized with respect to the arbitrariness in the way that the halves are allocated. Crocker and Algina (1986) point out that there are many ways to divide a set of items in half, and each split could lead to a different reliability estimate. The fourth measure, Cronbach's alpha, is the most popular reliability coefficient in the applied linguistics literature because it requires the least restrictive assumptions than the

other measures. Bollen (1989) points out, however, that Cronbach's alpha underestimates the reliability of congeneric measures as in the hypothesized model in Figure 4.5. A set of measures is said to be "congeneric" if each measure in the set purports to assess the same construct, except for measurement errors (Jöreskog, 1971b). For example, as indicated in Figure 4.5, x_1 , x_2 , and x_3 all served as measures of the latent variable POSS (perception of social status); they therefore represented a congeneric set of indicator variables. Taking these drawbacks of classical test theory into consideration, an alternative technique proposed by Bollen (1989) is employed to evaluate reliability of the measures ($x_1 - x_8$, $y_1 - y_{24}$) shown in Figure 4.5.

6.2.2 Squared Multiple Correlations (R^2_{xi})

The alternative reliability indicator employed is the squared multiple correlation for x_i , (R^2_{xi}), wherein the reliability of x_i is defined as "the magnitude of the direct relations that all variables (except δ 's) have on x_i " (Bollen, 1989, p. 221). This indicator allows correlated errors of measurement and observed variables depending on more than one latent variable. It can range from 0.00 to 1.00, thereby making its interpretation fairly easy. Indeed, values closer to one indicate higher reliability. Table 6.4 illustrates the reliability estimates of the measures at each time point in Figure 4.5.

Inspection of the R^2 values reported in Table 6.4 suggests that overall, the two latent variables POSS and PROF were represented by moderately strong measures, whereas the latent variable EXPO was represented by relatively weak measures with the weakest indicator being R^2_{83} ($= .27$) at Time 3. Interpretation of the R^2_{83} value indicates

Table 6.4 *Estimates of the Squared Multiple Correlations for the Measurement Model*

<u>Latent variables</u>	<u>Time 1</u>	<u>Time 2</u>	<u>Time 3</u>	<u>Time 4</u>
POSS				
R^2_{11}	.37	.49	.44	.54
R^2_{21}	.73	.63	.65	.54
R^2_{31}	.70	.54	.60	.70
PROF				
R^2_{42}	.38	.53	.64	.65
R^2_{52}	.34	.66	.66	.65
R^2_{62}	.89	.67	.67	.69
EXPO				
R^2_{73}	.33	.50	.50	.35
R^2_{83}	.49	.53	.27	.39

Note. N = 97.

that for this observed variable (exposure through receptive uses of English) at Time 3, only 27% of its variance was explained by the latent factor EXPO, and all else was error. The question raised here is how to deal with the variable which has low reliability. Should it be deleted from the hypothesized model? Loehlin (1998) points out that simply dropping a variable would produce a shift in the meaning of the latent variable which makes it unsuitable for testing the original theory. Thus the prudent stance is taken here: that is, the paths between the latent variable EXPO and its measures may be worth reassessing in future studies but should not be changed in the study.

6.3 Summary

In this chapter, the validity and reliability of the measures in the overall model shown in Figure 4.5 were estimated by a structural equations approach proposed by Bollen (1989). This approach is more general than the traditional validity and reliability test in that it works even when an observed variable has multiple latent causes or when

the error term for the observed variable correlates with other error terms. The estimated validity of the measures indicated the moderately strong relationship between the three latent variables and their respective measures. The estimated reliability, on the other hand, suggests that the reliability of the measures of POSS (perception of social status) and PROF (English proficiency) were moderately high, whereas the reliability of the measures of EXPO (exposure to English) was relatively low to moderate at best. One way of dealing with an observed variable which has very low validity and reliability is simply to drop it. However, taking such a step means the cease to operate in a confirmatory mode of analysis. Thus, with the original measurement model in hand, Study 1 in the next chapter uses confirmatory factor analysis to test Hypotheses 1 to 3 stated in Chapter III.

CHAPTER VII

STUDY 1: MODELING THE RELATIONSHIPS AMONG PERCEPTION OF SOCIAL STATUS, ENGLISH PROFICIENCY, AND EXPOSURE TO ENGLISH

7.0 Overview

This chapter focuses on Study 1 in which Hypotheses 1 to 3 shown in Chapter III were tested by analyzing the cause-effect relationships among the three latent variables, perception of social status (POSS), English proficiency (PROF), and exposure to English (EXPO) depicted in Figure 4.5. The analytic strategy used for this study consisted of three separate stages: (a) evaluation of the measurement model that specifies the pattern of relationships between the three latent variables and their respective observed variables, (b) assessment of the consistency of the measurement model across time, and (c) comparison of structural models that differed in the pattern of cause-effect relationships among the three latent variables. LISREL 8.30 (Jöreskog & Sörbom, 1999) was used to perform a confirmatory factor analysis (a special case of SEM with latent variables) at each stage of analysis. Findings and limitations of this study and implications for future L2 socialization research are postponed until Chapter IX.

7.1 Restatement of Hypotheses

To clarify how to test Hypotheses 1 to 3 within the context of a confirmatory factor analysis, they are restated in SEM terms here. Hypothesis 1, “The change over time

in Japanese students' perception of social status when giving advice in English is a consequence of the increase of their English proficiency" could be phrased as follows:

POSS at Time $t+1$ and PROF at Time t ($t = 1-3$) would show significant interrelationship. That is, γ_{12} , β_{42} , and β_{75} would be significant.

Hypothesis 2, "The change over time in Japanese students' perception of social status when giving advice in English is a consequence of the increase of their amount of exposure to English" could be phrased as follows:

POSS at Time $t+1$ and EXPO at Time t ($t = 1-3$) would show significant interrelationship. That is, γ_{13} , β_{43} , and β_{76} would be significant.

Hypothesis 3, "The change over time in Japanese students' perception of social status when giving advice in English is a consequence of the increase of their amount of exposure to English mediated by the increase of English proficiency" could be phrased as follows:

POSS at Time $t+2$ and EXPO at Time t via PROF at Time $t+1$ ($t = 1-2$) would show significant interrelationship. That is, $\beta_{42}\gamma_{23}$ mediated by η_2 and $\beta_{75}\beta_{53}$ mediated by η_5 would be significant.

These three restated hypotheses are all relevant to the structural part of the overall model that represents relationships among the three latent variables (POSS, PROF, and EXPO). It should be kept in mind, however, that analyses of the structural models can be meaningful when the measurement model that adequately fits the data is established.

7.2 Analyses

7.2.1 Evaluation of the Measurement Model

The sequence of analyses in SEM with latent variables begins with evaluation of the measurement part of the overall model that specifies the pattern of relationships between the latent variables and the observed variables (Hoyle & Smith, 1994). Because Study 1 was conducted based on a four-wave longitudinal design as depicted in Figure 4.5, of primary interest in the evaluation of the measurement model was assessing whether measures were sufficiently invariant across time to permit hypothesis testing in the structural part of the overall model (Pentz & Chou, 1994). The measurement invariance question could be phrased, Does the meaning of variable x remain the same over the course of the investigation? or, in SEM terms, Does the same measurement model hold for variable x at each measurement occasion? (Hoyle & Smith, 1994; Pentz & Chou, 1994). The following hypotheses (Farrell, 1994) were used to answer the measurement invariance question: (a) The factor loadings are identical across each time point; (b) the factor loadings and measurement error variances are identical across each time point; and (c) the factor loadings and the measurement error variances and covariances are identical across each time point. It should be noted that these hypotheses are placed in a sequence such that increasing levels of consistency across time are imposed. Failure to support a hypothesis at any point in the sequence is a failure to support that hypothesis and all hypotheses subsequent to it (Hoyle & Smith, 1994). Farrell (1994) summarizes the logic of invariance testing as follows:

This sequence of constraints can be imposed until the resulting model has a significantly poorer fit than the model that precedes it. At that point, the less

constrained model is retained and no further constraints are imposed. This process can be used to arrive at the most parsimonious model possible. (p. 481)

7.2.1.1 Selecting the Measurement Model

In testing the above hypotheses concerning the measurement invariance, sequential chi-square difference tests (Anderson & Gerbing, 1988) were used. In the first step, comparisons were made between the saturated overall model in which all parameters relating the constructs to one another were estimated (i.e., no constraints were imposed) and the saturated model in which equality was imposed on the factor loadings across four time points (e.g., $\lambda_{x21} = \lambda_{y21} = \lambda_{y104} = \lambda_{y187}$). Table 7.1 shows the results of the comparison between these two models.

Table 7.1 *Summary of Tests for Invariance of Factor Loadings*

Competing Models	χ^2	df	$\Delta\chi^2$	Δdf	RMSEA	CFI
1 No invariance imposed*	711.61	401	---	---	.090 (.079; .10)	.87
2 Model with factor loadings of POSS, PROF, and EXPO held invariant	721.27	416	9.66	15	.087 (.077; .098)	.87

Note. * The solution did not converge. N = 97.

It should be noted that degrees of freedom for χ^2 (chi-square) are

$$[(p + q)(p + q + 1)]/2 - t$$

where $p + q$ is the number of observed variables analyzed and t is the total number of independent parameters estimated (Jöreskog & Sörbom, 1996). Two goodness-of-fit indices shown in Table 7.1, the RMSEA and the CFI, denote the Root Mean Square Error of Approximation and the Comparative Fit Index, respectively. Because Steiger (1990) and MacCallum et al. (1996) have urged the use of confidence intervals to assess the precision of the RMSEA value, a 90% confidence interval around each RMSEA value was reported in a parenthesis.

Inspection of the results shown in Table 7.1 indicates that the difference between the two models was not significant-- $\chi^2 (15, N = 97) = 9.66, p > .80$, resulting in not rejecting the null hypothesis (Model 2 - Model 1 = 0). This finding suggests that the factor loadings, as a set, were not significantly different across the four time points. From the perspective of statistical parsimony, Model 2 with the factor loadings held invariant across time was selected for further analyses.

In the second step of the sequence of analyses, a more restrictive invariance hypothesis was tested. Specifically, the model including the invariant factor loadings (i.e., Model 2 selected in the first step) was compared with a model in which particular measurement error variances as well as the factor loadings were constrained to be identical across the four time points. Table 7.2 below exhibits the results of the hierarchical imposition of the invariance of the error variances. An examination of the results of sequential chi-square tests shown in Table 7.2 indicates that adding the equality imposition to measurement error variances made the model fit significantly worse. No additional constraints across time were therefore imposed and all further analyses were conducted on Model 2 in which the factor loadings were constrained to be identical

across the four time points and the measurement error variances were estimated individually at each time point. The LISREL input file related to the selected model is shown in Appendix F.

Table 7.2 *Summary of Tests for Invariance of Measurement Errors*

Competing Models	χ^2	df	$\Delta\chi^2$	Δdf	RMSEA	CFI
Model 2 with factor loadings of POSS, PROF, and EXPO held invariant	721.27	416	---	---	.087 (.077; .098)	.87
Model 2 with:						
• Measurement error variances of POSS, PROF, and EXPO invariant*	815.83	440	94.56**	24	.094 (.084; .10)	.84
• Measurement error variances of POSS and PROF invariant	773.26	434	51.99**	18	.090 (.080; .10)	.86
• Measurement error variances of PROF and EXPO invariant*	794.18	431	72.91**	15	.094 (.083; .10)	.84
• Measurement error variances of POSS and EXPO invariant*	786.08	431	64.81**	15	.093 (.082; .10)	.84
• Measurement error variances of POSS invariant	745.39	425	24.12**	9	.088 (.078; .099)	.87
• Measurement error variances of PROF invariant	750.71	425	29.44**	9	.089 (.079; .10)	.87
• Measurement error variances of EXPO invariant	769.05	422	47.78**	6	.093 (.082; .10)	.84

Note. * The solution did not converge. ** $p < .01$. N = 97.

7.2.1.2 Assessing the Fit of the Selected Measurement Model

Let us turn now to the goodness-of-fit statistics for Model 2, namely the saturated model including the invariant factor loadings across time. Farrell (1994) stated that the fit of the saturated model is extremely important in that all possible latent variable models are nested within it. Model A is said to be nested within Model B, when one or more parameters that are freely estimated in Model B are fixed at zero or constrained to have the same value in Model A (Anderson & Gerbing, 1988). Table 7.3 shows a summary of selected goodness-of-fit indices for Model 2. It is important to note here that the selection of goodness-of-fit indices were arbitrary with no intention that the indices used here were better than others. The rationale for the use of multiple indices is that one should avoid the decision depending on a specific index (Tanaka, 1993).

Table 7.3 *Summary of Selected Goodness-of-Fit Indices for Model 2*

χ^2	df	p-value	RMSEA	SRMR	PGFI	CFI
721.27	416	<.000	.087 (.077; .098)	.076	.54	.87

Note. N = 97. The SRMR and the PGFI denote the standardized Root Mean Square Residual and the Parsimony Goodness-of-Fit Index, respectively.

Given the known sensitivity of chi-square statistic to sample size (e.g., Cohen, 1990, 1994; Kirk, 1996), use of that index provides little guidance in determining the extent to which the model does not fit. Thus it is more beneficial to rely on fit as represented by the other indices shown in Table 7.3. Let us begin with Root Mean Square Error of Approximation (RMSEA). Model 2 showed the RMSEA value of .087, with the 90% confidence interval ranging from .077 to .098. According to Steiger's (1989)

guidelines for interpreting RMSEA values, those below .10 and .05 are considered to be “good” and “very good”, respectively. Moreover, Browne and Cudeck (1993) suggest that a model with a RMSEA greater than .10 not be employed. By either of these standards, it is safe to say that the fit of Model 2 to the data collected was not very good but marginally acceptable.

The conclusion drawn from inspection of the RMSEA value was consistent with the results of the other goodness-of-fit statistics shown in Table 7.3. Specifically, the standardized Root Mean Square Residual (SRMR) represents the average value across all standardized residuals and ranges from 0.0 to 1.00; in a well-fitting model this value will be small, say, .05 or less (Byrne, 1998). The standardized RMR value of .076 for Model 2 can be interpreted as meaning that the fit was marginal. The next index, the Parsimony Goodness-of-Fit Index (PGFI) takes into account the number of estimated parameters of the hypothesized model in the assessment of overall model fit (James, Mulaik, & Brett, 1982; Mulaik et al., 1989). Mulaik et al. (1989) suggest that nonsignificant chi-square statistics and goodness-of-fit indices in the range of .90, accompanied by parsimonious-fit indices in the range of .50, are not unexpected. By this standard, the PGFI value of .54 indicates the parsimoniously acceptable fit of Model 2 to the data. The last index shown in Table 7.3, the Comparative Fit Index (CFI) is a revised version of Bentler and Bonnett's (1980) Normed Fit Index (NFI) such that sample size is taken into account in assessing the model fit. This index provides a measure of complete covariation of a hypothesized model with the independence model,¹¹ a value > .90 indicating an

¹¹ The independence model is one of complete independence of all variables in the model (i.e., in which all correlations among variables are zero) and is the most restricted (Byrne, 1998).

acceptable fit to the data (Bentler, 1992). The CFI value of .87 reveals once again that the model fit was not wonderful but marginal.

Review of these criteria suggests that overall, the model fit was marginally acceptable, but some modification in specification may enable Model 2 to represent the data better. Inspection of modification indices provided by LISREL indicates that allowing a path between measurement errors of x_2 and x_3 (see Figure 4.5) would reduce the chi-square value by 22.34. As emphasized in the LISREL manual (Jöreskog & Sörbom, 1996), however, one should not just free paths blindly. Jöreskog (1993) pointed out that the specification of correlated error terms for purposes of achieving a better fitting model is not an acceptable practice; as with other parameters, such specification must be supported by a strong substantive rationale, empirical rationale, or both. Adhering to this caveat, further analyses were conducted on Model 2 without the correlated errors between x_2 and x_3 being included.

7.2.2 Comparison of Structural Models

Once the fit of the selected measurement model has been confirmed to be acceptable, focus is given next to the structural part of the model that is directly relevant to hypothesis testing. Five structural models were compared on the basis of the sequence of analyses proposed by Anderson and Gerbing (1988). Five models examined in this study included the saturated model (M_s) in which all parameters relating the latent variables to one another were freely estimated. M_s can be located at one end of the continuum concerning imposed restrictions. That is, M_s can be defined as the least constrained structural model. Obversely, a null model (M_n) in which all parameters

relating the latent variables to one another were fixed at zero can be located at the other end of the continuum. That is, M_n can be defined as the most constrained structural model. The researcher's theoretical model of interest (M_t), that is, the model including the hypothesized structural parameters shown in Figure 4.5, can be located in the middle of the continuum. A constrained model (M_c) can be defined as one in which a parameter estimated in M_t is constrained, whereas a unconstrained model (M_u) can be defined as one in which a parameter constrained in M_t is estimated. Given these definitions, the five models can be placed from most to least constrained in such a sequence as M_n , M_u , M_t , M_c , and M_s . It should be noted that the difference among the five models can be found only in the pattern of the structural paths, not in the measurement parts of the model; otherwise, any comparison of structural models would become invalid.

Sequential chi-square difference tests were used to compare the fit of each model to the data and to determine which structural model should be selected for further analyses. Each test can be framed as testing a null hypothesis of no significant difference between two nested structural models (e.g., $M_t - M_s = 0$). The sequence of tests was determined on the basis of the decision-tree framework proposed by Anderson and Gerbing (1988). First of all, comparison was made between M_t and M_s , because it provided an assessment of fit for the theoretical model of interest to the estimated construct covariances. The results of a chi-square difference test indicated that the comparison was not significant-- $\chi^2 (9, N = 97) = 2.41$, resulting in not rejecting the null hypothesis ($M_t - M_s = 0$).

Given the nonsignificant difference between M_t and M_s , M_c and M_t were compared next. M_c differed from M_t such that the effects of EXPO on PROF (i.e., γ_{23} , β_{53} ,

and β_{86}) were not included. This $M_c - M_t$ comparison was relevant to testing the significance of the indirect effects of EXPO on POSS via PROF. If the paths linking EXPO and PROF were significantly different from zero, then the indirect effects may exist. The results of a chi-square difference test indicate that M_t fit the data significantly better than M_c , $\chi^2(3, N = 97) = 76.74, p < .01$, resulting in rejecting the null hypothesis ($M_c - M_t = 0$). This finding suggests that the paths linking EXPO and POSS were meaningful and indirect effects of EXPO on POSS may be significant.

Because the $M_c - M_t$ comparison was significant, the $M_t - M_u$ comparison was assessed next. M_u differed from M_t in that the effects of PROF on EXPO (i.e., γ_{32} , β_{62} , and β_{95}) were freely estimated in M_u . By the $M_t - M_u$ comparison it was posited that higher English proficient students sought out more opportunities to be exposed to English. The results of a chi-square difference test indicates that M_u and M_t were not significantly different, $\chi^2(3, N = 97) = 4.84, p > .10$. This finding suggests that relaxing the next most likely constraint from a theoretical perspective in M_t did not significantly add to its explanation of the construct covariances. Moreover, as far as the preference for a more parsimonious model was concerned, M_t was the model to be accepted. Thus further analyses were conducted on the originally hypothesized theoretical model (M_t) as shown in Figure 4.5.

7.3 Results

Figure 7.1 illustrates on the next page standardized path coefficients representing the cross-time relationships among the latent variables in the selected structural model. Paths associated with significant coefficients at .05 level were represented by solid lines.

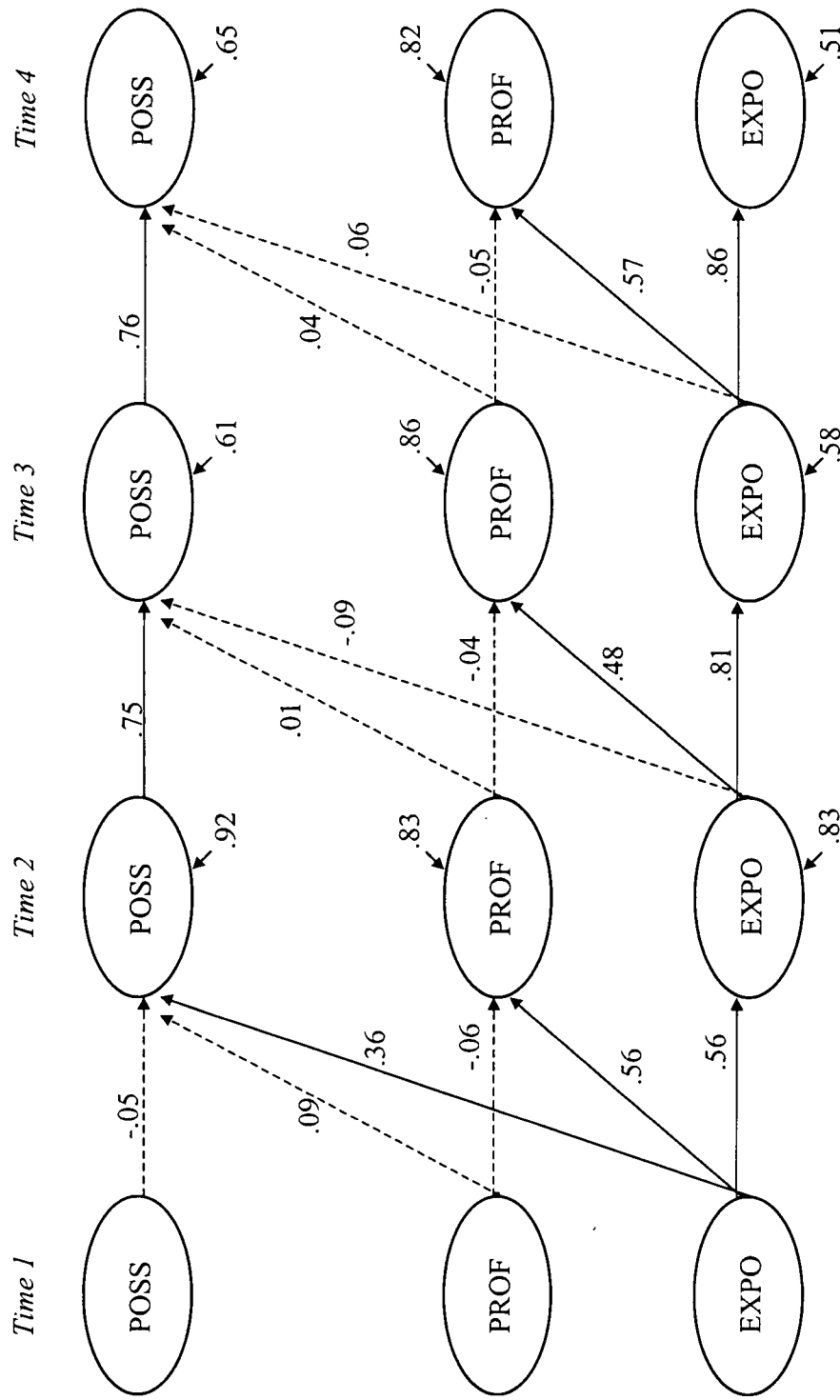


Figure 7.1 Standardized parameters representing the cross-time relationships among the latent variables in the final model.

Note. Paths associated with significant and nonsignificant coefficients are represented by solid and dashed lines, respectively. Parameters associated with the measurement model and within-time correlations among residuals are not illustrated to decrease the complexity of this figure.

To decrease the complexity of Figure 4.5, parameters associated with the measurement model and the within-time correlations among the residuals were not included in Figure 7.1. Inspection of path coefficients shown in Figure 7.1 reveals several characteristics of change over time in the three latent variables, POSS, PROF and EXPO.

POSS at Time 1 did not have much impact on POSS at Time 2, as represented by the dashed line with the value of the path coefficient being $-.05$. As for POSS at Time 2, EXPO showed significant direct impact on it, as illustrated by the solid line linking EXPO at Time 1 and POSS at Time 2 ($= .36$). This pattern, however, was not consistent at subsequent time points, as shown by the dashed line linking EXPO at Time 2 and POSS at Time 3 ($= -.09$) and the one linking EXPO at Time 3 and POSS at Time 4 ($= .06$). The impact of PROF on POSS, on the other hand, was very weak and nonsignificant, as shown in the dashed lines linking PROF at Time 1 and POSS at Time 2 ($= .09$), PROF at Time 2 and POSS at Time 3 ($= .01$), and PROF at Time 3 and POSS at Time 4 ($= .04$). An examination of these results suggests that support was not found for the relationship between EXPO and POSS, or between PROF and POSS. Thus Hypotheses 1 and 2 were both rejected.

Further examination of Figure 7.1 reveals that except for the relationship between POSS at Time 1 and POSS at Time 2, the autocorrelation effects (i.e., relationships between the same variables over time) were the strongest and most consistent effects in POSS. Two path coefficients linking POSS and subsequent levels of POSS were both significant, showing the high degree of stability (i.e., one of $.75$ and the other of $.76$). It was therefore suggested that change in POSS occurred sometime between Time 1 and Time 2, and such altered POSS was upheld for the periods from Time 2 to Time 4.

As for PROF, autocorrelation effects as seen in POSS were very weak, and none of the path coefficients linking PROF and the subsequent level of PROF were significant, as illustrated by the dashed lines (i.e., the one of $-.06$, the second of $-.04$, and the third of $-.05$). In contrast, the effects of EXPO on PROF were all significant, showing the moderately strong impact on PROF across the four time points (i.e., all paths linking PROF and EXPO were $.48$ or above). Moreover, the paths ranged from $.48$ to $.57$, indicating the moderate degree of stability over time. Given that the relationship between PROF and EXPO were moderately strong, and that the effects of PROF on POSS were very weak and nonsignificant, support was not found for the indirect effects of EXPO on POSS via PROF. In fact, the indirect effect of EXPO at Time 1 on POSS at Time 3 via PROF at Time 2 was $.01$, and the other indirect effect of EXPO at Time 2 on POSS at Time 4 via PROF at Time 3 was $.02$. Thus Hypothesis 3 was rejected.

Figure 7.1 also reports the standardized residuals for each endogenous variable. Farrell (1994) explains how to interpret standardized residuals as follows:

Squaring these provides an estimate of the proportion of variance in each endogenous variable not predicted by the model. Alternatively, subtracting the squared values from 1.00 indicates the proportion of variance predicted by the model. (p. 484)

These coefficients shown in Figure 7.1 reveal that the model accounted for 15% to 62% of the variance in POSS, 26% to 31% of the variance in PROF, and 31% to 73% of the variance in EXPO.

7.4 Summary

This chapter has focused on Study 1 that was designed to test Hypotheses 1, 2, and 3 discussed in Chapter III. A confirmatory factor analysis was performed at three separate stages of data analysis, namely evaluation of the measurement model, evaluation of the measurement model across time, and comparison of the structural models. Inspection of the results indicated that change in UBC-Rits students' perception of social status occurred at the early stage of studying abroad, although it was not a function of English proficiency or amount of exposure to English. All three hypotheses were therefore rejected. The results shown in this chapter are elaborated upon from L2 socialization perspectives in Chapter IX, in conjunction with results shown in Study 2 which is discussed in the next chapter.

CHAPTER VIII

STUDY 2 : TESTING FOR INVARIANT LATENT MEAN STRUCTURES

8.0 Overview

This chapter focuses on Study 2 that was designed to assess Hypothesis 4 and Research questions 1 and 2 discussed in Chapter III. Study 2 aimed to investigate the impact of L2 learning environment on pragmatic development, that is, how the route and rate of pragmatic development differs between an ESL group and an EFL group. To this end, UBC-Rits group in an ESL environment and Kyoto-Rits group in an EFL environment were compared using a multigroup structured latent means model within the framework of LISREL 8.30.

8.1 Basic Concepts Underlying Tests of Latent Means

In the multigroup comparisons using statistical techniques such as ANOVA, focus is given to the extent to which the differences among the means of the *observed* variables representing the groups are statistically significant. As can be seen in Figure 4.5, however, looking solely at the observed variable means may be problematic because they are functions of the other parameters in the model. An advantage of using the structured latent means model is that the focus is on the means of *latent* variables rather than observed variables. More specifically, the means of latent variables derive not only from the means of observed variables but also from the structured coefficients in the model. Byrne (1998) states, "The intent is to test for the equivalence of means related to each underlying construct or factor" (p. 304). Indeed, applications of the structured latent

means model involve testing the model fit simultaneously across two or more groups. The application to be discussed in this chapter is to test for group differences in the means of the latent variable POSS (i.e., perception of social status when giving advice in English).

8.2 Evaluating the Baseline Model

The first step in multigroup comparisons is to assess the goodness-of-fit of the hypothesized model separately for each group. This is because any discussion of latent mean differences is problematic if the measures and the structure of the construct under study are not equivalent across groups (Alwin & Jackson, 1981; Byrne, 1988). Figure 8.1 on the next page represents the hypothesized model. It should be noted that it included serial correlations among measurement errors not only at adjacent time points (e.g., $e_{x1} - e_{x4}$) but also at non-adjacent time points (e.g., $e_{x1} - e_{x7}$, $e_{x1} - e_{x10}$). This is because repeated measurement of the same variable often results in correlated measurement errors (Judd & Milburn, 1980; Kessler & Greenberg, 1981).

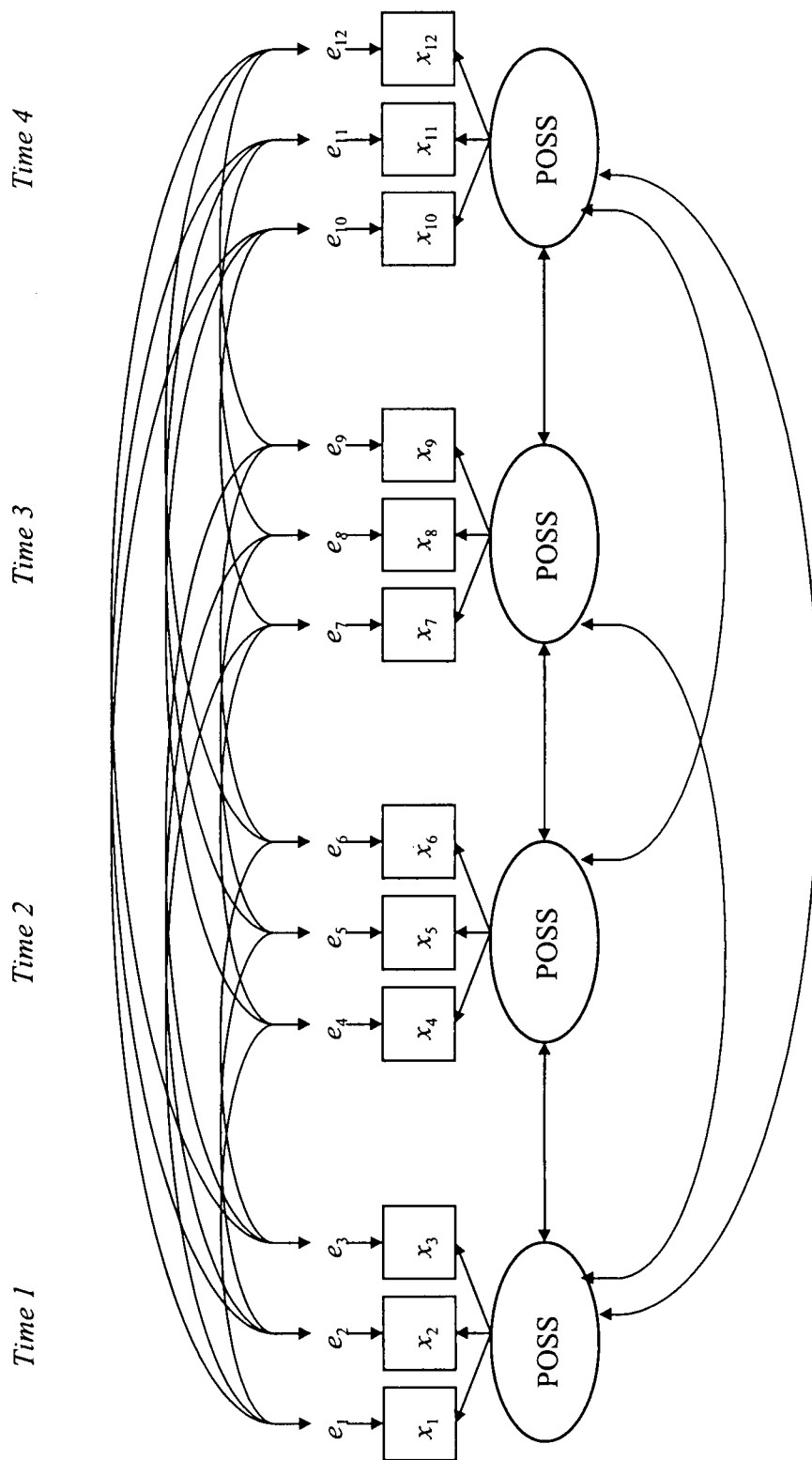


Figure 8.1 Baseline model of change over time in POSS (perception of social status).

Table 8.1 exhibits the results of the fit statistics for the hypothesized model shown in Figure 8.1.

Table 8.1 *Summary of Selected Goodness-of-fit Indices for the Hypothesized Model*

Group	χ^2	df	p-value	RMSEA	SRMR	GFI	CFI
UBC-Rits Group	46.82	29	.019	.080 (.033; .12)	.042	.92	.98
Kyoto-Rits Group	28.83	29	.474	.000 (0.0; .075)	.014	.95	1.00

Note. N = 97 for UBC-Rits group, and 102 for Kyoto-Rits group. The RMSEA, SRMR, GFI, and CFI denote the Root Mean Square Error of Approximation, the standardized Root Mean Square Residual, the Goodness-of-Fit Index, and the Comparative Fit Index, respectively. A 90% confidence interval around each RMSEA value is reported in a parenthesis.

As with examples in the previous chapter, selection of goodness-of-fit indices was arbitrary with no intention that the indices used here were better than others. Inspection of the values of the SRMR, GFI, and CFI shown in Table 8.1 suggests that the hypothesized model fit the data fairly well for both Kyoto-Rits and UBC-Rits groups. On the other hand, an examination of the confidence interval of the RMSEA value for UBC-Rits group indicates the possibility of the misspecification of the model. A review of modification indices provided by LISREL 8.30 suggests that the incorporation of two measurement error covariances (i.e., the one between x_1 and x_2 , and the other between x_2 and x_3) into the model could result in substantive drops in the chi-square value. However, there were no substantive theoretical and empirical rationales for adding those covariances to the model. Thus the initially hypothesized model shown in Figure 8.1 was selected as the baseline model that the two groups shared.

8.3 The Logic of the Structured Latent Means Model

Once the baseline model has been selected, the next step is to transform it into a model that represents latent mean structures. Figure 8.2 on the next page illustrates the latent mean structures model used in the present study. Several features of the model and technical terms used in Figure 8.2 may be worthy of a brief explanation here. First, the τ s (taus) and κ s (kappas) represent the regression coefficients of the observed variables onto the constant and the regression coefficients of the latent variables onto the constant, respectively. Byrne (1998) stated that factor intercepts (κ s) for one group is fixed to zero and therefore, this group operates as a reference group against which latent means for the other group are compared. In other words, factor intercepts are interpretable only in a relative sense.

Second, CONSTANT enclosed in rectangles can be defined as a dummy variable that “provides the mechanism for parameterizing the necessary intercepts in the model and, thus, plays a key function in the estimation of latent mean values; its variance remains constrained to zero” (Byrne, 1998, p. 308).

A third point to be noted here is that the multigroup comparison with respect to the means of latent variables is first performed on a model in which all λ s are constrained equally across groups, all intercepts for the observed variables (i.e., τ s) are constrained equally across groups, variance associated with the CONSTANT remains fixed to 1.00, and all factor intercepts (κ s) are freely estimated in one group and constrained equally to zero in the other group.¹²

¹² For further understandings of the logic of this statistical technique, see Byrne (1998).

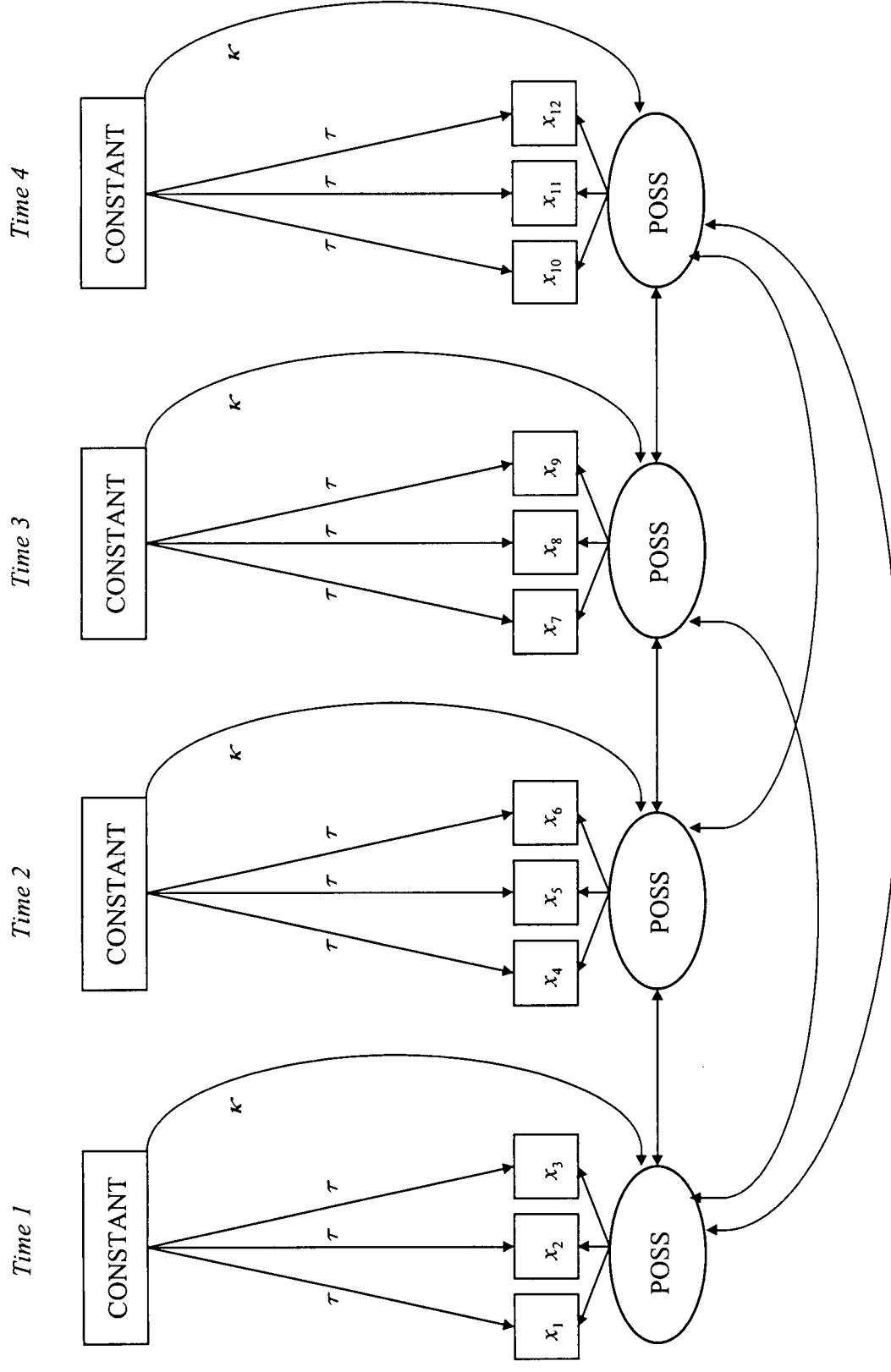


Figure 8.2 Latent mean structures model.

Note. Measurement errors and serial correlations among them were not illustrated to decrease the complexity of this figure.

In the structured latent means model used in this study, the Kyoto-Rits group was defined as the reference group by fixing its Kappa matrix to zero. Thus the primary focus was on estimating the Kappa values for UBC-Rits group that represented latent mean differences between the two groups.

8.4 Evaluating the Structured Latent Means Model

Table 8.2 reports the results of the goodness-of-fit statistics for the structured latent means model shown in Figure 8.2.

Table 8.2 *Summary of Selected Goodness-of-fit Indices for the Structured Means Model*

χ^2	<i>df</i>	<i>p</i> -value	RMSEA	SRMR	GFI	CFI
141.18	76	.000	.094 (.069; .12)	.033	.95	.97

Note. N = 199.

A review of information reported in Table 8.2 reveals that the model fit was marginally acceptable. Although the SRMR, GFI, and CFI values indicated the fairly good fit of the model to the observed data, the 90% confidence interval around the RMSEA value exceeded .10, that is, the upper bound of the acceptable fit. This finding suggests that the equality constraints imposed on both all factor loadings and the variable intercepts across the two groups may be excessively stringent. As with the invariance testing strategy used in Study 1, the initially hypothesized model was compared next with the less constrained model in which the factor loadings of CJ's were estimated freely at each group. The rationale for relaxing those loadings was that the degree of change over

time in the observed means of CJ's differed substantially between the two groups (see Table 5.1 and 5.2 in Chapter V). This phenomenon might be indicative of increasing differentiation as part of developmental change in perception of social status only among UBC-Rits students. If so, imposing equality across the two groups on those loadings would be unrealistic. Table 8.3 exhibits the results of the goodness-of-fit statistics for the less constrained model.

Table 8.3 *Summary of Selected Goodness-of-fit Indices for the Less Constrained Structured Means Model*

χ^2	df	p-value	RMSEA	SRMR	GFI	CFI
130.37	72	.000	.091 (.065; .12)	.029	.95	.98

Note. N = 199.

The results of a chi-square difference test between the less constrained model and the initially hypothesized theoretical model indicated that the former fit the data significantly better-- χ^2 (4, N=199) = 10.81, $p < .05$, although the RMSEA value shown in Table 8.2 indicated slight improvement of the fit.

Given the possibility of the misfit suggested by the upper bound of the RMSEA value, the model examined next was the one in which the factor loadings XL's as well as CJ's were freely estimated at each group and the factor loadings of PD's were fixed to one. Table 8.4 displays the results of the goodness-of-fit statistics for this model.

The results of a chi-square difference test between the model in Table 8.3 and the model in Table 8.4 indicated that the latter fit the data significantly better-- χ^2 (4, N = 199)

= 14.46, $p < .01$. Moreover, the RMSEA value dropped substantially, although the upper bound of the confidence interval (= .11) was still beyond the range of the acceptable fit.

Table 8.4 *Summary of Selected Goodness-of-fit Indices for the Structured Means Model Involving the Variant Factor Loadings of CJ's and XL's*

χ^2	df	p-value	RMSEA	SRMR	GFI	CFI
115.91	68	.000	.085 (.040; .11)	.015	.95	.98

Note. N = 199.

To seek the most parsimonious model, the model in Table 8.4 was compared next to the model in which equality between groups was imposed on the factor loadings of CJ's and XL's at Time 1 but not on those at Time 2, Time 3 or Time 4. Table 8.5 shows the results of the goodness-of-fit statistics.

Table 8.5 *Summary of Selected Goodness-of-fit Indices for the Structured Means Model Involving the Invariant Factor Loadings of CJ's and XL's at Time 1*

χ^2	df	p-value	RMSEA	SRMR	GFI	CFI
116.44	70	.000	.082 (.055; .11)	.016	.95	.98

Note. N = 199.

The results of the chi-square difference test between the models in Tables 8.4 and 8.5 indicated that there was no significant difference between the two models-- χ^2 (2, N=199) = .53, $p > .70$. Moreover, the RMSEA value in Table 8.5 indicated the better fit of the model to the observed data than that in Table 8.4. Theoretically, measurement

invariance at Time 1 is appropriate because the observed variables at baseline (i.e., Time 1) were expected to be equally valid indicators of the latent variables for each group and because intervention (i.e., UBC-Rits students' studying abroad during the period from Time 2 to Time 4) were expected to change the means and variance-covariance structures of the latent variables.

In pursuit of the most parsimonious model, the model in Table 8.5 was compared next to the model in which equality between the two groups were imposed on the factor loadings of CJ's and XL's at Time 1 and Time 2. This invariance testing posited that the intervention effect may not appear until Time 3. Table 8.6 shows the results of the goodness-of-fit statistics.

Table 8.6 *Summary of Selected Goodness-of-fit Indices for the Structured Means Model Involving the Invariant Factor Loadings of CJ's and XL's at Time 1 and Time 2*

χ^2	<i>df</i>	<i>p</i> -value	RMSEA	SRMR	GFI	CFI
124.21	72	.000	.086 (.060; .11)	.033	.95	.98

Note. N = 199.

The results of the chi-square difference test revealed the significant difference between the models in Tables 8.5 and 8.6-- $\chi^2 (2, N = 199) = 7.77, p < .05$. Furthermore, the values of RMSEA and SRMR shown in Table 8.6 also indicated that the less constrained model (the model in Table 8.5) was fitting better than the constrained model (the model in Table 8.6), suggesting that equality imposition at both Time 1 and Time 2 were too stringent. No further constraints were therefore imposed on the model and

further analyses were conducted on the model that involved the invariant factor loadings of CJ and XL at Time 1 only.

In sum, interpretation of the results of a series of chi-square difference tests suggests that the factor loadings of CJ's and XL's were significantly different between UBC-Rits and Kyoto-Rits groups at Times 2, 3, and 4. As mentioned above, this conclusion can be supported because the former group had lived in an ESL environment since Time 2. Methodologically, the conclusion was not an unexpected phenomenon because in longitudinal research involving subjects and theoretical constructs that are expected to change over time, total measurement invariance may be an unrealistic goal and partial invariance may be an acceptable goal (Byrne, Shavelson, & Muthen, 1989; Pentz & Chou, 1994). It was thus reasonable to conduct further analyses of latent means on the model in which the factor loadings of CJ and XL at Time 1 were constrained to be equal across the two group and those of CJ's and XL's at Times 2, 3, and 4 were estimated freely at each group.

8.5 Assessing the Latent Means

To answer the question of whether the latent variable means were significantly different for UBC-Rits and Kyoto-Rits groups, estimates of Kappa parameters for the former group is reported in Table 8.7. The LISREL input file related to this analysis is shown in Appendix G.

Table 8.7 *Summary of Estimates of the Kappa Values in the Final Structured Latent Means Model*

Time	Kappa	<i>t</i> -value
1	-.46	-2.24*
2	.27	1.39
3	1.24	5.70*
4	1.55	6.41*

Note. * $p < .05$.

As explained earlier in this chapter, the values reported in Table 8.7 represent latent mean differences between the UBC-Rits and Kyoto-Rits groups. The Kyoto-Rits group was designated as the reference group and therefore, the Kappa parameters for the group were fixed to zero. Inspection of the Kappa values shown in Table 8.7 reveals that the latent means were statistically different between the two groups at Times 1, 3 and 4, as indicated by *t*-values reported together with Kappa values. Given the negative value of the Kappa parameter at Time 1, it can be said that at Time 1, UBC-Rits students had significantly lower levels of pragmatic competence with respect to perception of social status when giving advice in English, than did Kyoto-Rits students. It was also revealed that at Time 2, there was little difference in the level of pragmatic competence between the two groups, and that as time further went by, UBC-Rits students came to show significantly higher levels of pragmatic competence than Kyoto-Rits students. Given these findings, Hypothesis 3 stated in Chapter III was not rejected.

To answer research questions addressed in Chapter III, let us review Table 8.8 in which parameter estimates for UBC-Rits and Kyoto-Rits groups are summarized. It

should be noted that the values at Time 1 were identical between the two groups because equality was imposed on those parameters in the selected model.

Table 8.8 *Parameter Estimates for UBC-Rits and Kyoto-Rits Groups*

	<u>UBC-Rits Group</u>				<u>Kyoto-Rits Group</u>			
	<u>T1</u>	<u>POSS</u> <u>T2</u>	<u>T3</u>	<u>T4</u>	<u>T1</u>	<u>POSS</u> <u>T2</u>	<u>T3</u>	<u>T4</u>
<u>Parameters</u>								
λ_1	1.00	0	0	0	1.00	0	0	0
λ_2	.97	0	0	0	.97	0	0	0
λ_3	.94	0	0	0	.94	0	0	0
λ_4	0	1.00	0	0	0	1.00	0	0
λ_5	0	1.49	0	0	0	.83	0	0
λ_6	0	1.26	0	0	0	.90	0	0
λ_7	0	0	1.00	0	0	0	1.00	0
λ_8	0	0	1.85	0	0	0	.94	0
λ_9	0	0	1.58	0	0	0	1.03	0
λ_{10}	0	0	0	1.00	0	0	0	1.00
λ_{11}	0	0	0	1.76	0	0	0	.99
λ_{12}	0	0	0	1.71	0	0	0	1.00

Note. Unstandardized solution; all zero values represents fixed parameters. T1, T2, T3, and T4 represent Times 1, 2, 3, and 4, respectively. λ_1 , λ_4 , λ_7 , and λ_{10} are associated with indicators of PD (x_1 , x_4 , x_7 , and x_{10}), λ_2 , λ_5 , λ_8 , and λ_{11} are associated with indicators of CJ (x_2 , x_5 , x_8 , and x_{11}), and λ_3 , λ_6 , λ_9 , and λ_{12} are associated with indicators of XL (x_3 , x_6 , x_9 , and x_{12}). N=199.

Inspection of Table 8.8 reveals that the factor loadings of the measures at Times 2, 3, and 4 differed between the two groups. Given that the loadings of CJ's at Times 2, 3, and 4 differed substantially between the two groups (i.e., 1.49 vs. .83 at Time 2; 1.85 vs. .94 at Time 3; and 1.76 vs. .99 at Time 4), and given the above-mentioned finding on the basis of the Kappa values--that UBC-Rits students came to show increasingly and significantly higher levels of pragmatic competence than Kyoto-Rits students, it can be

said that especially in the scenarios relevant to CJ, UBC-Rits students came to show the same preferences for advice type as native speakers of English. This finding was consistent with change in the observed means of CJ's as shown in Table 5.1 in Chapter V. The similar story holds true for UBC-Rits students' perception of XL, although the difference in the factor loadings of XL's between the two groups was not as big as that in CJ's. These findings do not imply, however, that UBC-Rits students were less competent in the scenarios relevant to PD than Kyoto-Rits students. On the contrary, as represented by high observed means shown in Table 5.1 in Chapter V (i.e., PD1 = 12.13), UBC-Rits students' perception of PD were already similar to English native speakers' at Time 1. This is true for Kyoto-Rits students as displayed by high observed means shown in Table 5.2 (i.e., PD1 = 12.66). Unfortunately, these findings on the basis of the observed means cannot be confirmed with respect to the factor loadings of PD's because as shown in Table 8.8, the parameters depending on PD's were fixed to one for statistical identification in the present model.

8.6 Summary

Study 2 discussed in this chapter attempted to compare the different levels of pragmatic competence that resulted from study in an EFL and an ESL environments. The results based on the structured latent means model revealed that there was an impact of living and studying in the target speech community on pragmatic competence to give advice to equal-status (CJ) and lower-status (XL) persons. The results also revealed that when the subjects were in Japan, students had pragmatic competence to give advice to higher-status (PD) persons, although this finding was confirmed only at the observed

means level. The results shown in Study 2 are interpreted from L2 socialization perspectives in the next chapter, in connection with those shown in Study 1.

CHAPTER IX

CONCLUSION

9.0 Overview

This chapter summarizes this dissertation. The results from Study 1 and Study 2 are elaborated upon within a L2 socialization perspective accompanied by an appraisal of the approach to language instruction in Japan discussed in Chapter I. Limitations of the studies are discussed and implications for further research into L2 socialization conclude the chapter.

9.1 Summary

9.1.1 Purpose

The present study focused on changes over time in university-level Japanese students' sociocultural perceptions of social status during their year abroad in Canada, and the impact of such changes at subsequent time points. The sociocultural perception examined was perceived "social status" which Brown and Levinson (1987) suggested was a contributory factor in the perception of social asymmetry, power and authority. The study attempted to examine (1) whether (and to what extent) Japanese students, before they came to study in Canada, had recognized English native speakers' understanding of social status and had learned how to offer advice appropriately in English to individuals of various social statuses, (2) what proportion of differential pragmatic development among Japanese students in Canada was accounted for by their *English proficiency* and *amount of exposure to English*, and (3) whether (and to what extent) living and studying

in Canada facilitated Japanese students' pragmatic development, which was assessed by the degree of approximation to native speech act behavior in various advice-giving situations repeated over the course of an academic year. To this end, the study compared the development of Japanese exchange students' pragmatic competence during their year abroad in Canada with peers in Japan who did not undertake a year abroad.

9.1.2 Theoretical Background

Over the last decade Schieffelin and Ochs' (1986) language socialization model, developed to study children's first language (L1) acquisition within their own culture, has been applied to various English-as-a-second-language (ESL) contexts within a largely qualitative research tradition, centered on case studies. The model relates second language (L2) acquisition to the sociocultural competence that L2 learners acquire over time. For various methodological reasons, however, previous studies have revealed little about the characteristics of the L2 socialization process. Indeed, the primary methodological problem is that few studies have been designed to examine the extent to which L2 learners have acquired pragmatic competence before they enter the target speech community. A second problem is that previous L2 socialization studies have adopted a taken-for-granted view of culture as the basis for interpretation and explanation of a L2 learner's culture of origin and the target culture. A third problem is that few studies have employed an adequate number of subjects to examine 'intracultural variance.' A fourth problem is that few studies have explored L2 socialization that takes place in L2 learners' home countries. The present study was designed to begin to address these problems.

9.1.3 Methodology

9.1.3.1 Subjects

The subjects consisted of two groups enrolled in the same Japanese university: one group of 97 students who came to the University of British Columbia to study for eight months in an English immersion environment (called UBC-Rits students) and the other of 102 students who continued to study in Japan (called Kyoto-Rits students).

9.1.3.2 Data Collection

The researcher tracked the groups from the period prior to the departure of one group for Canada through its return to Japan. In-class questionnaires, designed to focus on learners' preferences for resolving problems requiring giving advice to individuals of various social statuses (i.e., higher status, status equal, and lower status), were administered four times (July, October, January, and April) during the academic year. The same data collection procedures were used in both Japan and Canada. The questionnaire on uses of English that was designed to obtain information about the amount of exposure to English and English proficiency was administered to UBC-Rits students four times during the academic year.

9.1.4 Analyses

Hypotheses and research questions were:

Hypotheses 1:

The change over time in Japanese students' perception of social status when giving advice in English is a consequence of the increase of their English proficiency.

Hypothesis 2:

The change over time in Japanese students' perception of social status when giving advice in English is a consequence of the increase of their amount of exposure to English.

Hypothesis 3:

The change over time in Japanese students' perception of social status when giving advice in English is a consequence of the increase of their amount of exposure to English mediated by the increase of English proficiency.

Hypothesis 4:

The Japanese students studying in the target speech community come to show increasingly and significantly higher levels of pragmatic competence to offer advice in English than those studying in Japan.

Research question 1:

Do the students studying in the target speech community come to show the same preferences for advice type as native speakers of English, depending on the status relationship of the conversational participants?

Research question 2:

Do the students studying in Japan come to show the same preferences for advice type as native speakers of English, depending on the status relationship of the conversational participants?

Structural Equation Modeling (SEM) with latent variables based on a four-wave longitudinal design was used to test Hypotheses 1, 2, and 3, whereas a Multi-group Structured Latent Means Model was used to assess Hypotheses 4 and Research Questions 1 and 2.

9.1.5 Results

9.1.5.1 Results of Study 1

Study 1 sought to examine the relationships among UBC-Rits students' perception of social status when giving advice in English, English proficiency, and amount of exposure to English. Hypotheses 1, 2, and 3 posited that change in their perception would be functions of the other two factors. All three hypotheses were rejected, and it was revealed that change in their perception of social status occurred at the early stage of studying abroad, sometime between Time 1 (when they were in Japan) and Time 2 (when they spent two months in Canada) and such altered perception continued to affect their perception until the end of their stay in the target speech community.

9.1.5.2 Results of Study 2

Study 2 examined the impact of living and studying in the target speech community on pragmatic development, while assessing Hypothesis 4 and Research

Questions 1 and 2. For the purpose of the study, UBC-Rits group in an ESL environment and Kyoto-Rits group in an EFL environment were compared. The difference in means of a latent variable (perception of social status) indicated that when both groups were in Japan, UBC-Rits students had significantly lower levels of pragmatic competence to offer advice appropriately in English to individuals of various social status, than did Kyoto-Rits students. As time went by, however, UBC-Rits students came to show increasingly and significantly higher levels of pragmatic competence. Thus Hypothesis 4 was not rejected. Moreover, as represented by the results of the measurement invariance testing, the drastic change observed among UBC-Rits students' perception of social status occurred sometime between Time 1 and Time 2, that is, the early stage of their studying abroad. This finding is consistent with what was observed in Study 1.

As for Research Questions 1 and 2, UBC-Rits group came to show similar preferences as native English speakers when giving advice to lower-status and status-equal persons. This was not true for the Kyoto-Rits group. As far as advice-giving to higher-status persons is concerned, both UBC-Rits and Kyoto-Rits groups showed similar preferences as native speakers during the entire observation period.

9.2 Interpreting the Results from L2 Socialization Perspectives

The results of Study 1 and Study 2 were consistent in that L2 socialization as evidenced by change in UBC-Rits students' perception of social status (i.e., their increasing understanding of how English native speakers perceive social status) occurred by the time they had spent two months in Canada. Indeed, two major questions are posed here. What caused that change? Why did their perception of social status change so soon

after their arrival in the target speech community? Inspection of the results shown in Figure 7.1 and Table 8.8 provides some clues to explain these findings.

The studies examined whether (and to what extent) Japanese students in an academic exchange program, before they came to study in Canada, had learned the target sociocultural rules of offering advice through communication-based classes in school. They had learned English in the communication-based classes that were designed to enhance their pragmatic competence. Perhaps, they had acquired, to a certain extent, how to offer advice appropriately in English to higher, equal, or lower status persons and had understood how English native speakers perceive social status. If their understanding had reached the level to allow them to function efficiently in the target speech community, a strong interrelationship would have been observed between their perception of social status observed in Japan and in Canada (see Figure 7.1). The reality was, however, that the significant interrelationship was observed between the amount of exposure to English observed in Japan in July and their perception of social status observed in Canada in October. Interpretation of this finding suggests that the students who sought out more opportunities to be exposed to English even when they were in Japan had acquired a higher level of the competence to give advice appropriately to individuals of various social statuses. Put another way, the competence acquired through communication-based classes in Japan alone was not sufficient for the students to function competently in Canada and perhaps, other extra exposure than that received in the classes was necessary to become competent at the early stage of their study abroad. That is, L2 socialization occurred among the students who were eager to be exposed to English even when they

were in Japan. Thus the change during the early stage of their studies abroad was accounted for partly by the effect of their perceptions of social status.

The study attempted to account for differential pragmatic development among Japanese students in a target speech community as functions of their English proficiency as well as the amount of exposure to English. It should be kept in mind, however, that because only 15% of the variance in perception of social status at Time 2 was accounted for by the hypothesized model shown in Figure 7.1, it is highly likely that there were some other direct or indirect factors that caused the change. This finding gives rise to some speculation. Given that the students who tried to be exposed to English even in Japan were likely to be highly motivated to learn English, then motivation might have been a better indicator of the change.

It should also be noted that although the students received more exposure to English in Canada than in Japan as shown in Table 5.1, the amount of exposure was significantly associated only with their levels of English proficiency but not with their perception of social status at all while they were in Canada (see Figure 7.1). Given these nonsignificant interrelationships between amount of exposure to English and the students' perception of social status in Canada, it was not supported that the more exposure to English they received, the higher the level of their understanding of English native speakers' perception of social status. What factor, then contributed to the UBC-Rits students' increasingly and significantly higher levels of understanding of social status as represented by latent mean difference between UBC-Rits and Kyoto-Rits groups shown in Table 8.7? A key to answer this question is related to the following question: In what respect, other than learning environments, did these two groups differ? One

possibility is that UBC-Rits students might have been more motivated to learn English than the Kyoto-Rits students. Once again, motivation is likely associated with the pragmatic development of UBC-Rits students, although this remains speculation in the present study.

This study also examined whether (and to what extent) living and studying in the target speech community facilitated Japanese students' pragmatic development, which was assessed by the degree of approximation to native speech act behavior in various advice-giving situations repeated during the course of an academic year. Inspection of their preferences for advice type in each status relationship shown in Table 8.8 revealed that UBC-Rits group came to show the same preferences as native speakers of English when giving advice to lower-status and equal-status persons. An examination of the observed means suggested that across all four time points both UBC-Rits and Kyoto-Rits groups had similar preferences for advice type as native speakers when offering advice to higher-status persons. These findings relevant to their preferences for advice type in each status relationship contradict what was found in a preliminary study--that the students in the exchange program did not give advice in English to higher-status individuals in a socially appropriate manner. What was revealed in the present study was that the L2 socialization that took place in the target speech community was evidenced by change in their understanding of English native speakers' perception of status-equal and lower-status persons, and L2 socialization that had taken place in their home country was represented by the acquisition of the pragmatic competence to offer advice to higher-status persons. The question is, how did they acquire the competence concerning higher-status persons in Japan? There are several possibilities. First, similar perceptions of

higher social status are shared between English and Japanese native speakers. It might be the case that they had acquired the pragmatic competence naturally in the L1 socialization process and had applied it to L2 socializing contexts in the target speech community. Second, they may have acquired the competence through their communication-based classes, although English natives and Japanese natives have different perceptions of higher social status. Given that Japanese people tend to use polite expressions when talking to individuals of higher status, the first possibility is more likely than the second.

From a methodological point of view, the findings in Study 1 and Study 2 verified the importance of the modified longitudinal research design in which data collection begins before the subjects enter the target speech community. If the subjects had been observed in the target speech community only, that is, if focus had been given exclusively to synchronic L2 socialization in the target speech community, the changes that occurred at the early stage of their studying abroad would not have been observed, so that different conclusions would have been drawn. In other words, the importance of incorporating a diachronic perspective into L2 socialization research was confirmed in the present studies.

As a result of employing a relatively large number of subjects, the studies were able to illuminate the variance of the three latent variables, namely perception of social status, English proficiency and amount of exposure to English and the interrelationships among them. The results of the studies indicated the risk of making unsubstantiated generalization of the findings from a small sample to a population. Furthermore, the studies demonstrated the importance of employing a reference group in L2 learners' countries of origin in order to clarify the L2 socialization process.

The methodology used in the present studies was established to begin to overcome weaknesses of frequently used qualitative approaches to L2 socialization. It should be noted, however, that there is no intention of dismissing the findings of all qualitative studies. However, it is clear that qualitative research approaches are not sufficient to elaborate upon or to generalize about the dynamic quality of the L2 socialization process. Observing the same socialization events from both a qualitative and a quantitative standpoint must remain an innovative venture in L2 socialization research.

9.3 Limitations

There are several limitations to Study 1 and Study 2. First, given the complexity of the model as shown in Figure 4.5, the number of subjects was too small. Although SEM with latent variables is useful in analyzing longitudinal data as discussed in Chapter IV, large samples are strongly recommended to reduce the bias in estimating parameters. With data from over 200 subjects, parameter estimates in the complex model as shown in Figure 4.5 would become more reliable. Second, since the subjects were sampled in a non-random manner from the pool of students at only one Japanese university, research findings should not be generalized to other populations. Third, although a latent variable, English proficiency, was assessed using three sections of one kind of test, namely the TOEFL, it would be better to use three different tests, each of which measures a different aspect of English proficiency. Fourth, although the present study ended up with observations in Canada, it might be interesting to know what happened after UBC-Rits students returned to Japan. Keep tracking of subjects after their return to Japan would make it possible to interpret the L2 socialization process more diachronically.

Thus, future L2 socialization research should be designed with these limitations in mind, especially when it is quantitative in nature. Furthermore, it would be ideal to conduct research in which both quantitative and qualitative approaches are employed to investigate the same research question. Such an approach would produce findings that could be cross-validated and corroborated. It is hoped that the present study has contributed to demonstrating specifically how important it is to incorporate a quantitative approach into L2 socialization research and how findings from a quantitative approach can advance our understanding of the complexities of L2 socialization.

BIBLIOGRAPHY

- Alwin, D. F., & Jackson, D. J. (1981). Applications of simultaneous factor analysis to issues of factorial invariance. In D. D. Jackson & E. F. Borgatta (Eds.), *Factor analysis and measurement in sociological research: A multidimensional perspective* (pp. 249-280). Beverly Hills, CA: Sage.
- Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin*, 103, 411-423.
- Atkinson, D., & Ramanathan, V. (1995). Cultures of writing: An ethnographic comparison of L1 and L2 university writing/language programs. *TESOL Quarterly*, 29, 539-568.
- Bachman, L. F. (1991). *Fundamental considerations in language testing* (2nd ed.). Oxford University Press.
- Barnlund, D. C. (1975). *Public and private self in Japan and the United States: Communicative styles of two cultures*. Tokyo: Simul Press.
- Becker, H., Geer, B., Hughes, E. C., & Strauss, A. (1961). *Boys in white*. Chicago: University of Chicago Press.
- Bentler, P. M. (1986). Structural modeling and Psychometrika: An historical perspective on growth and achievements. *Psychometrika*, 51, 35-51.
- Bentler, P. M. (1992). On the fit of models to covariances and methodology to the *Bulletin*. *Psychological Bulletin*, 112, 400-404.
- Bentler, P. M., & Bonnett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, 88, 588-606.

- Bentler, P. M., & Chou, C. P. (1987). Practical issues in structural modeling. *Sociological Methods and Research*, 16, 78-117.
- Bialystok, E. (1993). Symbolic representation and attentional control in pragmatic competence. In G. Kasper & S. Blum-Kulka (Eds.), *Interlanguage pragmatics* (pp. 43-57). New York: Oxford University Press.
- Blum-Kulka, S., House, J. & Kasper, G. (1989). Investigating cross-cultural pragmatics: An introductory overview. In S. Blum-Kulka, J. House & G. Kasper (Eds.), *Cross-cultural pragmatics: Requests and apologies* (pp. 1-34). Norwood, NJ: Ablex.
- Blum-Kulka, S., & Olshtain, E. (1986). Too many words: Length of utterance and pragmatic failure. *Studies in Second Language Acquisition*, 8, 47-61.
- Bollen, K. A. (1989). *Structural equations with latent variables*. New York: John Wiley & Sons.
- Brown, P., & Levinson, S. C. (1987). *Politeness: Some universals in language usage*. New York: Cambridge University Press.
- Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 445-455). Newbury Park, CA: Sage.
- Byrne, B. M. (1988). Adolescent self-concept, ability grouping, and social comparison: Reexamining academic track differences in high school. *Youth and Society*, 20, 46-67.

- Byrne, B. M. (1998). *Structural equation modeling with LISREL, PRELIS, and SIMPLIS: Basic concepts, applications, and programming*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Campbell, R., & Wales, R. (1970). The study of language acquisition. In J. Lyons (Ed.), *New horizons in linguistics*. Harmondsworth, England: Penguin Books.
- Clancy, P. M. (1986). Acquiring communicative style in Japanese. In B. B. Schieffelin & E. Ochs (Eds.), *Language socialization across cultures*. New York: Cambridge University Press.
- Cohen, A. D. (1996). Developing the ability to perform speech acts. *Studies in Second Language Acquisition*, 18, 253-267.
- Cohen, J., & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Connell, J. P. (1987). Structural equation modeling and the study of child development: A question of goodness of fit. *Child Development*, 58, 167-175.
- Crago, M. B. (1992). Communicative interaction and second language acquisition: An Inuit example. *TESOL Quarterly*, 26, 487-505.
- Crago, M. B., Annahatak, B., & Ningiuruvik, L. (1993). Changing patterns of language socialization in Inuit homes. *Anthropology and Education Quarterly*, 24, 205-223.
- Crocker, L., & Algina, J. (1986). *Introduction to classical & modern test theory*. Orland, FL: Harcourt Brace Jovanovich.
- Doi, T. (1973). *Amae no kozo*. Tokyo: Kodansha.

- Doi, T. (1974). Some psychological themes in Japanese human relationships. In J. C. Condon & M. Saito (Eds.), *Intercultural encounters with Japan: Communication -contact and conflict* (pp. 17-26). Tokyo: Simul Press.
- Duran, R. P., Canale, M., Penfield, J., Stansfield, C. W., & Liskin-Gasparro, J. (1985). *TOEFL from a communicative viewpoint on language proficiency: A working paper*. TOEFL Research Report 17. Princeton, NJ: Educational Testing Service.
- Eisenstein, M. & Bodman, J. W. (1986). "I very appreciate": Expressions of gratitude by native and non-native speakers of American English. *Applied Linguistics*, 7, 167-185.
- Eisenstein, M. & Bodman, J. W. (1993). Expressing gratitude in American English. In G. Kasper & S. Blum-Kulka (Eds.), *Interlanguage pragmatics* (pp. 64-81). New York: Oxford University Press.
- Ellis, R. (1990). *Instructed second language acquisition*. Oxford: Blackwell.
- Ely, R., & Gleason, J. B. (1995). Socialization across contexts. In P. Fletcher & B. MacWhinney (Eds.), *The handbook of child language* (pp. 251-270). Oxford: Blackwell.
- Farrell, A. D. (1994). Structural equation modeling with longitudinal data: Strategies for examining group differences and reciprocal relationships. *Journal of Consulting and Clinical Psychology*, 62, 477-487.
- Fleiss, J. L., & Shrout, P.E. (1977). The effects of measurement errors on some multivariate procedures. *American Journal of Public Health*, 67, 1181-1191.
- Fraser, B. (1990). Perspectives on politeness. *Journal of Pragmatics*, 14, 219-236.

- Gee, J. P. (1992). *The social mind: Language, ideology and social practice*. New York: Bergin & Garvey.
- Gee, J. P. (1996). *Social linguistics and literacies: Ideology in discourses* (2nd ed.). London: Taylor & Francis.
- Glass, G. V., & Hopkins, K. D. (1996) *Statistical methods in education and psychology* (3rd ed.). Needham Heights, MA: Allyn & Bacon
- Gollob, H. F., & Reichardt, C. S. (1991). Interpreting and estimating indirect effects assuming time lags really matter. In L. M. Collins & J. L. Horn (Eds.), *Best methods for the analysis of change: Recent advances, unanswered questions, future directions* (pp. 243-259). Washington, DC: American Psychological Association.
- Gronlund, N. E. (1985). *Measurement and evaluation in teaching* (5th ed.). New York: Macmillan.
- Halliday, M. A. K., & Hasan, R. (1985). *Language, context, and text: Aspects of language in a social-semiotic perspective*. Oxford: Oxford University Press.
- Harklau, L. (1994). ESL versus mainstream classes: Contrasting L2 learning environments. *TESOL Quarterly*, 28, 241-272.
- Hayduk, L. A. (1987). *Structural equation modeling with LISREL*. Baltimore: Johns Hopkins University.
- Heath, S. B. (1982). Ethnography in education: Defining the essentials. In P. Gilmore & A. A. Glatthorn (Eds.), *Children in and out of school: Ethnography and education* (pp. 33-55). Washington, DC: Center for Applied Linguistics.

- Hinkel, E. (1997). Appropriateness of advice: DCT and multiple choice data. *Applied Linguistics*, 18, 1-26.
- Hoyle, R. H., & Smith, G. T. (1994). Formulating clinical research hypotheses as structural equation models: A conceptual overview. *Journal of Consulting and Clinical Psychology*, 62, 429-440.
- Huebner, T. (1979). Order of acquisition vs. dynamic paradigm: A comparison of method in interlanguage research. *TESOL Quarterly*, 13, 21-28.
- Hymes, D. H. (1971). *Piginization and creolization*. Cambridge: Cambridge University Press.
- Hymes, D. H. (1972a). Models of the interaction of language and social life. In J. J. Gumperz & D. H. Hymes (Eds.), *Directions in sociolinguistics: The ethnography of communication* (pp. 35-71). New York: Holt, Rinehart, & Winston.
- Hymes, D. H. (1972b). On communicative competence. In J. B. Pride & J. Holmes (Eds.), *Sociolinguistics*. Harmondsworth, England: Penguin Books.
- James, L. R., Mulaik, S. A., & Brett, J. M. (1982). *Causal analysis: Assumptions, models, and data*. Beverly Hills, CA: Sage.
- Jöreskog, K. G. (1971). Statistical analysis of sets of congeneric tests. *Psychometrika*, 36, 109-133.
- Jöreskog, K. G. (1993). Testing structural equation models. In J. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 294-316). Newbury Park, CA: Sage.
- Jöreskog, K. G., & Sörbom, D. (1996). *LISREL 8: User's reference guide*. Chicago, IL: Scientific Software International.

- Jöreskog, K. G., & Sörbom, D. (1999). *LISREL (Version 8.30)* [Computer software]. Chicago, IL: Scientific Software International.
- Judd, C. M., & Milburn, M. A. (1980). The structure of attitude systems in the general public: Comparisons of a structural equation model. *American Sociological Review*, 45, 627-643.
- Kasper, G. (1990). Linguistic politeness: Content research issues. *Journal of Pragmatics*, 14, 193-218.
- Kasper, G. (1997). Beyond reference. In G. Kasper & E. Kellerman (Eds.), *Communication strategies: Psycholinguistic and sociolinguistic perspectives* (pp. 345-360). Essex: Addison Wesley Longman.
- Kasper, G., & Dahl, M. (1991). Research methods in interlanguage pragmatics. *Studies in Second Language Acquisition*, 13, 215-247.
- Kasper, G., & Schmidt, R. (1996). Developmental issues in interlanguage pragmatics. *Studies in Second Language Acquisition*, 18, 149-169.
- Kessler, R. C., & Greenberg, D. F. (1981). *Linear panel analysis: Models of quantitative change*. New York: Academic Press.
- Kubota, R. (1999). Japanese culture constructed by discourses: Implications for applied linguistics research and ELT. *TESOL Quarterly*, 33, 9-35.
- Levinson, S. (1983). *Pragmatics*. Cambridge: Cambridge University Press.
- Loehlin, J. C. (1998). *Latent variable models: An introduction to factor, path, and structural analysis* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

- MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods, 1*, 130-149.
- Maeshiba, N., Yoshinaga, N., Kasper, G., & Ross, S. (1996). Transfer and proficiency in interlanguage apologizing. In S. Gass & J. Neu (Eds.), *Speech act across cultures* (pp. 155-187). New York: Newbury House.
- Matsumoto, Y. (1988). Reexamination of the universality of face: Politeness phenomena in Japanese. *Journal of Pragmatics, 12*, 403-426.
- Matsumoto, Y. (1989). Politeness and conversational universals: Observations from Japanese. *Multilingua, 82*, 207-221.
- Matsumura, S., & Takakuwa, M. (1999, August). *A critical framework for second-language socialization*. Paper presented at the 12th World Congress of Applied Linguistics. Waseda University, Tokyo.
- McMillan, J. H., & Schumacher, S. (1993). *Research in education: A conceptual introduction* (3rd ed.). New York: HarperCollins.
- Mulaik, S. A. (1972). *The foundations of factor analysis*. New York: McGraw-Hill.
- Nakane, C. (1967). *Tate-shakai no ningen-kankei: Tan'itsu-shakai no riron* [Personal relations in a vertical society: A theory of a homogeneous society]. Tokyo: Kodansha.
- Ninio, A., & Snow, C. E. (1996). *Pragmatic development*. Colorado: Westview Press.
- Ochs, E. (1988). *Culture and language development*. Cambridge: Cambridge University Press.

- Ochs, E., & Schieffelin, B. B. (Eds.). (1979). *Developmental pragmatics*. New York: Academic Press.
- Ochs, E., & Schieffelin, B. B. (1996). The impact of language socialization on grammatical development. In P. Fletcher & B. MacWhinney (Eds.), *The handbook of child language* (pp. 73-94). Oxford: Blackwell.
- Olshtain, E., & Blum-Kulka, S. (1985). Degree of approximation: Nonnative reactions to native speech act behavior. In S. M. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 303-325). Rowley, MA: Newbury House.
- Pedhazur, E. J. (1997). *Multiple regression in behavioral research: Explanation and prediction* (3rd ed.). Fort Worth, TX: Harcourt Brace.
- Pentz, M. A., & Chou, C. P. (1994). Measurement invariance in longitudinal clinical research assuming change from development and intervention. *Journal of Consulting and Clinical Psychology*, 62, 450-462.
- Pfaff, C. W. (1987). Functional approaches to interlanguage. In C. W. Pfaff (Ed.), *First and second language acquisition processes* (pp. 81-102). Cambridge, MA: Newbury House.
- Poole, D. (1992). Language socialization in the second language classroom. *Language Learning*, 42, 593-616.
- Raimes, A., & Zamel, V. (1997). Response to Ramanathan and Kaplan. *Journal of Second Language Writing*, 5, 21-34.
- Richards, J. C., Platt, J. & Platt, H. (1992). *Dictionary of language teaching & applied linguistics* (2nd ed.). Essex: Longman.

- Robinson, M. A. (1992). Introspective methodology in interlanguage pragmatics research. In G. Kasper (Ed.), *Pragmatics of Japanese as native and target language* (Second Language Teaching and Curriculum Center Technical Report, No. 3, pp. 27-82). Honolulu: University of Hawaii Press.
- Rose, K. (1994). On the validity of discourse completion tests in non-western contexts. *Applied Linguistics*, 15, 1-14.
- Schachter, J. (1986). In search of systematicity in interlanguage production. *Studies in Second Language Acquisition*, 8, 119-134.
- Schechter, S. R., & Bayley, R. (1997). Language socialization practices and cultural identity: Case studies of Mexican-descent families in California and Texas. *TESOL Quarterly*, 31, 513-541.
- Schieffelin, B. B., & Ochs, E. (Eds.). (1986a). *Language socialization across cultures*. New York: Cambridge University Press.
- Schieffelin, B. B., & Ochs, E. (1986b). Language socialization. *Annual Review of Anthropology*, 15, 163-191.
- Schumann, J. H. (1978). Social and psychological factors in second language acquisition. In J. C. Richards (Ed.), *Understanding second and foreign language learning: Issues and approaches* (pp. 163-178). Rowley, MA: Newbury House.
- Spack, R. (1997). The rhetorical construction of multilingual students. *TESOL Quarterly*, 31, 765-774.
- Stansfield, C. W. (Ed.). (1986). *Toward communicative competence testing: Proceedings of the second TOEFL invitational conference*. Princeton, NJ: Educational Testing Service.

- Steiger, J. H. (1989). *EzPATH: Causal modeling*. Evanston, IL: SYSTAT Inc.
- Steiger, J. H. (1990). Structural model evaluation and modification: An interval estimation approach. *Multivariate Behavioral Research*, 25, 173-180.
- Susser, B. (1998). EFL's othering of Japan: Orientalism in English language teaching. *JALT Journal*, 20, 49-82.
- Takahashi, S. (1992). Transferability of indirect request strategies. *University of Hawaii Working Papers in ESL*, 11(1), 69-124. (ERIC Document Reproduction Service No. ED 367 128).
- Takahashi, S. (1996). Pragmatic transferability. *Studies in Second Language Acquisition*, 18, 189-223.
- Takahashi, S., & DuFon, P. (1989). *Cross-linguistic influence in indirectness: The case of English directives performed by native Japanese speakers*. Unpublished manuscript, University of Hawai'i at Manoa, Honolulu. (ERIC Document Reproduction Service No. ED 370 439) .
- Takahashi, T., & Beebe, L. M. (1987). The development of pragmatic competence by Japanese learners of English. *JALT Journal*, 8, 131-155.
- Takahashi, T., & Beebe, L. M. (1993). Cross-linguistic influence in the speech act of correction. In G. Kasper & S. Blum-Kulka (Eds.), *Interlanguage pragmatics* (pp. 138-157). New York: Oxford University Press.
- Tanaka, J. S. (1993). Multifaceted conceptions of fit in structural equation models. In J. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 10-39). Newbury Park, CA: Sage.

- Triandis, H. C., Chen, X. P., & Chan, D. K. S. (1998). Scenarios for the measurement of collectivism and individualism. *Journal of Cross-Cultural Psychology*, 29, 275-289.
- Trosborg, A. (1987). Apology strategies in native/nonnatives. *Journal of Pragmatics*, 11, 147-167.
- Wentworth, W. M. (1980). *Context and understanding: An inquiry into socialization theory*. New York: Elsevier.
- Wierzbicka, A. (1991). *Cross-cultural pragmatics: The semantics of human interaction*. Berlin: Mouton de Gruyter.
- Willett, J. (1995). Becoming first graders in an L2: An ethnographic study of language socialization. *TESOL Quarterly*, 29, 473-504.
- Wolf, R. M. (1988). Questionnaires. In J. P. Keeves (Ed.), *Educational research, methodology, and measurement* (pp. 478-482). Oxford: Pergamon.
- Zamel, V. (1997). Toward a model of transculturation. *TESOL Quarterly*, 31, 341-352.

Appendix A**QUESTIONNAIRE ON JAPANESE STUDENTS' SPEECH ACT BEHAVIOR**

Through communication in English with native speakers of Japanese (e.g., your roommates or students), you might have recognized that some expressions, grammars and/or sentence structures that they frequently use sound awkward to you. Please list them below.

Thank you for your cooperation!

Appendix B

QUESTIONNAIRE ON YOUR BACKGROUND

1. Code: _____
2. Sex: Male / Female
3. Grade (Please circle one.): First-year / Second-year / Third-year / Forth-year
4. Department you are currently enrolled in: _____
5. Your parents' first language is Japanese. (Please circle one.): Yes / No
If not, please specify: _____
6. Do you have any experience staying or studying abroad. (Please circle one.): Yes / No
If yes, please write the name of the country, the length of stay, and the purpose.

Country	Length of Stay	Purpose
(example) <u>Australia</u>	<u>year(s)</u> <u>month(s)</u> <u>3 week(s)</u>	<u>sightseeing</u>
(1) _____	<u>year(s)</u> <u>month(s)</u> <u>week(s)</u>	_____
(2) _____	<u>year(s)</u> <u>month(s)</u> <u>week(s)</u>	_____
(3) _____	<u>year(s)</u> <u>month(s)</u> <u>week(s)</u>	_____
(4) _____	<u>year(s)</u> <u>month(s)</u> <u>week(s)</u>	_____
(5) _____	<u>year(s)</u> <u>month(s)</u> <u>week(s)</u>	_____

Thank you for your cooperation!

Appendix C**QUESTIONNAIRE ON CURRENT USES OF ENGLISH**

1. Code: _____
2. How often do you do the following activities?:
 - (a) Communicating in English with your friends.
_____ hour(s) _____ minute(s) per day.
 - (b) Communicating in English with your instructors.
_____ hour(s) _____ minute(s) per day.
 - (c) Communicating in English with _____ (Please specify.)
_____ hour(s) _____ minute(s) per day.
 - (d) Reading English newspaper such as "Japan Times."
_____ hour(s) _____ minute(s) per day.
 - (e) Reading English magazine such as "Newsweek."
_____ hour(s) _____ minute(s) per day.
 - (f) Reading English textbooks.
_____ hour(s) _____ minute(s) per day.
 - (g) Reading _____ in English. (Please specify.)
_____ hour(s) _____ minute(s) per day.
 - (h) Watching TV programs in English
_____ hour(s) _____ minute(s) per day.
 - (i) Watching movies in English.
_____ hour(s) _____ minute(s) per day.
 - (j) Watching _____ in English. (Please specify.)
_____ hour(s) _____ minute(s) per day.
 - (k) Listening to radio programs in English.
_____ hour(s) _____ minute(s) per day.
 - (l) Listening to English songs in CD.
_____ hour(s) _____ minute(s) per day.

- (m) Listening to _____ in English. (Please specify.)
_____ hour(s) _____ minute(s) per day.
- (n) Writing term papers in English.
_____ hour(s) _____ minute(s) per day.
- (o) Writing diary in English.
_____ hour(s) _____ minute(s) per day.
- (p) Writing _____ in English. (Please specify.)
_____ hour(s) _____ minute(s) per day.
- (q) Writing e-mails in English.
_____ hour(s) _____ minute(s) per day.
- (r) Others _____. (Please specify.)
_____ hour(s) _____ minute(s) per day.
3. What is your most recent TOEFL score?
Total: _____
Section I: _____ Section II: _____ Section III: _____

Thank you for your cooperation!

Appendix D

A MULTIPLE - CHOICE QUESTIONNAIRE (Japanese version)

Code: _____.

(Please do not put your name on the questionnaire!)

Instructions: 次ページから 1 2 個の場面が設定されています。それぞれの場面には 4 つの選択肢が与えられています。あなたがもっとも適切だと思う表現を 1 つ選んでマルをつけてください。何も言わないのが適切だと思えば選択肢 D.Nothing にマルをつけてください。ただし、場面はすべてカナダだと仮定し、登場人物もすべて英語のネイティブスピーカーだと仮定してください。それぞれの場面に登場する人物は次の 3 人の中の 1 人です。

Supervisor: P.D. is your supervisor. You have been taking P.D.'s seminar for three months. You and P.D., together with other students, have gone out for dinner several times after the seminar. You have visited P.D.'s office several times to talk about the topic you would present in the seminar.

(P.D.はあなたのゼミの先生です。P.D.のゼミを受け始めて 3 ヶ月になります。あなたは他の生徒と共に、先生と授業後何度か夕食にでかけたことがあります。あなたは自分がゼミで発表するテーマについて話をするために先生のオフィスを何度か訪れたことがあります。)

Classmate: C.J. is your classmate. You and C.J. often go out for lunch together after the class. You have borrowed C.J.'s notebook several times before. You regard C.J. as a good friend.

(C.J.はあなたのクラスメートです。あなたと C.J.は授業後よく一緒にランチを食べに出かけます。あなたは C.J.のノートを以前何度か借りたことがあります。あなたは C.J.のことを仲のいい友達だと思っています。)

First-year university student: X.L. is a first-year student. You and X.L. belong to the same club. You and X.L. often go out for dinner together after the club activity. You regard X.L. as a good friend.

(X.L.は大学 1 年生でああなたのクラブの後輩です。あなたと X.L.はクラブ活動後よく一緒に夕食を食べに出かけます。あなたは X.L.のことを仲のいい友達だと思っています。)

[Supervisor: P.D.; Classmate: C.J.; First-year university student: X.L.]

Situations

1. You and the instructor P.D. are in a restaurant. The instructor says something about ordering a hamburger. You ordered a hamburger in this restaurant before and, in your opinion, it was really greasy. What do you think would be appropriate to say in this situation?

(あなたとゼミの先生 P.D. はレストランにいます。先生はハンバーガーを注文するといっています。ところが、以前あなたがこのレストランでハンバーガーを注文した際、ひどく油っぽかったのを覚えています。この場面で先生に何か言うとしたら、あなたなら次のどの表現を使いますか。)

- A. You shouldn't order the hamburger. I had it here before, and it was really greasy.
- B. Maybe it's not a good idea to order a hamburger. I had one here before, and it was really greasy.
- C. I had it here before, and it was really greasy.
- D. Nothing

2. Your classmate C.J. considers skipping today's afternoon class. You happened to know that one absence loses five points from one's final marks in the class. What do you think would be appropriate to say in this situation?

(あなたのクラスメートの C.J. が今日の午後のクラスをサボろうと考えています。そのクラスの授業を 1 回サボると最終の成績から 5 点引かれることをあなたは何かの機会に知りました。この場面でクラスメートの C.J. に何か言うとしたら、あなたなら次のどの表現を使いますか。)

- A. I've heard one absence loses five points from the final marks.
- B. You should come to class. I've heard one absence loses five points from your final marks.
- C. I think it's better to come to class. I've heard one absence loses five points from your final marks.
- D. Nothing

3. X.L. is considering taking a course. You have heard that the course is really difficult. What do you think would be appropriate to say in this situation?

(クラブの後輩の X.L. がある授業を取ろうと考えています。ところが、そのクラスが非常に難しいクラスであることをあなたは耳にしました。この場面で後輩の X.L. に何か言うとしたら、あなたなら次のどの表現を使いますか。)

- A. I don't think it's a good idea to take this course. I've heard it's really difficult.
- B. I've heard it's really difficult.
- C. You shouldn't take this course. I've heard it's really difficult.
- D. Nothing

[Supervisor: P.D.; Classmate: C.J.; First-year university student: X.L.]

4. You see the supervisor P.D. working in the office late at night and looking pale. What do you think would be appropriate to say in this situation?

(あなたのゼミの先生 P.D.がオフィスで夜遅くに仕事をしているのをみかけました。先生の顔色はよくありません。この場面で先生に何か言うとしたら、あなたなら次のどの表現を使いますか。)

- A. I'm going home soon. It's very late.
- B. You shouldn't work so hard. It's very late.
- C. Maybe it's better to go home. It's very late.
- D. Nothing

5. You see your classmate C.J. put a one-dollar coin into the slot of a broken vending machine. C.J. couldn't get a pop or the money back from the machine. What do you think would be appropriate to say in this situation?

(あなたのクラスメートの C.J.が自動販売機に 1 ドルコインを入れたところ、コーラが出てこず、お金も戻ってきません。この場面でクラスメートの C.J.に何か言うとしたら、あなたなら次のどの表現を使いますか。)

- A. Maybe it's better to complain about it. The office is downstairs.
- B. You should complain about it. The office is downstairs.
- C. The office is downstairs.
- D. Nothing

6. X.L. is thinking of taking a car to a repair shop downtown. However, you know it's notorious for a sloppy job. What do you think would be appropriate to say in this situation?

(クラブの後輩の X.L.がダウンタウンにある車の修理工場に自分の車を持っていこうと考えています。ところが、その修理工場はまともな仕事ができないことで有名なことをあなたは知っています。この場面で後輩の X.L.に何か言うとしたら、あなたなら次のどの表現を使いますか。)

- A. You shouldn't take your car to that shop. It has a really bad reputation.
- B. Maybe it's better to take your car to another shop. It has a really bad reputation.
- C. I usually don't take my car to that shop. It has a really bad reputation.
- D. Nothing

[Supervisor: P.D.; Classmate: C.J.; First-year university student: X.L.]

7. You see the supervisor P.D. is considering buying an expensive book without knowing that another bookstore sells it at a 20 percent discount. What do you think would be appropriate to say in this situation?

(あなたのゼミの先生 P.D.が高い本を買おうとしています。どうやら、先生は別の本屋で同じ本を20パーセントオフで売っていることを知らない様子です。この場面で先生に何か言うとしたら、あなたなら次のどの表現を使いますか。)

- A. You should buy the book at another store. This store is over-priced.
- B. This store is over-priced.
- C. Maybe, it's not a good idea to buy the book here. This store is over-priced.
- D. Nothing

8. You see your classmate C.J. working on the assignment late at night and is visibly tired. What do you think would be appropriate to say in this situation?

(あなたのクラスメートの C.J.が夜遅くに宿題に取り組んでいるのを見かけました。C.J.は目に見えて疲れています。この場面でクラスメートの C.J.に何か言うとしたら、あなたなら次のどの表現を使いますか。)

- A. Maybe it's better to go home. It's very late.
- B. I'm going home soon. It's very late.
- C. You shouldn't work so hard. It's very late.
- D. Nothing

9. You have heard from X.L. that X.L. didn't get the exact amount of change at the cashier of the cafeteria. What do you think would be appropriate to say in this situation?

(クラブの後輩の X.L.がカフェテリアのレジで正確なおつりをもらえなかったことをあなたは X.L.から聞きました。この場面で後輩の X.L.に何か言うとしたら、あなたなら次のどの表現を使いますか。)

- A. Maybe it's better to complain about it. That person is still there.
- B. You should complain about it. That person is still there.
- C. That person is still there.
- D. Nothing

[Supervisor: P.D.; Classmate: C.J.; First-year university student: X.L.]

10. You have just heard from the supervisor P.D. that the supervisor is considering a trip to Banff from Vancouver in a car which breaks down frequently. What do you think would be appropriate to say in this situation?

(あなたのゼミの先生 P.D.がバンクーバーからバンフまでの旅行を計画していることを知りました。先生は頻繁に故障する自分の車で行くと言っています。この場面で先生に何か言うとしたら、あなたなら次のどの表現を使いますか。)

- A. I think it may be risky for you to take such a long trip in this car.
- B. Taking such a long trip in this car may be risky.
- C. You shouldn't take this car for such a long trip. It may be risky.
- D. Nothing

11. You have noticed that your classmate C.J. has forgotten to leave a tip when leaving. What do you think would be appropriate to say in this situation?

(あなたはクラスメートの C.J.がチップを置くのをすっかり忘れてレストランを出て行こうとしていることに気づきました。この場面でクラスメートの C.J.に何か言うとしたら、あなたなら次のどの表現を使いますか。)

- A. A tip is important. You shouldn't forget to leave one.
- B. A tip is important.
- C. Maybe it's better to leave a tip. It's important.
- D. Nothing

12. You see X.L. studying in the library late at night and looking pale. What do you think would be appropriate to say in this situation?

(クラブの後輩の X.L.が夜遅くに勉強しているのを見かけました。X.L.の顔色はよくありません。この場面で後輩の X.L.に何か言うとしたら、あなたなら次のどの表現を使いますか。)

- A. You shouldn't work so hard. It's very late.
- B. Maybe it's better to go home. It's very late.
- C. I'm going home soon. It's very late.
- D. Nothing

Thank you for your cooperation!

Appendix E

A MULTIPLE - CHOICE QUESTIONNAIRE (English version)

Code: _____

(Please do not put your name on the questionnaire!)

Instructions: Several situations are described in the items below. Following the description of a situation, you will find a multiple choice selection of three possible statements, A, B, and C. Choose the statement that you think would be most appropriate to say in the situation. If you think it would be most appropriate to say nothing, choose option D. Please circle one of the four options in each situation.

When you are responding to the questions, please keep in mind that all situations happen in Canada and all three imaginary characters are native-speakers of English. The three imaginary people are:

Supervisor: P.D. is your supervisor. You have been taking P.D.'s seminar for three months. You and P.D., together with other students, have gone out for dinner several times after the seminar. You have visited P.D.'s office several times to talk about the topic you would present in the seminar.

Classmate: C.J. is your classmate. You and C.J. often go out for lunch together after the class. You have borrowed C.J.'s notebook several times before. You regard C.J. as a good friend.

First-year university student: X.L. is a first-year student. You and X.L. belong to the same club. You and X.L. often go out for dinner together after the club activity. You regard X.L. as a good friend.

[Supervisor: P.D.; Classmate: C.J.; First-year university student: X.L.]

Situations

1. You and the instructor P.D. are in a restaurant. The instructor says something about ordering a hamburger. You ordered a hamburger in this restaurant before and, in your opinion, it was really greasy. What do you think would be appropriate to say in this situation?
 - A. You shouldn't order the hamburger. I had it here before, and it was really greasy.
 - B. Maybe it's not a good idea to order a hamburger. I had one here before, and it was really greasy.
 - C. I had it here before, and it was really greasy.
 - D. Nothing
2. Your classmate C.J. considers skipping today's afternoon class. You happened to know that one absence loses five points from one's final marks in the class. What do you think would be appropriate to say in this situation?
 - A. I've heard one absence loses five points from the final marks.
 - B. You should come to class. I've heard one absence loses five points from your final marks.
 - C. I think it's better to come to class. I've heard one absence loses five points from your final marks.
 - D. Nothing
3. X.L. is considering taking a course. You have heard that the course is really difficult. What do you think would be appropriate to say in this situation?
 - A. I don't think it's a good idea to take this course. I've heard it's really difficult.
 - B. I've heard it's really difficult.
 - C. You shouldn't take this course. I've heard it's really difficult.
 - D. Nothing
4. You see the supervisor P.D. working in the office late at night and looking pale. What do you think would be appropriate to say in this situation?
 - A. I'm going home soon. It's very late.
 - B. You shouldn't work so hard. It's very late.
 - C. Maybe it's better to go home. It's very late.
 - D. Nothing

[Supervisor: P.D.; Classmate: C.J.; First-year university student: X.L.]

5. You see your classmate C.J. put a one-dollar coin into the slot of a broken vending machine. C.J. couldn't get a pop or the money back from the machine. What do you think would be appropriate to say in this situation?

- A. Maybe it's better to complain about it. The office is downstairs.
- B. You should complain about it. The office is downstairs.
- C. The office is downstairs.
- D. Nothing

6. X.L. is thinking of taking a car to a repair shop downtown. However, you know it's notorious for a sloppy job. What do you think would be appropriate to say in this situation?

- A. You shouldn't take your car to that shop. It has a really bad reputation.
- B. Maybe it's better to take your car to another shop. It has a really bad reputation.
- C. I usually don't take my car to that shop. It has a really bad reputation.
- D. Nothing

7. You see the supervisor P.D. is considering buying an expensive book without knowing that another bookstore sells it at a 20 percent discount. What do you think would be appropriate to say in this situation?

- A. You should buy the book at another store. This store is over-priced.
- B. This store is over-priced.
- C. Maybe, it's not a good idea to buy the book here. This store is over-priced.
- D. Nothing

8. You see your classmate C.J. working on the assignment late at night and is visibly tired. What do you think would be appropriate to say in this situation?

- A. Maybe it's better to go home. It's very late.
- B. I'm going home soon. It's very late.
- C. You shouldn't work so hard. It's very late.
- D. Nothing

9. You have heard from X.L. that X.L. didn't get the exact amount of change at the cashier of the cafeteria. What do you think would be appropriate to say in this situation?

- A. Maybe it's better to complain about it. That person is still there.
- B. You should complain about it. That person is still there.
- C. That person is still there.
- D. Nothing

[Supervisor: P.D.; Classmate: C.J.; First-year university student: X.L.]

10. You have just heard from the supervisor P.D. that the supervisor is considering a trip to Banff from Vancouver in a car which breaks down frequently. What do you think would be appropriate to say in this situation?

- A. I think it may be risky for you to take such a long trip in this car.
- B. Taking such a long trip in this car may be risky.
- C. You shouldn't take this car for such a long trip. It may be risky.
- D. Nothing

11. You have noticed that your classmate C.J. has forgotten to leave a tip when leaving. What do you think would be appropriate to say in this situation?

- A. A tip is important. You shouldn't forget to leave one.
- B. A tip is important.
- C. Maybe it's better to leave a tip. It's important.
- D. Nothing

12. You see X.L. studying in the library late at night and looking pale. What do you think would be appropriate to say in this situation?

- A. You shouldn't work so hard. It's very late.
- B. Maybe it's better to go home. It's very late.
- C. I'm going home soon. It's very late.
- D. Nothing

Thank you for your cooperation!

Appendix F

LISREL Input File 1

```

! Model 2
DA NI=32 NO=97 MA=CM
LA
PD1 CJ1 XL1 PD2 CJ2 XL2 PD3 CJ3 XL3 PD4 CJ4 XL4
L1 G1 R1 L2 G2 R2 L3 G3 R3 L4 G4 R4
PRO1 REC1 PRO2 REC2 PRO3 REC3 PRO4 REC4
RA=C:/RESEARCH/DISSERTATION.DAT FO
(24F2.0, 8F3.0)
SE
4 5 6 16 17 18 27 28 7 8 9 19 20 21 29 30 10 11 12 22 23 24 31 32 1 2 3 13 14 15 25 26/
MO NY=24 NX=8 NE=9 NK=3 LX=FU,FI LY=FU,FI GA=FU,FI BE=FU,FI PH=SY,FI
PS=SY,FI TD=SY,FI TE=SY,FI
LE
POSS2 PROF2 EXPO2 POSS3 PROF3 EXPO3 POSS4 PROF4 EXPO4
LK
POSS1 PROF1 EXPO1
VA 1 LY 1 1 LY 4 2 LY 7 3
VA 1 LY 9 4 LY 12 5 LY 15 6
VA 1 LY 17 7 LY 20 8 LY 23 9
FR LY 2 1 LY 3 1 LY 5 2 LY 6 2 LY 8 3
FR LY 10 4 LY 11 4 LY 13 5 LY 14 5 LY 16 6
FR LY 18 7 LY 19 7 LY 21 8 LY 22 8 LY 24 9
FR PS 1 1 PS 2 2 PS 3 3
FR PS 4 4 PS 5 5 PS 6 6
FR PS 7 7 PS 8 8 PS 9 9
FR PS 2 1 PS 3 2 PS 3 1 PS 5 4 PS 6 5 PS 6 4 PS 8 7 PS 9 8 PS 9 7
FR TE 1 1 TE 2 2 TE 3 3 TE 4 4 TE 5 5 TE 6 6 TE 7 7 TE 8 8
FR TE 9 9 TE 10 10 TE 11 11 TE 12 12 TE 13 13 TE 14 14 TE 15 15 TE 16 16
FR TE 17 17 TE 18 18 TE 19 19 TE 20 20 TE 21 21 TE 22 22 TE 23 23 TE 24 24
FR TE 9 1 TE 17 9 TE 10 2 TE 18 10 TE 11 3 TE 19 11
FR TE 12 4 TE 13 5 TE 14 6 TE 20 12 TE 21 13 TE 22 14
FR TE 15 7 TE 23 15 TE 16 8 TE 24 16
VA 1 LX 1 1 LX 4 2 LX 7 3
FR LX 2 1 LX 3 1 LX 5 2 LX 6 2 LX 8 3
FR PH 1 1 PH 2 2 PH 3 3
FR PH 2 1 PH 3 2 PH 3 1
FR TD 1 1 TD 2 2 TD 3 3 TD 4 4 TD 5 5 TD 6 6 TD 7 7 TD 8 8
FR TH 1 1 TH 2 2 TH 3 3 TH 4 4 TH 5 5 TH 6 6 TH 7 7 TH 8 8

```

(To be continued)

(Continued)

EQ LX 2 1 LY 2 1 LY 10 4 LY 18 7
EQ LX 3 1 LY 3 1 LY 11 4 LY 19 7
EQ LX 5 2 LY 5 2 LY 13 5 LY 21 8
EQ LX 6 2 LY 6 2 LY 14 5 LY 22 8
EQ LX 8 3 LY 8 3 LY 16 6 LY 24 9
FR GA 1 1 GA 2 2 GA 3 3 GA 1 2 GA 1 3 GA 2 3
FR BE 4 1 BE 5 2 BE 6 3 BE 4 2 BE 4 3 BE 5 3
FR BE 7 4 BE 8 5 BE 9 6 BE 7 5 BE 7 6 BE 8 6
PD
OU AD=OFF RS EF MI SC

Appendix G

LISREL Input File 2

```

! Testing for Invariance of Latent Means across Groups
! Group 1--Rits in Canada
DA NG=2 NI=12 NO=97 MA=CM
LA
PD1 CJ1 XL1 PD2 CJ2 XL2 PD3 CJ3 XL3 PD4 CJ4 XL4
KM
1.00
0.44 1.00
0.52 0.72 1.00
0.57 0.61 0.60 1.00
0.62 0.62 0.56 0.54 1.00
0.59 0.55 0.54 0.53 0.57 1.00
0.53 0.57 0.63 0.83 0.43 0.49 1.00
0.64 0.61 0.54 0.49 0.73 0.53 0.47 1.00
0.65 0.53 0.50 0.51 0.53 0.80 0.48 0.67 1.00
0.55 0.58 0.59 0.79 0.53 0.57 0.89 0.55 0.56 1.00
0.61 0.53 0.41 0.45 0.62 0.49 0.43 0.90 0.65 0.50 1.00
0.67 0.58 0.56 0.52 0.53 0.75 0.48 0.68 0.93 0.55 0.67 1.00
ME
12.13 10.78 11.04 12.82 11.43 11.74 12.99 12.98 12.75 13.12 13.26 13.00
SD
1.91 2.08 1.82 1.64 1.78 1.55 1.82 1.71 1.55 1.54 1.57 1.49
SE
1 2 3 4 5 6 7 8 9 10 11 12/
MO NX=12 NK=4 LX=FU,FI PH=SY,FR TD=SY TX=FR KA=FR
LK
POSS1 POSS2 POSS3 POSS4
FR LX 2 1 LX 3 1 LX 5 2 LX 6 2 LX 8 3 LX 9 3 LX 11 4 LX 12 4
FR TD 4 1 TD 5 2 TD 6 3 TD 7 1 TD 8 2 TD 9 3 TD 10 1 TD 11 2 TD 12 3
FR TD 7 4 TD 8 5 TD 9 6 TD 10 4 TD 11 5 TD 12 6
FR TD 10 7 TD 11 8 TD 12 9
ST .1 LX 2 1 LX 3 1 LX 5 2 LX 6 2 LX 8 3 LX 9 3 LX 11 4 LX 12 4
ST .02 PH 1 1 PH 2 2 PH 3 3 PH 4 4
ST .01 PH 2 1 PH 3 1 PH 4 1 PH 3 2 PH 4 2 PH 4 3
ST .8 TD 1 1 TD 2 2 TD 3 3 TD 4 4 TD 5 5 TD 6 6 TD 7 7 TD 8 8 TD 9 9 TD 10 10 TD
11 11 TD 12 12
ST .4 TD 4 1 TD 5 2 TD 6 3 TD 7 1 TD 8 2 TD 9 3 TD 10 1 TD 11 2 TD 12 3
ST .4 TD 7 4 TD 8 5 TD 9 6 TD 10 4 TD 11 5 TD 12 6
ST .4 TD 10 7 TD 11 8 TD 12 9

```

(To be continued)

(Continued)

VA 1 LX 1 1 LX 4 2 LX 7 3 LX 10 4

PD

OU NS AD=OFF

Group 2--Rits in Japan

DA NO=102 MA=CM

LA

PD1 CJ1 XL1 PD2 CJ2 XL2 PD3 CJ3 XL3 PD4 CJ4 XL4

KM

1.00

0.87 1.00

0.86 0.84 1.00

0.90 0.85 0.87 1.00

0.73 0.73 0.76 0.75 1.00

0.86 0.81 0.83 0.80 0.67 1.00

0.81 0.85 0.88 0.78 0.77 0.82 1.00

0.83 0.78 0.79 0.76 0.59 0.82 0.85 1.00

0.88 0.86 0.84 0.86 0.76 0.84 0.88 0.86 1.00

0.78 0.81 0.83 0.72 0.70 0.81 0.87 0.83 0.86 1.00

0.82 0.78 0.84 0.81 0.54 0.80 0.83 0.76 0.82 0.83 1.00

0.83 0.84 0.80 0.82 0.73 0.70 0.88 0.81 0.83 0.83 0.83 1.00

ME

12.66 11.53 11.48 12.66 11.52 11.44 12.63 11.56 11.57 12.66 11.56 11.61

SD

1.63 1.57 1.55 1.82 1.73 1.67 1.85 1.85 1.91 2.15 2.16 2.12

SE

1 2 3 4 5 6 7 8 9 10 11 12/

MO LX=IN PH=SY,FR TD=SY TX=IN KA=FI

LK

POLITE1 POLITE2 POLITE3 POLITE4

FR TD 4 1 TD 5 2 TD 6 3 TD 7 1 TD 8 2 TD 9 3 TD 10 1 TD 11 2 TD 12 3

FR TD 7 4 TD 8 5 TD 9 6 TD 10 4 TD 11 5 TD 12 6

FR TD 10 7 TD 11 8 TD 12 9

ST .02 PH 1 1 PH 2 2 PH 3 3 PH 4 4

ST .01 PH 2 1 PH 3 1 PH 4 1 PH 3 2 PH 4 2 PH 4 3

ST .8 TD 1 1 TD 2 2 TD 3 3 TD 4 4 TD 5 5 TD 6 6 TD 7 7 TD 8 8 TD 9 9 TD 10 10 TD 11 11 TD 12 12

ST .4 TD 4 1 TD 5 2 TD 6 3 TD 7 1 TD 8 2 TD 9 3 TD 10 1 TD 11 2 TD 12 3

ST .4 TD 7 4 TD 8 5 TD 9 6 TD 10 4 TD 11 5 TD 12 6

ST .4 TD 10 7 TD 11 8 TD 12 9

VA 1 LX 1 1 LX 4 2 LX 7 3 LX 10 4

(To be continued)

(Continued)

FR LX 5 2 LX 6 2 LX 8 3 LX 9 3 LX 11 4 LX 12 4

PD

OU RS EF MR SS SC AD=OFF