A PILOT STUDY OF THE EFFECTS OF BACKGROUND CHARACTERISTICS OF INTERVIEWERS ON THE INTER-RATER RELIABILITY OF THE ORAL TESTING PROCEDURE FOR THE SENIOR HIGH SCHOOL FRENCH PROGRAM IN THE PROVINCE OF NEWFOUNDLAND

CENTRE FOR NEWFOUNDLAND STUDIES

TOTAL OF 10 PAGES ONLY
MAY BE XEROXED

(Without Author's Permission)

KEVIN FRANCIS FLYNN

National Library of Canada
Bibliothèque nationale du Canada

Acquisitions and Bibliographic Services Branch
Direction des acquisitions et des services bibliographiques

395 Wellington Street
Ottawa, Ontario
K1A 0N4

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

Canada

A Pilot Study Of The Effects Of Background Characteristics Of Interviewers On The Inter-Rater Reliability Of The Oral Testing Procedure For The Senior High School French Program In The Province Of Newfoundland.

by

Kevin Francis Flynn, B.A., B.Ed.

A thesis submitted to the school of Graduate Studies in partial fulfillment of the requirements for the degree of Master of Education.

Department of Curriculum
Memorial University of Newfoundland
March 1991

St. John's *Newfoundland*

National Library of Canada
Bibliothèque nationale du Canada

Acquisitions and Bibliographic Services Branch
Direction des acquisitions et des services bibliographiques

395 Wellington Street
Ottawa, Ontario
K1A 0N4

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

The author has granted an irrevocable non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of his/her thesis by any means and in any form or format, making this thesis available to interested persons.

L'auteur a accordé une licence irrévocable et non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de sa thèse de quelque manière et sous quelque forme que ce soit pour mettre des exemplaires de cette thèse à la disposition des personnes intéressées.

The author retains ownership of the copyright in his/her thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without his/her permission.

L'auteur conserve la propriété du droit d'auteur qui protège sa thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

ISBN 0-315-82605-3

Canada

EDUCATIONAL TESTING SERVICE PRINCETON, N.J. 08541

June 19, 1991

Mr. Kevin Flynn
26 Brookfield Road
St. John's, NF
Canada A1A 3T7

Dear Mr. Flynn:

Thank you for your letter in which you requested permission to use portions of our publication, The Oral Proficiency of Teachers in the United States in the 1980s - An Empirical Study. Self-Assessment of Oral Proficiency. Final Report by Thomas L. Hilton (1985).

Educational Testing Service is pleased to grant royalty - free, nonexclusive, nontransferable permission to reproduce the item listed on the attached appendix. The following terms apply to this permission:

1. The material is to be used only for the research purposes described in your letter and is not to be distributed, published, or used in any other manner without written permission from ETS;

2. Please use the following credit line following the source citation on the page where you use the material and/or other appropriate location:

   Reprinted by permission of Educational Testing Service, the copyright owner.

4. The test materials must be placed in an appendix to the dissertation, and the appendix must not be made available to University Microfilms, Inc.

5. You will assume responsibility for the analyses and conclusions of your study and, other than acknowledgment of the source of the questions, you will not use ETS' name in such a way as to imply participation in or responsibility for your research.

6. Please send a copy of your research report to my attention, when it becomes available.

7. This Agreement shall be considered null and void if not signed and returned within 30 days of the date of this letter.

If the above arrangements are satisfactory, please sign both copies of this letter, and return to me at the above listed address. A fully executed copy will be returned to you for your records.

Sincerely,

Anne F. Marcantonio
Copyrights and Permissions
Administrator

cc: Thomas Hilton, ETS

ACCEPTED AND AGREED TO:

BY: _____

TITLE: _____

DATE: July 19/91

ACCEPTED AND AGREED TO:

BY: _____
Educational Testing Service

TITLE: Copyrights & Permissions Administrator

DATE: _____

APPENDIX

MATERIALS COVERED BY THIS PERMISSION

The Oral Language Proficiency of Teachers in the United States in the 1980s - An Empirical Study. Self-Assessment of Oral Proficiency. Final Report (1985) Thomas L. Hilton Ed269426

ABSTRACT

Although numerous studies have shown a high degree of validity and reliability in oral proficiency testing, not all areas have been investigated. This paper is, essentially, the report of a pilot study investigating one such area: the relationship between background characteristics of teacher interviewers and the reliability of the rating of an oral proficiency test. It was hypothesized that, despite the formal training sessions offered to the interviewers, differences in the ratings of the oral proficiency of interviewers would occur, and that these differences could be related to the diverse characteristics of the teacher interviewer. The questions asked were:

(1) Are there significant differences in the way interviewers
   (a) rank the five factors of speaking proficiency?
   (b) rate the acceptability of errors?
   (c) rate proficiency levels?

(2) If significant variations do occur, could these differences be associated with the language proficiency or the background characteristics of the interviewers?

A questionnaire was distributed to one-half of the trained interview population on the island portion of the province. A response rate of 84 percent was achieved. The background characteristics investigated were:

(1) demographic information,

(2) number of years of teaching experience,

(3) number of university French courses taken,

(4) time spent in a French milieu,

(5) extent of speaking French in the milieu with friends and acquaintances and at home,

(6) a self-rating of aural and oral proficiency levels.

The dependent variables employed were:

(1) a ranking of the importance of five factors (vocabulary, grammar, pronunciation, fluency and comprehension) of oral proficiency;

(2) a rating of the acceptability of examples of various types of errors, and

(3) a rating of described proficiency levels as employed in the oral testing manual for the French 3200 Senior High School program.

The data were tabulated and a profile of the average respondent constructed; different variables were analyzed to determine significant differences, then crosstabulated with the independent variable and a chi test performed.

The results indicated that:

(1) there were differences in the rankings of the five factors of oral proficiency,

(2) there were significant differences in the rating of the acceptability of errors and the levels of oral proficiency,

(3) these differences were primarily in the areas of the rating of vocabulary and grammar examples,

(4) these differences were associated primarily with the language proficiency of the interviewers.

Further study of this area is recommended in order to determine to what extent differences would affect student evaluation when using global rather than discrete item evaluation techniques.

DEDICATION

This study is dedicated to the author's parents,

Mildred M. Flynn and Kevin S. Flynn,

for their constant support and encouragement.

ACKNOWLEDGEMENTS

The writer wishes to express his gratitude to Mrs. Joan Netten for her valuable guidance and encouragement.

Appreciation is also expressed to the following people:

(1) Paula Hogan who undertook the typing of this thesis.

(2) The teachers who completed the questionnaire for their cooperation made this study possible.

Finally, the writer wishes to thank his wife, Laura, for her constant support and encouragement during the completing of this thesis.

TABLE OF CONTENTS

ABSTRACT ii

ACKNOWLEDGEMENTS v

TABLE OF CONTENTS vi

LIST OF TABLES ix

LIST OF FIGURES xii

CHAPTER ONE 1

STATEMENT OF PROBLEM 1

Competence and performance 2

Role of language tester 4

Language tester profile 6

Questions for investigation 6

Procedures 8

Hypotheses 8

Significance of the study 9

Limitations for the study 9

Definitions of terms 10

CHAPTER TWO 12

REVIEW OF THE LITERATURE 12

Historical development and identification 12

Pre-Scientific Era 13

Psychometric - Structuralist Era 14

Teaching 14
Testing 16
The good test 17
Discrete point testing 17

Integrative-Sociolinguist era 20

Proficiency movement 21

Proficiency testing 23

Indirect proficiency tests 24
Semi-direct proficiency tests 26
Direct proficiency tests 27

Research questions 29

Components of the speaking skill 29
Errors 29
Reliability of student performance 30
Interviewer reliability 31
Summary 32

CHAPTER THREE 34

THE STUDY 34

Introduction 34

Instrument 35
Sample population 40
Selection 42
Data analysis 44

CHAPTER FOUR 47

PROFILE OF THE RESPONDENTS 47

Teaching experience 48

Number of years 48
Gender 49

Academic background 49

University courses 49
Time spent in a French milieu 50
Comparison between time spent in a French milieu and number of years of teaching experience 54

Oral proficiency 56

Speaking extent in a French milieu 56
Speaking extent at home and with acquaintances 61

Oral Proficiency (self-rated) 63
'Can do' statements 69

Aural proficiency (self-rated) 72

Profile of a typical respondent 76

Other general information 78

CHAPTER FIVE 81

ANALYSIS OF VARIANCE 81

Ranking of speaking proficiency factors and independent variables 84

Acceptability of errors and independent variables 89

Ratings of levels of oral proficiency and independent variables 92

Location 92
Gender 93
Time spent in milieu 94
Years teaching experience 96
Speaking extent in the milieu with friends or acquaintances and in the home 98
Oral proficiency (self-rated) 99
Aural proficiency (self-rated) 103

Conclusion 105

CHAPTER SIX 107

SUMMARY AND DISCUSSION 107

Problem 107

Methodology 107

Results 108

Conclusion 115

Global ratings versus discrete example ratings .116

Recommendations 118

Conclusion 119

BIBLIOGRAPHY 122

APPENDIX A 133

APPENDIX B 146

APPENDIX C 147

LIST OF TABLES

Table 1 Distribution of sample population 43

Table 2 Questionnaires: mailed out/returned 43

Table 3 Years teaching experience compared with gender . 48

Table 4 Academic qualifications as determined by university courses 50

Table 5 Time spent in a French milieu 51

Table 6 Time spent in a French milieu compared with university French courses 54

Table 7 Time spent in a French milieu compared with years teaching experience 55

Table 8 Speaking extent in a French milieu 56

Table 9 Speaking extent in a French milieu compared with university French courses 57

Table 10 Speaking extent in a French milieu compared with teaching experience 58

Table 11 Speaking extent at home and with acquaintances . 62

Table 12 Speaking extent in a French milieu compared with speaking extent with friends and acquaintances 62

Table 13 Speaking extent in a French milieu compared with speaking extent in the home 63

Table 14 Level of oral proficiency (self-rated) 64

Table 15 Level of oral proficiency compared with time spent in a French milieu 65

Table 16 Level of speaking ability compared with speaking extent in the milieu 66

Table 17 Level of speaking ability compared with university French courses 68

Table 18 Level of speaking ability compared with gender 69

Table 19 'Can Do' statements 71

Table 20 Significant levels for 'can do' statements compared with speaking extent and levels of oral proficiency 72

Table 21 Aural proficiency (self-rated) 73

Table 22 Significance levels for aural proficiency compared with speaking extent and levels of oral proficiency 76

Table 23 Awareness of goals and methodological change ... 78

Table 24 Respondents familarity with proficiency scale .. 79

Table 25 Aids employed in determining proficiency level . 79

Table 26 Actual numbers of significant values at or below .05 82

Table 27 Independent variables showing significant relationships with items measuring the importance of speaking proficiency factors 82

Table 28 Independent variables showing significant relationships with items measuring the acceptability of errors 83

Table 29 Independent variables showing significant relationships with items measuring level of oral proficiency 83

Table 30 Ranking of speaking proficiency factors compared with oral and aural proficiency 85

Table 31 Ranking of speaking proficiency factors compared with years of teaching experience 87

Table 32 Mean rankings for acceptability of errors compared with oral proficiency (self-rated) 90

Table 33 Mean rankings for acceptability of errors compared with aural proficiency (self-rated) ... 91

Table 34 Mean rankings for proficiency levels (grammar) compared with location 93

Table 35 Mean rankings for proficiency levels (vocabulary) compared with gender 94

Table 36 Mean rankings for proficiency levels (vocabulary) compared with time spent in a French milieu 95

Table 37 Mean rankings for proficiency levels (vocabulary) compared with years of teaching experience 97

Table 38 Mean rankings for proficiency levels (grammar) compared with speaking extent in the milieu 99

Table 39 Mean rankings for proficiency levels (vocabulary, grammar, fluency) compared with speaking extent with friends and acquaintances .100

Table 40 Mean rankings for proficiency levels (vocabulary, pronouns) compared with speaking extent in the home 101

Table 41 Mean rankings for proficiency levels (vocabulary, grammar, fluency, pronunciation) compared with oral proficiency (self-rated) 102

Table 42 Mean rankings for proficiency levels (grammar, fluency) compared with aural proficiency (self-rated) 104

FIGURES

Figure 1 Distribution of university French courses completed and time spent in a French milieu 52

Figure 2 Distribution of years teaching experience by the speaking extent of teachers in the French milieu 59

Figure 3 Distribution of University courses by the speaking extent of teachers in the French milieu 60

Figure 4 Distribution of oral performance level of teacher/interviewers for five 'can do' statements 74

Figure 5 Distribution of aural performance level of teacher/interviewers for listening situations .. 75

Figure 6 Distribution of the rating of the factors in speaking proficiency for years teaching experience 88

Figure 7 Pairs of independent variables with significant values below the .05 level 111

Figure 8 Background characteristics significantly related to proficiency in French 114

CHAPTER ONE

Statement of Problem

During the last two decades universal interest in the testing of L2 oral proficiency has grown significantly. In recent years throughout the United States and Canada, and particularly in the province of Newfoundland, second language oral interview procedures have been developed and refined to meet the educational needs of government agencies, industry and the academic setting of secondary schools and post-secondary institutions (Lowe and Liskin-Gasparro 1986, Newfoundland and Labrador Department of Education 1986, New Brunswick Department of Education 1984, 1974).

A number of research questions have been raised about the validity and reliability of oral interview tests (Underhill 1982, Carroll 1979, 1961, Spolsky 1978, Ingram 1968). Other questions have been asked concerning the structure and procedure of the tests, such as the appropriateness of direct and indirect testing techniques, simultaneous or delayed scoring, method of scoring, length of testing time, number of interviews per day by interviewers and number of raters per interview (Oller 1975, Wilds 1975, Upshur 1973, Brière 1971). More recently questions have been raised with respect to the employment of the oral interview for the development, the creation and the evaluation of curriculum guidelines and new courses (Shapson 1988), curriculum materials (Bragger 1985), and teaching and testing strategies appropriate for use in the proficiency-oriented classroom (Harlow 1990, Gonzalez Pino

1989). Additional queries concern the use of oral interviews by agencies to assist in the selection of employees such as certification of language teachers (Leblanc 1990, Hiple and Manley 1987, Magnan 1986, Van Patten 1986, Byrnes 1985, Groves 1985, Hirsch 1985, Manley 1985, Medley 1985, Valdes 1985, Higgs 1984, Johnson and La Bouve 1984, Kaplin 1984, Liskin-Gasparro 1984, Medley 1984, Omaggio 1984, Stevick 1984, Stern 1983).

Although agencies which employ the oral interview give formal and intensive training to interviewers, varied backgrounds such as professional training, language experience and proficiency account for individual interpretation of language behaviour. It is these varied profiles, and the concerns outlined above, upon which this study of the oral interview is founded.

Competence and Performance

The testing of oral proficiency may be characterized by the phenomena of competence and performance. Carroll (1968) suggested that performance varied with respect to the following dimensions: speed and diversity of response, complexity of information processing and awareness of features of linguistic competence. However, no systematic chart or taxonomy of behaviours has been developed in order to rate an individual's language performance or ability. The distinction between competence and performance was formulated by Chomsky as cited in Stern (1983:129) "'Performance' refers to the infinitely

varied individual acts of verbal behaviour with their irregularities, inconsistencies, and errors."

'Competence' according to Chomsky is the individual's capacity to abstract and develop performance behaviours in a systematic and orderly set of rules for the language. Further to and, for some linguists, contrary to Chomsky's linguistic competence, is 'communicative competence'. Although implying linguistic competence, it reflects the social and cultural view of language. Hymes (1972:277) formulates it as a competence of "when to speak, when not, and as to what to talk about with whom, when, where, in what manner."

The concept of communicative competence has far reaching implications in language pedagogy and language testing. Allen (1983), Savignon (1983), Canale and Swain (1980), among others have developed a framework for communicative language proficiency which comprises/includes grammatical competence (proficiency/capacity to master the grammatical features of a language), sociolinguistic competence (proficiency/capacity to employ a language in appropriate sociolinguistic and cultural contexts), discourse competence (proficiency/capacity to achieve unified discourse), and strategic competence, (proficiency/capacity to employ communicative strategies in order to enhance communication and compensate for breakdowns in communication). While the communicative approach has broadened the comprehension of the nature of language, language pedagogy and communication terminology is imprecise. Bachman and Palmer (1984) define the components of the Canale and Swain (1980)

communicative approach within a general framework of measurement. This writer employed the terminology as defined by Bachman and Palmer (1984) and Canale (1984) in this thesis.

Role of Language Tester

A language tester must distinguish between competence and performance and decide whether she/he is testing one or the other or a combination thereof. If she/he is testing the ability of an individual to attain communicative competence, she/he examines the acquisition of interrelated habits. If performance is the objective of inquiry, the tester measures the strength or effectivness of the habit or behaviour. Carroll suggests the following:

> The single most important problem confronted by the language tester is that he cannot test competence in any direct sense; he can measure it only through manifestations of it in performance. (1968:51).

Stern (1983:346) enhances this idea when he writes that the "concept of competence (or proficiency) is accessible only through inference from the language behaviour of the individual, his 'performance' in listening, speaking, reading and writing."

The language tester is interested in measuring valid competence and performance qualities or variables with reference to criteria, standards or performance levels which have been established. The performance and/or competence qualities of an individual are not 'unitary', that is to say, the same proficiency level for one language factor such as

pronunciation, grammar, vocabulary or fluency is not necessarily the same or equally developed for another. This notion is due to the various heterogeneous learning abilities, time and motivational factors which affect learning or acquisition of the language. Therefore, a well-articulated, all-encompassing testing procedure or rating scale must be employed in order to meet the language tester's needs.

Given the above it is important to classify or chart systematically performance outcomes at various levels or dimensions. A hierarchy or taxonomy of language behaviour is virtually impossible "because of the very difficult problem of identifying precisely many of the complex variables which define the competence of a speaker or listener in any act of communication" (Brière 1971:387). Therefore, proficiency levels, which are descriptions or characteristic profiles of individual language users, such as the ones proposed in the Newfoundland and Labrador Department of Education French 3200 Oral Testing: A Manual for Interviewers (1986:8), are utilized.

These proficiency levels have been developed for the Newfoundland and Labrador educational context and are based on the most widely accepted measures, both from a theoretical and empirically validated point of view, and are the state-of-the-art mechanism for making inferences about individuals' language characteristics or profiles. It is, therefore, both fitting and proper that an investigation of the reliability of the use of these measures for evaluating student performance be undertaken.

Language Tester Profile

Bachman and Palmer (1984:42) write that "many educators and researchers have lost sight of the fact that communication involves two parties" and success in communicative performance is dependent upon the abilities of the interlocuter(s). Thus a measure of communicative performance is not an absolute measure of success or communicating ability on the part of the testee. An interlocutor's success or failure is always, to some extent, dependent upon his or her audience. Therefore, the interviewer's background and language facility is critical to the interviewee's communicative performance.

Questions for Investigation

In this study, the writer investigated interrater reliability in three dependent variables and their relationship to interviewer profile variables thought to be of importance, namely gender, number of university courses, teaching experience, language proficiency, and time spent in a French milieu. The three dependent variables which the survey subjects rated were: acceptability of errors, characteristics of proficiency levels, and factors of speaking proficiency.

The relationship between the sets of independent and dependent variables was the primary focus for the study. The major questions of the study were:

1. Are there significant variations in the way interviewers rank in importance the five factors of speaking proficiency (vocabulary, pronunciation, grammar, fluency, comprehension)?

2. Do interviewers' scores vary significantly when rating the acceptability of different types of errors?

3. Do interviewers' scores vary significantly when rating proficiency levels as established by the Newfoundland and Labrador Department of Education in the document French 3200 Oral Testing: A Manuel For Interviewers (1986)?

4. If interviewer ratings do vary significantly, can these differences be associated with the oral/aural proficiency level of interviewers?

5. If interviewer ratings do vary significantly, can these differences be associated with background characteristics such as gender, location, university courses, years of teaching experience, and time spent in a French milieu?

A further question was added to the data collection instrument which pertained to the effect of the introduction of the oral interview on the strategies used by the teacher in the classroom. This question was: Have subjects changed their teaching strategies as a result of the implementation of the oral interview?

Procedures

The writer conducted a survey of the general population of interviewers trained by the Newfoundland and Labrador Department of Education. The survey solicited specific information which described subjects' language experiences and skills, and their professional training and experiences. The survey collected general demographic information about the sample population. The sample population were also asked to rate themselves on certain aspects of their language proficiency. The results of the subjects' ratings of themselves on the dependent variables provided data for the comparisons of interrater reliabilities.

Hypotheses

The writer hypothesized that, despite the formal training sessions offered to the interviewers by the province, differences in the ratings of the dependent variables outlined in the survey would be related to the diverse backgrounds of subjects as described by the survey's interviewer profiles. Teaching experience, time spent in a French milieu and language proficiency were the major individual variables which it was thought would support the hypothesis. The remaining profile variables were considered valuable in completing a total profile of the interviewers.

Significance of the Study

Results of the study could contribute to an understanding not only of second language oral testing but also of theory about language learning behaviours and teaching strategies. Similarities or differences in interviewer profiles may indicate reasons for differences in correlations of interrater reliability. These differences may also account for some variation in the final scores of interviewees. These types of issues are of considerable importance when comparing the reliability of examination scores for individual students.

Results of the study could also expose new areas of inquiry into language learning behaviours and second language learners. Curriculum design and other pedagogical considerations, such as teacher training, are areas of inquiry which could also be enhanced by the findings of this study.

Limitations of the Study

A population limited to only a small number of trained interviewers restrict generalization of the findings. However, results of this pilot study may encourage further inquiry. The use of a sample oral interview, as opposed to the ratings of actual oral examinations, may also effect outcomes.

The types of errors as outlined in Part Two, Question 2 of the survey are not classified or grouped. This procedure allows for only overall generalizations on acceptability of errors as opposed to insight into the acceptance of specific classifications of errors.

The question concerning teaching strategies, outlined in Part Two, Question 3 of the general section of the survey, is limited to either a negative or affirmative response. For further study of this aspect, additional or complementary information would have to be sought in a future inquiry.

Although there are several limitations to the study, one of the pilot study's strengths is an invitation to further examine the issue of the reliability of the oral interview format in assessing second language oral proficiency.

Definition of Terms

Many of the terms employed in this thesis are defined as they appear in the text. However, the following terms may need clarification so that the reader can understand the meaning attributed to them in the context of this study.

> CORE FRENCH - Core French is defined as a program of instruction in which students study the various aspects of French language for a limited number of periods per week/cycle. In most areas of Canada the regular French-as-a-second language program is designated by this term.

> PROFICIENCY - Proficiency can be seen as a goal or learning outcome and be defined in terms of objectives and/or standards which are to be attained at a particular level in a student's academic program. Assessing proficiency as an empirical fact is thus "the actual performance of given individual learners or groups of learners." Stern (1983:341).

ORAL TEST - An oral/speaking test is a formal means of measurement demonstrating achievement in the skill. The specific aural/oral skills which are generally accepted as being a part of communication are the following:

1) Aural comprehension encompassing sound discrimination and meaning.

2) Oral production encompassing:

a)	grammar	-	the discrete and/or holistic grammatical features of a language.
b)	vocabulary	-	selection and integration of lexical items for discourse.
c)	comprehension	-	skill in understanding fluent or native speakers.
d)	pronunciation	-	perception and production of sounds and basic intonation patterns.
e)	fluency	-	spontaneity of delivery, stress and intonation in appropriate socialinguistic and cultural contexts.

CHAPTER TWO

Review of the Literature

Historical Development and Identification

Trends in second language testing are inclined to follow methodological and linguistic trends in teaching (Upshur 1975). As suggested by Ingram (1968), tests are designed for specific purposes, as are language courses and pedagogy.

Ingram (1968:74), in outlining the shape of a language test, suggests that:

> "the shape of a test is determined by the purpose for which one is testing. Do we wish to predict future behaviour, measure comparative achievement...diagnose strengths and weaknesses for remedial teaching...the shape of a test is (further) determined by the field which is being tested, more precisely by the kind of behaviour we wish to sample, whether it is language behaviour or mathematical behaviour..."

Just as language testing is molded and shaped for various purposes, unique historical, cultural and geographical characteristics also influence the structure and development of courses and teaching strategies.

It may be useful to divide language teaching, testing and pedagogy into three eras or trends for purposes of identification and progress, as has been proposed by Stern (1983), Spolsky (1978), Chastain (1976) and Brière (1971).

Pre-Scientific Era

Proponents for the introduction of modern languages into the school curriculum had first to challenge the Renaissance tradition of studying the classics, Latin and Greek, from a strictly religious viewpoint. The goal of these educational reformers was to bring modern foreign languages to school curricula on their own terms. Foreign or second language learning, when first introduced into the school curriculum, was considered synonymous with literature, the learning of literary passages, translated prose and memorization of verb paradigms (Brière 1971). Methodology of the era could be described as a grammar translation method. Kelly (1969), cited in Liskin-Gasparro (1984:15), writes: "Language skill was equated with ability to conjugate and decline". No effort was made to meet the communicative needs of students. Curricula of the era dictated students' needs as being of a mental or intellectual nature and modern language study had to prove its value as a discipline.

This first era, the pre-scientific period, may be characterized, as its name suggests, by a lack of development or perhaps concern for statistical elements or notions, such as objectivity and reliability in the evaluation of results. The prevalent testing procedures developed and extensively employed were composition and dictation of literary passages or grammar translation exercises of textual or cultural interest. The scoring was basically subjective, based on the personal judgment of the examiners who were, in many cases, the

teachers. No special skill was required, and the predominant assumption was that a teacher, because of her/his teaching expertise, was well equipped to judge the proficiency or achievement level of an individual. As a result, objectives were ill-defined and each test constructor could employ various criteria for grading. Personal prejudices of the constructor, such as the relevances or importance of grammatical competence, stylistics and/or any other considerations were common and acceptable. If an oral examination was employed, it was the exception and "...the ensuing 'tests' frequently consisted of unstructured interviews or oral compositions which lead to the same chaotic conditions as those for written tests" (Brière 1971:386).

Psychometric-Structuralist Era

The second period of language teaching could be referred to as the psychometric-structuralist era. It is characterized by two groups of specialists, psychologists and structural linguists, who agreed on principles of testing but who had conflicting views on language learning theory.

Teaching.

The psychometric-structuralist period is further characterized by attempts to resolve the problem of teaching methods or strategies. The 'direct method', its principles being the use of the target language in direct relation to objects, visual aids or familiar target language words or

expressions, was the first venture from the traditional methodological style of grammar - translation. Unfortunately, the direct method was quickly abandoned, in part because it was considered too time consuming. Language teaching became very much influenced by linguistic research, with developments in vocabulary selection and reading approaches based on word counts. However, many teachers held to the security of the traditional techniques.

It was the debate on language learning theory, in particular the division between structuralists and functionalists, which encouraged the development of methods from language theory. Behaviourists, based on the functional lingustic school, who idealistically espoused the theory of complete scientific objectivity supported the stimulus - response theory of language learning. The product was the audio-lingual/visual approaches, the new methodologies which emphasized rote learning and drill procedures. In the development of these methodologies, language learning and teaching employed the new technologies of the era such as language laboratories, film strip projectors, television and computer assisted instruction. Thus, a revolution in modern language teaching characterized by educators experimenting with new ways of looking at language learning and an overall keen interest in technology predominated.

Testing.

Test construction in this climate produced universally precise, objective, reliable and scientific tests. The major concern was to provide "objective measures using various statistical techniques to assure reliability and certain kinds of validity" (Spolsky 1978:vi). The first significant product was the construction of multiple-choice "objective" tests with the common major goals of validity and reliability. Ingram (1968:74), in her review of the characteristics of the shape of language tests, writes that the initial criteria of all tests "whether predictive or non-predictive, language or non-language is primarily determined by the need to test the tests for reliability and validity. This is why, for instance, the multiple choice technique of answering is so common. In language testing it means that we usually draw upon the skills of reading and listening".

Ingram's concluding remark that multiple-choice techniques favour only reading and listening skills indicates a major deficiency in the state-of-the-art, scientific, language testing programme of the era. The new oral-aural emphasis being introduced at that time, replacing the original literary methods, was left abandoned. Thus a rift developed in what ought to have been a simultaneous expansion of teaching and testing techniques based upon linguistic principles. This division confirms Upshur's (1973) conclusion that testing tagged aimlessly along behind methods and theories of language teaching. Eventually multiple-choice and other short answer

type tests were adapted to test aural comprehension, but this type of testing, which generally can be refined to produce high levels of reliability and validity, is not necessarily appropriate to the testing of oral performance/proficiency.

The good test.

Oller (1975:184) originated the concept "that a 'good' test is one that not only provides valid and reliable information about the effectiveness of the students' learning and the teacher's instruction, but also functions as an integral part of the teaching-learning process by focusing attention on, and giving practice in, useful language skills" (Oller 1975:184). The 'good' language test is perhaps the key to teaching and testing. However, the development of the good language test continued to elude educators. The major question facing second language test developers was how to combine psychometric constructs, such as reliability and validity, with the multi-faceted outcomes of oral proficiency. Carroll (1979, 1961) makes a distinction between two language testing viewpoints. He refers to one as 'discrete-point testing' and the other as 'integrative skill or global type testing'.

Discrete point testing.

The discrete point item element was derived from structural linguistic notions of Robert Lado (1961), who suggests a non-universally agreed upon analysis of language

into three levels, phonology, syntax and semantics (Oller 1975). Carroll (1961), as cited in Spolsky (1978), notes that:

> "The work of Lado and other language testing specialists has correctly pointed to the desirability of testing for very specific items of language knowledge and skill judiciously sampled from the usually enormous pool of possible items. This makes for highly reliable and valid testing".

Brière (1971) writes that discrete point testing and teaching became the "order of the day" with multiple-choice tests allowing efficient administration, scoring, and provision for diagnostic and discriminatory purposes. There were serious disadvantages in that multiple-choice examinations require substantial expertise, time and effort on the part of the constructor in writing, proofing and pretesting. Discrete-point tests have their most profound fault in that they do not exhibit actual language usage. They request students "to perform highly artificial tasks and on the basis of one's performance attempt to infer one's level of competence for a different sort of task altogether" (Oller 1975:189). In addition to the critical remark with respect to the unnatural tasks, discrete-point testing does not complement the total program of language instruction in that such tests "fail to provide the student with practice in useful language skills...(and) may serve to confuse rather than to instruct the student concerning the points tested" (Oller 1975:189).

In addition to the above conclusions, Upshur (1973) further challenged discrete point testing on empirical grounds, concluding that results of studies (Educational Testing Service, 1968) have shown that correlations of proficiency

tests and academic grades are very low and that discrete point tests do not predict as well as work-sample tests. Oller (1978:43) writes that the "discrete-point method of teaching and testing is naive concerning the fact that a totality is greater than just a heap of unrelated parts". This view of 'totality' put forth by psycholinguistics interested as they are in language epitomizes the concepts of communicative competence and oral proficiency.

Upshur (1973:179) also challenged discrete point testing on theoretical grounds. He explicitly rejected the "notion that a linguistic description of a language can specify a set of items, the knowledge of which will assure unimpeded ability to be understood". This rejection introduces a theory proposed in the early part of this century by Jespersen who argued that, in order to teach a language, one has to teach learners to communicate in real-life situations (Ingram 1978, Oller 1973). Jespersen argued, as cited in Oller (1973):

> "...we ought to learn a language through sensible communications; there must be a certain connection in the thoughts communicated in the new language...one cannot say anything with mere lists of words. Indeed not even disconnected sentences ought to be used...when people say that instruction in languages ought to be a kind of mental gymnastics, I do not know if one of the things they have in mind is...sudden and violent leaps from one range of ideas to another."

This viewpoint typifies the theory of communicative competence, as is outlined above, and introduces the third era of language testing called the integrative-sociolinguistic.

Integrative-Sociolinguistic Era

The developments in modern language teaching parallel the trends occuring in society and education in general. In reaction to the rigorous drill techniques outlined in the previous period, curricula and course designers began to focus on the learner as an individual. Human values and relations between teacher and student called for research on the nature of the learner and the second language learning process.

The development of new approaches to second language testing have come to be of major importance with the advent of new methodologies. These new methods all reflect communicative language teaching strategies. Although referred to as new methodologies, several were developed during the last era but have received more recognition. Two such are the Silent Way (Gattegno 1972) and Community Language Learning (Curran 1976) as cited in Stern (1983). Other methods based on a similar communicative view are Suggestopaedia (Lozanov 1979), Total Physical Response (Asher 1969), the Natural Approach (Terrell 1986) and the Comprehension Approach (Winitz 1981). Although the present state of the art of language teaching theory may seem a little eclectic, its roots are founded in those disciplines which underlie language teaching: linguistics, psychology and pedagogy.

With the advent of communicative competence concept and the communicative approach to teaching, emphasis on communicative testing pushed integrative testing or global proficiency to the forefront. Carroll (1961) was the first to

construct a testing theory of "integrative testing" to deal with what he calls the "communicative effect" and "normal communication situation".

Proficiency Movement

It was the proficiency movement that best answered the question of accountability in the second language teaching profession. Its goals are an improved curriculum and a standardized evaluation based on functional and cognitive objectives where performance of learners is mandated (Magnan 1985). The face-to-face oral proficiency interview would undoubtedly be an appropriate evaluation instrument but not the sole testing technique.

The proficiency movement is more than just an oral proficiency test. Proficiency guidelines, The ACTFL Proficiency Guidelines (ACTFL 1986), have been written for all four skills (speaking, writing, reading, listening) and for cultural aspects of second language learning. The underlying theme of the proficiency guidelines is similar to that of the European approach of the functional/notional syllabus and the communicative competence theory (Lalande 1985, Wilkins 1976). This belief is further enhanced by the definition of language proficiency by Liskin-Gasparro (1984:12) as "the ability to function effectively in the language in real-life contexts". Walz (1986:13) believes that oral proficiency is "... a logical extension of the emphasis on communication we have been reading

about for fifteen years; it cannot be seen as a change in direction but rather as a concentration of effort".

As was inferred at the outset of this chapter, testing and methodogical trends should develop simultaneously. Omaggio (1983) suggests that teaching trends and methodologies should be identified by some "organizing principle" (Higgs 1984) by which various classroom methodologies, curriculum designs, program goals and instructional materials might begin to make collective sense. Clearly, more study must be undertaken in these areas if proficiency is to be the basis of future second language acquisition methodologies. Galloway (1987:36), in proposing a possible extension of the proficiency movement, suggests that there are no "sturdy bridges nor reliable stepping stones of research to join the two, one must make a speculative leap".

The proficiency oral interview and the associated ACTFL Proficiency Guidelines (ACTFL 1986) are presently being employed to develop new frontiers in language teaching (Scott 1989, Van Patten 1986, Higgs 1984, Liskin-Gasparro 1984, Medley 1984, Omaggio 1984, 1983), new courses (Magnan 1986, Hirsch 1985, Kaplin 1984, Stern 1983), new curriculum models (Byrnes 1985, Groves 1985, Medley 1985, Valdes 1985, Heilenman and Kaplan 1985, Stevick 1984, Stern 1983), and teacher training and certification requirements (Leblanc 1990, Hipple and Manley 1987, Manley 1985, Bernhardt and Hammadou 1982).

However, Kramsch (1986), Schulz (1986), Bachman and Savignon (1986) and Savignon (1985), not withstanding the

rebuttle by Lowe (1986), warn advocates that the oral inteview by itself does not have the well founded theoretical base needed in order to dictate the future direction for language acquisition and testing. They maintain that care, caution and collective strategy are the order for success when employing the ACTFL Guidelines (1986) as a framework for the development of language teaching and testing theory.

Byrnes and Canale (1987) and Byrnes (1987) propose that proficiency assessment is worthy of consideration and has great research potential. Byrnes (1987) singles out three areas which are representative of the numerous possible research issues: 1) language processes, 2) acquisition sequences, and 3) ultimate levels of attainment. Results from investigations "would be enhanced knowledge not only about learner-language and learner language use, but indirectly, about language, language learning and language instruction" (:49).

Proficiency Testing

Before discussing the issues related to, and the types of proficiency testing, it is important to make a distinction between "proficiency testing" and "achievement testing". Achievement tests examine specific features of language and are generally founded on limited amounts of knowledge presented in instructional materials. They are norm-referenced, and used primarily to provide specific feedback to both teacher and student. Proficiency tests are criterion-referenced, based on functional language ability and rated globally. They are not

limited by course materials and instructional variables (d'Anglejan, Harley and Shapson 1990, Magnan 1985, Larson and Jones 1984).

Communicative competence is one of the goals of the proficiency movement. The testing of communicative competence can be divided into two major types of oral proficiency testing, direct and indirect.

Indirect proficiency tests.

Indirect proficiency tests "derive validity, through a correlational relationship with direct proficiency tests rather than through the face/content validity of the instruments themselves" (Carroll 1978:29). If a given indirect test is found to correlate highly and consistently with direct proficiency tests, then it becomes useful as a surrogate measure of that proficiency. This means that it can be considered as a reasonably accurate predictor of performance levels (Clark 1975).

In some cases indirect tests may involve the students in some quasi-realistic oral activities, such as describing printed pictures aloud or using taped questions to elicit answers. Other indirect methods commonly employed are the pragmatic proficiency tests as proposed by Oller (1978).

"Pragmatic facts of language", as defined by Oller cited in Ingram (1978:6), are "those having to do with the relations between linguistic units, speakers and extralinguistic facts". According to this definition, the term

covers any of the interrelated activities of linguistics, psycholinguistics, sociolinguistics, and discourse and strategic capacities. Oller proposed the use of cloze tests and dictation for testing second language proficiency. One of the indications of the validity of such tests is their strong intercorrelation with each other, near the .90 level as cited in Oller (1978).

A cloze test is constructed by deleting every Nth word from a passage of prose. The examinee's task is to replace the missing words. Hanzeli (1977:865) writes "Cloze tests have a high degree of statistical reliability, are relatively easy to administer, and clearly call into play the kind of global communicative competence most language teachers are aiming for today." Although cloze tests are easily prepared and scored and appear to be a valid measure of language proficiency, Heilenman (1983) proposes a cautious use of the cloze test procedure as an alternative or supplement to other methods of student placement.

The second pragmatic proficiency test, dictation, is one that is "administered at a normal conversational rate over segments that challenge the short-term memory span of the examinees (and) ... has proved repeatedly to be an excellent device for the measurement of language proficiency" (Oller 1978, Oller and Streiff 1975).

Pragmatic tests such as the cloze test and dictations are integrative global proficiency tests which, according to Oller (1978:56) "sample the non-native speaker's ability to do

what native speakers dc in the normal use of language". This "normal use of language" is what Jespersen proposed to be the goal of the language teaching exercise (Farhady 1979, Ingram 1978, Oller 1973).

The ever increasing employment of computers in today's schools has given use to this technology as both a teaching and testing instrument. Dandonoli (1987) proposes the use of computers for the administation of proficiency tests. Each test would require a bank of items to be readily available, thus providing a more efficient, accurate and effectively rewarding test than traditional indirect proficiency tests. Computerized adaptive tests are a "process of estimation of ability and subsequent item administration continues until some arbitrarily predetermined criterion of measurement precision has been achieved" (:85). This is similar to level checks and probes employed in the oral proficiency interview. Although Dandonoli reported on computerized adaptive tests solely for the receptive skills, this writer believes that innovative teachers and programmers can develop computerized surrogate oral proficiency tests.

Semi-direct proficiency tests.

An alternative to direct, face-to-face oral testing is a technique proposed by Larson (1984) at Brigham Young University. Semi-direct testing is based on testing techniques which elicit active speech by the examinee through means such as tape recordings, printed test booklets, or elicitation

techniques other than direct interviews. Results of a study comparing semi-direct and direct tests showed correlations of .69 for an indirect testing technique for German and .89 for a direct testing technique for Spanish. However, student preference indicated that although direct testing provoked nervousness it was the overall preferred technique.

Direct proficiency tests.

Direct proficiency testing is usually interpreted to mean a replication of the real-life language use situation, and is characterized by high face/content validity. In the testing of oral proficiency, possible direct testing procedures include reading aloud, presenting a prepared speech, a small group discussion, playing a game, conducting a survey, talking on the telephone and, of course, conversing face-to-face with one or more interlocutors.

Although the face-to-face interview may be the most life-like oral assessment interview technique, and is believed to be a valid measure of proficiency (Backman and Palmer 1981:68), contextual, psychological and affective elements may differ somewhat from the natural setting. As Perren (1968) points out" ... both participants know perfectly well that it is a test and not a tea party, and both are subject to psychological tensions, and what is more important, to linguistic constraints of style and realistic thought appropriate to the occasion by both participants" (:26). The

interview situation is, however, the most practical of any natural setting in order to observe individuals.

Byrnes (1987) poses an interesting question as to whether the oral interview in fact taps real language proficiency rather than test-taking behaviour. Byrnes (1987:44) reports that the issue is unresolved as the "evidence goes both ways, for an enhancing as well as a debilitating influence of testing."

In addition to being valid, direct, indirect and semi-direct, proficiency tests must be reliable in that they must provide consistent and replicable information about an individual's performance. Results of studies have demonstrated that direct measures of oral proficiency may be as reliable as the indirect methods of more structured or standardized tests. The oral proficiency interview has generally been shown to yield respectable evaluation results of interrater reliability (Bachman and Palmer 1981, Clifford 1980, Adams 1978, Clark 1972, and Carroll 1967). However, a number of questions have still been raised about the validity and reliability of the oral interview as an evaluative measure.

Research Questions

Components of the Speaking Skill

Harris (1969), as cited in Mullen (1980), proposed five basic factors or components of the speaking skill: pronunciation, grammar, vocabulary, fluency, and comprehension. Higgs and Clifford (1982) employed the factors of the oral proficiency interview as variables in an attempt to investigate the relative contribution of the different factors to global language proficiency. Results indicated a fluctuating relative importance of the factors with an equal contribution of subskills at the global performance rating at Level 5, the highest oral proficiency level. Other researchers such as Magnan (1988), Byrnes (1987), Bachman and Savignon (1986), Kramsch (1986) and Savignon (1985) felt that advocates of the oral proficiency interview do not apply an equal balance of emphasis to all components of the speaking skill as described by Canale and Swain (1980) in their definition of communicative competence. Conclusions such as these raise questions about the validity of the oral interview.

Errors

Concern about student errors, both in teaching and testing for proficiency, has gained support from research on native speakers' judgement of errors (Lantolf and Frawley 1985). Studies by Gynan (1985), Khalil (1985), Delisle (1982), Ensz (1982), Ludwig (1982), Piazza (1980) as cited in Omaggio (1984:78), Chastain (1980,) Galloway (1980), and Guntermann

(1978) have all revealed that both lexical and grammatical errors are considered obstructive to communication. Ensz (1982:138) also found that "...the French people...irrespective of sex, age, occupation, or home region in France, expressed a significant intolerance for grammatical errors." Chastain (1980) wrote that the errors which are conceived to be the most troublesome are generally the basis of the major focus of the classroom language curriculum. Instructor perception of errors, therefore, not only affects the selection of curriculum content and activities but also grading procedures. If there are major differences between instructor perceptions and native intolerance of error, the validity of the oral interview could also be affected.

Reliability of Student Performance

The validity of the test itself is a major concern as discussed above. However, there are three elements common to all types of language tests, the student, the test and the marker (Underhill 1982). The student is a major factor in the direct oral proficiency test. There are numerous affective variables such as attitude, motivation, fatigue, mood, health, related to the interviewer which may alter the reliability of a test. These considerations affect the reliability of the result for the particular student being interviewed.

Interviewer Reliability

There is also a third element in the case of the oral interview, the marker or the interviewer. As affective variables of students alter their performance during an oral examination, certain characteristic traits of interviewers may alter ratings of oral tests. Problems of rater reliability are not related to overall rater reliability of the same performance or with simultaneous or delayed scoring. The problem lies with certain characteristics or personal traits of the interviewer. The question may be raised: Does the proficiency level of raters, or interviewers, cause their judgements or ratings of acceptability of errors, of proficiency levels, and of the relative importance of the five factors in speaking proficiency to vary? If indeed there is a significant correlation between the proficiency levels of interviewers' judgements in these matters, more importance ought to be attached to the question of competency and proficiency of examiners. Another question as posed by Wilds (1975) concerns the possible daily fluctuations of the rating of oral interviews and whether these fluctuations are affected by background characteristics of interviewers and or raters. If the results of research studies indicate that background characteristics do affect rater or interrater reliability, then these characteristics need to be given the importance they deserve in selecting appropriate interviewers.

Summary

The historical development of second language learning/acquisition has made a full revolution. The initial goal of foreign language teaching and learning was a "usable level of skill." Liskin-Gasparro (1984:13) reports that this goal was initially proposed in 3000 B.C. and again in Renaissance Europe where authors of text books written in the target language envisaged today's goal of practical communicative ability.

Testing of a second language has also completed the revolution. The initial oral testing technique used was a direct practical and/or functional use of the target language. Conquering or conquered peoples of Spartan dynasties were mandated to employ foreign languages during the initial integration periods of the cultures. Today's oral interview is the new 'state of the art' of second language proficiency testing. Thirty years of use and research have demonstrated that the oral interview does possess a degree of validity and reliability as a procedure for testing second language oral proficiency. Advocates of the procedure believe that the problems which remain are for the most part logistical rather than theoretical. However, not all aspects of the oral interview have been fully researched or investigated. It is to some of the uninvestigated areas that the various questions raised in this thesis are related.

Beginning with Carroll (1961), there has been a developmental trend in the field of language testing from the

focus on discrete-point item testing to the current global or interactive skill type testing of the language proficiency movement. Prior to the developments of very recent years, little progress was made as initiatives to incorporate oral assessment procedures were stymied to a great extent by the limited language proficiency of the instructor. However, with the expansion of the use of the oral interview technique, the oral test has become a major issue of concern for language educators. The question of the validity and the reliability of an oral test which attempts to determine communicative competence are the major focus of study for investigators.

Since the language behaviour of both interlocutors in an interview situation is dependent upon each other, the role of the interviewer is of considerable importance in determining the language proficiency of the interviewee. It would therefore appear that the background, language and professional development of the interviewer is of major importance in determining the reliability of the oral interview as a measurement device. It is on these characteristics of the interviewer that this thesis is focussed.

CHAPTER THREE

The Study

Introduction

The curriculum area of interest in this study is that of French-as-a-second language with a focus on the oral proficiency interview. The literature indicates that discrete point testing is a highly reliable testing procedure (Spolsky 1978, Brière 1971, Lado 1961) but it does not test communicative competence. Numerous studies have also shown a high degree of validity and reliability for second language oral proficiency testing (Bachman and Palmer 1981, Clifford 1980, Adams 1978, Clark 1972, and Carroll 1967). However, some questions still remain as not all aspects of the oral interview have been fully investigated. The effects of professional background and language proficiency on interrater reliability is one of these areas. This study was developed in order to examine more closely the relationship between interviewers' characteristics and the rating of an oral interview.

In designing the study the researcher hypothesized that significant variations in interrater reliability for specific aspects of a sample oral interview test might be associated with the interviewers' background characteristics, in particular affecting both aural and oral proficiency levels. A survey was conducted of teachers in the province of Newfoundland who had been trained to use the oral interview for the provincial senior high school French program (French 3200).

The researcher chose to employ a survey method in order to include as large a number of the qualified interviewers in the province as possible. Operational and economic constraints made employment of actual test interviews, either taped or live, unfeasible. Therefore, a sample test was composed and employed, similar to the procedure used in the research of Higgs and Clifford (1982). The present study (as well as that of Higgs and Clifford) was not based on actual test interviews but on the choices and ratings of trained interviewers when given a description of interview situations.

Instrument

The instrument was designed to collect information about the interviewer's background and about his/her judgements in a described interview situation. A copy of the questionnaire is included as Appendix A. The questionnaire was composed of two parts: the first section requested information about the background characteristics of the teacher interviewers; and the second section asked subjects to rate various aspects of a sample oral interview.

The first part of the questionnaire employed the theoretical basis and structural format of a national language teachers' survey conducted by the Educational Testing Service of New Jersey (E.T.S.) in 1985. E.T.S. reported that it was possible to design a questionnaire that subjects could complete in a short period of time and in which teacher/interviewers

could rate their oral language proficiency with validity and reliability.

In Part One of the questionnaire demographic information, such as gender and location, was sought from the subjects. Although these areas were not thought to be major factors in the E.T.S. survey, it was thought that they might be of significance for this study. The researcher believed that the high number of experienced male secondary school teachers in the province might not place as much emphasis on the communicative aspects of the interview situation, as they had probably completed their academic training prior to the emphasis on the development of the oral skill. Information concerning location of subjects, rural or urban, was also sought because it was believed that teachers who were more proficient in oral skills might reside in the urban areas. This point is further discussed below.

Additional biographical information such as language courses completed, and teaching experience was also sought. E.T.S. (1985) established that the question of formal instruction and the number of university language courses, in comparison to direct experience, are of vital importance in establishing the self-rating proficiency and biographical profiles of subjects.

A series of questions was posed in order to establish subjects perceptions of the oral/aural language proficency. The initial questions established the length of time spent in the French milieu and also the extent to which the target language

was spoken in the milieu. E.T.S. (1985) reported that in assessing oral language proficiency it was useful to inquire about experiences in the milieu. It is believed that those experiences are related to language proficiency.

Complementing experiences in the milieu the next question established the extent of speaking the target language while at home and with acquaintances. Oral language proficiency is also believed to be related to subjects' use of the target language at home and with friends. Thus practice, even when not perfect, improves oral language proficiency. E.T.S. (1985) reported that little or no contact with the target language may indicate low oral proficiency levels. It was felt that speaking extensively, at home and in the milieu, would provide some significant correlations affecting interrater reliability.

Complementing the above, information on the actual proficiency levels of oral and aural skills was sought. The subjects were asked to judge their own level/ability in each skill. The questionnaire presented a series of paragraphs describing speaking and listening levels of spoken French. Each subject was asked to indicate which description most accurately described his/her listening and speaking ability. The descriptions were successfully employed by E.T.S. (1985).

The questions posed in the first part of the survey were believed to provide both an overall and discrete item profile of the background characteristics of the teacher/interviewers.

The second part of this questionnaire was composed of two major themes. One theme posed several general questions

pertaining to the procedures of the French 3200 oral test and related teaching strategies or methodologies. The other theme explored the ranking of various aspects of oral proficiency, certain described speech samples.

The general questions in Part Two of the questionnaire were believed to give, first of all, an indication of some administrative concerns about the French 3200 Oral Interview test. In addition, subjects were asked to indicate specific aids employed in determining interviewee proficiency levels and also to cite any possible external factors which might influence the overall rating of an interview. It was argued in the literature that oral proficiency testing and the proficiency movement as a whole is the basis for the L2 proficiency orientation of new courses and the development of new methodologies and curriculum guidelines. Therefore, inquiries about these items, would produce relevant information about whether new testing methods influenced teaching strategies, and also general knowledge of program guidelines. This information was requested in order to assist in improving the administration of the oral interview, and the results are discussed separately.

The second theme pertained to the major focus of this study. In this section, respondents were requested to perform three tasks: (1) to rank the five factors in speaking proficiency; (2) to rate the acceptability of certain described errors; and (3) to rate described proficiency levels. Each item featured specific examples for subjects to rank or to rate.

As was discussed in Chapter Two, the literature has indicated that researchers such as Higgs and Clifford employed the ranking of five factors of oral proficiency in their studies in order to describe tester perception of the roles of the subskills. The subskills or factors employed here are similar to those used in the Harris (1969) study and are also characterized as generally acceptable in the term 'oral test' under the heading, Definition of Terms, in Chapter One of this study. Respondents were asked to rank according to their perceptions, the order of importance of the five factors: pronunciation, grammar, vocabulary, fluency and comprehension. It was felt that significant differences exist in the importance given to the five factors which might influence interviewer ratings, and which might be related to specific background characteristics of the teacher interviewer.

The second question in Part Two of the survey asked subjects to rate the acceptability of a range of errors. The literature indicated that in proficiency testing, errors are a concern for researchers because lexical and grammatical errors are believed to be obstructive to communication. The writer believed that, because of possible differences in background characteristics and proficiency levels of interviewers, there might be both significant and interesting differences with regard to the acceptability of various types of errors.

The last question of the survey asked inteviewers to rate, according to the proficiency levels outlined in the French 3200 oral examination manual, descriptions of sample interviewees'

oral proficiency characteristics. These sample interviewee descriptions were composed from descriptions outlined in the French 3200 Oral Testing: A Manual for Interviewers (1986). Although the literature has indicated generally high correlations of interrater reliability in oral proficiency testing, the researcher felt that, if the background characteristics and proficiency levels of raters varied widely, there might be significant differences in the ratings given to the described speech samples.

As previously indicated, the general framework and theoretical basis of the questionnaire were developed from the E.T.S. (1985) national language teachers' questionnaire. The language descriptions used in the survey were adapted from the interviewer's manual for the French 3200 oral test, thus ensuring validity for the Newfoundland situation. During its development, the questionnaire was subject to analysis by the writers' thesis supervisor. It was also administered to several graduate students and resource persons. This critical analysis resulted in revisions being made.

Sample Population

The sample population of 90 subjects for the study was selected from a group of one hundred eighty senior high school French teachers and program coordinators. The entire population had participated in an oral interview training session. The training session for test interviewers was an intensive two day workshop sponsored by the Instruction Division of the

Department of Education. The Instruction Division workshops were held in the spring of 1987 and 1988 in several locations throughout the province. The Instruction Division training workshops were comprised of sessions on the theory, historical background and development of the oral interview. Each participant had the opportunity to observe and conduct live face-to-face interviews. Participants were required to rate five taped interviews of various levels and, in consultation with a workshop leader, discuss and/or defend their ratings of the taped interviews. This procedure was in preparation for the interviewing and rating of their own students as part of the development of the shared evaluation system for the French 3200 Public Examination.

The survey was conducted in early June during the time scheduled for the oral interviews. The oral interviews are scheduled in individual schools throughout the province during the first two weeks of June. This time was chosen to ensure that all those surveyed were as familiar as possible with the procedure and proposed rating scale of the French 3200 Oral Testing Manual.

The total population of interviewers throughout the island portion of the province was examined and lists of participants selected geographically. The lists employed were those composed by the Curriculum Division for the workshop training sessions. As a little more than half the participants attending the regional training workshops were located in the central and western regions of the province and the remaining participants

were located in the eastern and Avalon regions, two major groups were constructed, east and west. This division was made to ensure an equal distribution throughout the island. No participants were surveyed in Labrador. These two major groups were again subdivided into rural and urban groups. Rural and urban groups were defined by the existence or the non-existence of a French immersion program. Therefore, teacher/interviewers allocated to urban groups were those who were assigned to schools in communities or neighbouring communities where a French immersion program was offered. It was also felt that such a division would be helpful in profiling the subjects' speaking extent at home and with acquaintances, as it is believed that the presence of a French immersion program may increase opportunities to speak the target language.

Selection.

As can be seen in Table 1, there was a nearly equal distribution of gender in each of the four quadrants and a proportionate division of urban/rural in both the east and west divisions. Such divisions permit a full view of the target population. The researcher believed that this selection process ensured a broad and representative sample of the teacher/interviewer population of the province.

TABLE 1 DISTRIBUTION OF SAMPLE POPULATION

	Rural		Urban		
	Male	Female	Male	Female	TOTAL
West	37	33	10	10	90
East	33	29	11	17	90
TOTAL	70	62	21	27	180

Table 2 shows the number of questionnaires mailed out to subjects and the number returned. Subjects were grouped by gender and location, and urban/rural divisions (as opposed to east/west divisions), as the former division was of more importance to the researcher than the latter.

TABLE 2 Questionnaires: Mailed Out/Returned

Divisions	Number Mailed Out	Number Returned	Percent Returned
Male rural	39	37	95
Female rural	27	22	81.4
Male urban	10	7	70
Female urban	14	10	71.0
TOTAL	90	76	84.4

Although the majority of the questionnaires were completed and returned within a short period of time, there were some questionnaires still outstanding. The researcher endeavoured to contact all those who had not yet returned these questionnaires

and was successful in retrieving a number of those not returned. This follow-up activity aided in achieving the high rate of return of 84 percent.

Although 76 questionnaires were returned, one completed questionnaire was removed before the commencement of the analysis of the data. The reason for this was the extremely high number of hours the subject had spent in a French milieu. In addition to this fact, the other biographical information and high levels of proficiency reported led the researcher to assume that the subject was a francophone. Therefore, it was felt that this particular subject should not be included in the group being studied in order to retain a homogenous group.

Data Analysis

There were five major questions posed in the initial chapter of this thesis. Three of these questions inquired directly into the reliability of interviewers' scores as related to rankings of the three dependent variables of the study. The three dependent variables were: (1) a ranking in order of importance of the five factors of speaking proficiency: pronunciation, grammar, vocabulary, fluency and comprehension; (2) the acceptability of certain types of errors; and (3) ratings of sample descriptors of the oral test proficiency levels as outlined in the French 3200 Oral Testing: A Manual For Interviewers.

A fourth question was dependent upon significant variations in interviewers' ratings for the dependent variables in the study. It postulated that there would be significant variations and asked if these differences were associated with the interviewers' profile.

The fifth and final question concerns the possible change in teaching strategies as a result of the implementation of the oral interview. This is a self assessment type of question requiring only a yes/no answer.

A descriptive analysis of the data collected in the first part of the questionnaire was prepared based on a comparison of observed and expected values of the independent variables: location; gender; time spent in a French milieu; number of university French courses; number of years teaching experience; speaking extent in the milieu; speaking extent at home and with friends; level of speaking ability (self-rated); "can-do" statements; and listening comprehension. This procedure was followed in order to attempt to develop a profile of the interviewers' background and his/her communicative competence.

In addition, an analysis of variance was employed as a means of studying the interviewers' ratings to determine whether or not significant difference exist in the three dependent variables: (1) factors of oral proficiency; (2) acceptability of errors; and (3) described oral proficiency levels. Tests of significance were performed to determine significant differences in the interviewers rankings of the selected dependent variables.

These procedures were followed in order to discover whether there were significant variations in the rankings given to the dependent variables by the interviewers. Where significant differences did occur, the attempt was made to establish a relationship between the differences in the rankings and the background characteristics of the interviewers involved.

CHAPTER FOUR

Profile of the Respondents

The first part of the questionnaire was designed to elicit information on the demographic and background characteristics of the respondents. These data were then analyzed in order to develop profiles of the respondents which would later, in the analysis of the data, be applied in interpreting the results of the analysis of variance of the dependent variables.

Results of the questionnaire indicated that somewhat more than half, 57 percent or 43 of the respondents, were male. Seventy-seven percent or 58 of the respondents were from rural areas. This result is consistent with the overall teacher/interviewer population of which 50 percent was male and 76 percent was teaching in rural areas at the time the survey was being conducted.

In reporting the findings in this section, the frequencies for the responses to the questionnaire are first discussed. Where there are interesting relationships with the responses to other questions, these relationships are then indicated, and the results of cross-tabulations of responses discussed. Tables indicating cross-tabulation relationships are also presented. These tables show the observed values and, in brackets, the expected value figures which represent the predicted normal values for each category for the number of respondents surveyed. Differences between the observed values and the expected values indicate a relationship of interest between the two variables involved.

Teaching Experience

Number of years.

Table 3 indicates that respondents to the questionnaire had a wide variety of years of teaching experience.

TABLE 3

Years Teaching Experience Compared With Gender

Gender	Years of Experience					Total Number of Respondents
	1-2 years	3-5 years	6-10 years	11-15 years	16+ years	
Male	3(4)*	5(9)	14(14)	9(9)	12(7)	43
Female	4(3)	10(6)	10(10)	7(7)	1(5)	32
TOTAL	7	15	24	16	13	75

*The numbers in brackets are the expected values.

The range of years of teaching experience extended from the category 1 to 2 years to that of 16 years or more. The largest number of teachers were in the category 6 to 10 years of teaching experience. This result denotes a somewhat younger than average teaching population as the average number of years teaching experience for the general teaching population in the province would be in the would be in the 11 - 15 years category. The median for full-time teachers in the province is 16.2 years (Newfoundland and Labrador Department of Education. 1991).

It is also interesting to note that 17 percent, or 13 of the respondents, were already teaching at the period of change to an emphasis on oral and aural skills. This change was initiated by the methodology introduced for the teaching of French programs in Newfoundland in the early 1970's.

Gender.

It is interesting to note the higher than expected number of males in the most experienced teacher category (16 or more years). This finding is explained by the fact that fewer female teachers taught at the senior high school level in past years. The differences at the lower extremity may be explained by the fact that more females now graduate with a major in French from Memorial University and enter the teaching profession at the Senior High School level.

Academic Background

University courses.

As indicated in Table 4 most of the subjects had a good academic second language background. At Memorial University a major in French consists of 12 courses. Seventy-three percent of the respondents had completed enough courses for a major in French. The average number of courses completed was 16.4.

TABLE 4

Academic Qualification as Determined by University Courses

Number of Courses	Number of Respondents	Percentage of Respondents
0 - 3	0	0
4 - 9	6	8
10 - 12	14	19
13 - 18	29	38
19 - 24	14	19
25 (or more)	12	16
TOTAL	75	100

Time spent in a French milieu.

Table 5 indicates that, in addition to academic training, many subjects have spent a considerable amount of time in a French milieu, as a student, working or on vacation. The range of time spent in a French milieu extended from less than 2 months to more than two years. However, the average length of stay in a French milieu was 11.2 months. Only 12 percent or 9 of the respondents had spent less than 3 months, the equivalent of one academic semester, in a French milieu. Students completing a major in French have been required to spend at least six weeks in a French milieu. Since three quarters of the teachers have completed the required number of courses for a French major, they probably have spent at least one semester in a French milieu.

TABLE 5

Time Spent in a French Milieu

Number of Months	Number of Respondents	Percentage of Respondents
0 - 2	9	12
3 - 4	14	19
5 - 9	21	28
10 - 14	7	9
15 - 24	20	27
25 (or more)	4	5
TOTAL	75	100

Figure 1 shows a superimposed histogram comparing the distribution of academic qualifications, or number of university courses, and the number of months spent in a French milieu. It is interesting to note that the distribution for the number of courses is approximately symmetrical or normal while the distribution for the time spent in the milieu is bimodal.

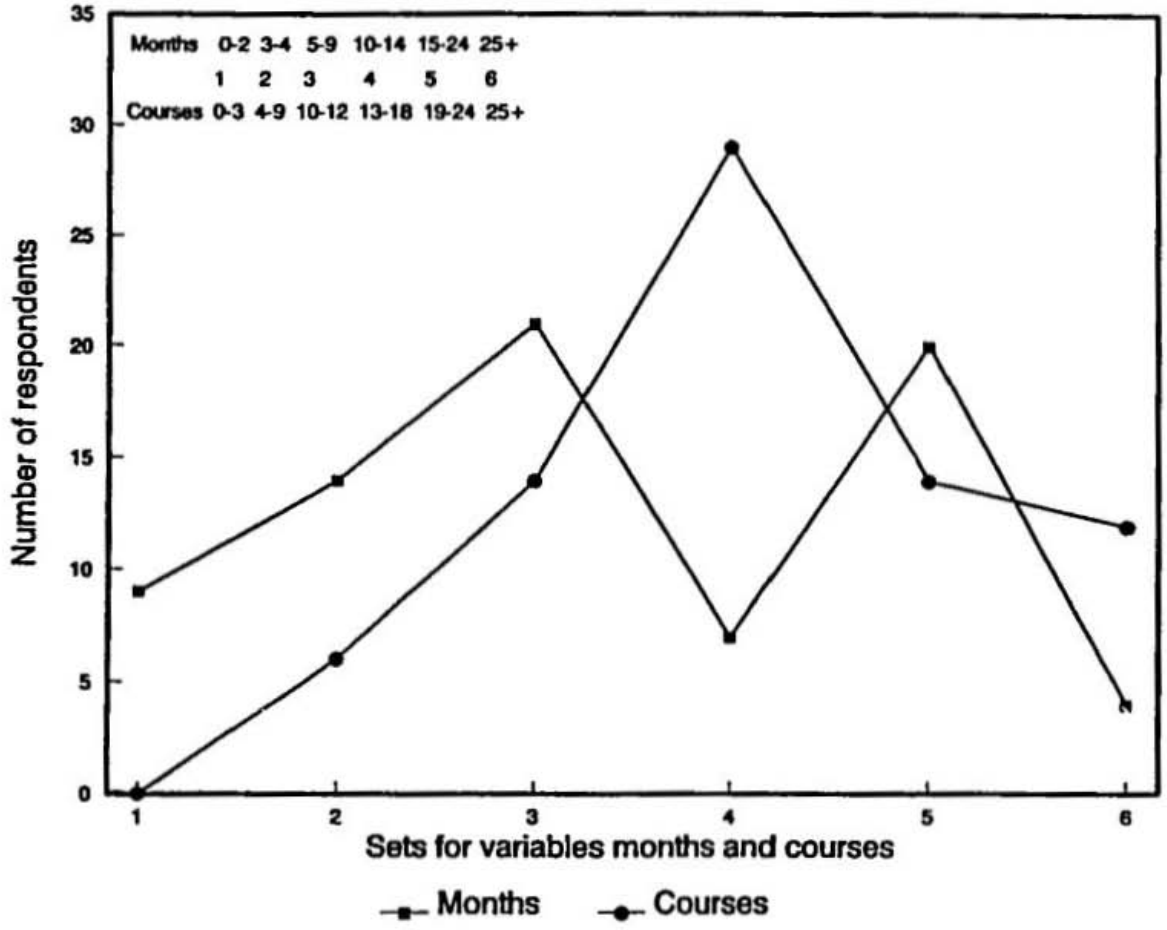

Figure 1: Distribution of university courses completed and time spent in a French milieu.

Table 6 indicates that a number of teachers completing séjours in a French milieu have taken more than the expected number of academic courses. A possible explanation of this finding for respondents who have spent only one semster in a French milieu may be the organization and the delivery of the French program at Memorial University where students are encouraged to spend one semester of study at the Institut Frecker in St. Pierre in the second year of their university program. This Institut provides an opportunity for students to spend up to three months in a French milieu and receive credit for five academic courses. Another explanation may be that experienced teachers who have completed a degree program are now beginning to develop oral French skills as they may have completed their academic training prior to the promotion of the emphasis on oral/aural skills.

Table 6 also indicates that subjects tend to continue to complete courses as they spend more time in the milieu. These findings may indicate that the links teachers have with the French milieu are mainly for academic purposes. For those subjects who have completed more courses than would be expected compared to the time spent in a French milieu, it may be that they have taken courses at home in place of travelling to a French milieu. Distance, and financial and time constraints, provide some of the reasons for these results.

TABLE 6

Time Spent in a French Milieu Compared With University French Courses

Months in Milieu	Number of University French Courses				Number of Respondents
	1 - 9	10 - 12	13 - 18	19+	
0 - 2	2(1)*	4(2)	3(4)	0(3)	9
3 - 4	1(1)	6(3)	6(5)	1(5)	14
5 - 9	1(2)	3(4)	11(8)	6(7)	21
10 - 14	0(1)	0(1)	3(3)	4(2)	7
15 - 24	1(2)	1(4)	6(8)	12(7)	31
25+	1(0)	0(1)	0(2)	3(1)	4
TOTAL	6	25	29	26	86

*The numbers in brackets are the expected values

These findings may help explain the distributions in Figure I, where the variable university French courses follows a normal pattern, while the variable, time spent in a French milieu, is shown to be bimodal.

Comparison between time spent in a French milieu and number of years of teaching experience.

Table 7 indicates that a larger number than expected of teachers with limited years teaching experience are spending more time in the milieu. It may be that the 3 month duration for the Institut Frecker at St. Pierre influences this finding. It may also be that the teachers in the less experienced

category feel the need to develop oral competence to teach the new oral programs.

Furthermore, it may be that teachers having made the initial contact of 3 to 4 months with the French milieu by the completion of their first five years teaching experience continue to maintain a contact with the French milieu.

The more experienced group of teachers may have already completed their degree at Memorial University before the commencement of the 3 months Institut at St. Pierre. However, Table 7 indicates that nearly half or 14 of the most experienced group of 29 teachers either continue to maintain a contact with the French milieu or had spent a large amount of time in the milieu during their academic training.

TABLE 7

Time Spent in a French Milieu Compared With Years Teaching Experience

Months in Milieu	Years Teaching Experience					Total Number of Respondents
	1 - 2 Years	3 - 5 Years	6 - 10 Years	11 - 15 Years	16+ Years	
0-2	0(1)*	1(2)	2(3)	3(2)	3(2)	9
3-4	3(1)	4(3)	3(4)	1(3)	3(2)	14
5-9	3(2)	5(4)	8(7)	3(5)	2(4)	21
10-14	0(1)	1(1)	2(2)	2(1)	2(1)	7
15+	1(1)	4(5)	9(8)	7(5)	3(4)	24
TOTAL	7	15	24	16	13	75

*The numbers in brackets are the expected values

Oral Proficiency

Speaking extent in a French milieu.

Complementing their length of stay in a French milieu, respondents indicated that considerable effort was made to speak the target language as shown in Table 8. Forty-seven percent or nearly half of the respondents reported speaking French for all situations while in a French milieu.

TABLE 8

Speaking Extent in a French Milieu

Situation	Number of Respondents	Percentage of Respondents
Occasional social situation	7	10
Consistently in social situations	32	43
Consistently in all situations	35	47
Total	74	100

When compared with the number of university courses taken as indicated in Table 9, those respondents who had completed only ten to twelve academic courses appear to have made less effort to use French when in a French milieu than would have been expected. This group would represent those who have followed the regular program for French majors at Memorial University. It may be that their motivation to improve their oral skills was somewhat less than those who are the other extremes of the continuum, that is, just beginning or continuing to develop their oral competence.

TABLE 9

Speaking Extent in a French Milieu Compared With University French Courses

Speaking Extent	Number of University Courses					Total Number of Respondents
	1-9	10-12	13-18	19-24	25+	
Occasional social situation	0(1)*	6(1)	1(3)	0(1)	0(1)	7
Consistent social situation	4(2)	5(6)	12(12)	7(6)	4(5)	32
Consistent in all situations	1(2)	3(7)	16(14)	7(7)	8(6)	35
TOTAL	5	14	29	14	12	74

*The numbers in brackets are the expected values

As indicated in Table 10, it appears that the most experienced teachers seemed to make less effort to speak French in the milieu than would be expected. On the other hand, respondents with a fewer number of years experience seem to have made considerable effort to use French in the milieu. Teachers with 16 or more years teaching experience had most likely already completed their professional university training when the promotion and development of aural and oral skills commenced. Teachers with 1 to 5 years teaching experience would need to qualify themselves for the new emphasis.

TABLE 10

Speaking Extent in a French Milieu Compared With Years Teaching Experience

Speaking Extent	Number of Respondents					Total Number of Respondents
Speaking Extent	1-2 Years	3-5 Years	6-10 Years	11-15 Years	16+ Years	
Occasional social situations	1(1)*	2(1)	0(2)	0(1)	4(1)	7
Consistent social situations	2(3)	3(7)	12(10)	9(7)	6(6)	32
Consistent in all situations	4(3)	10(7)	12(11)	6(7)	3(6)	35
TOTAL	7	15	24	15	13	74

*The numbers in brackets are the expected values

Figures 2 and 3 show histograms comparing teachers' speaking extent in the milieu with years teaching experience and university courses completed. It is interesting to compare the distributions for teachers speaking only occasionally with the distributions for the other groups. Figures 2 and 3 are similar in that both have a normal distribution for those who spoke French consistently in social or in all situations.

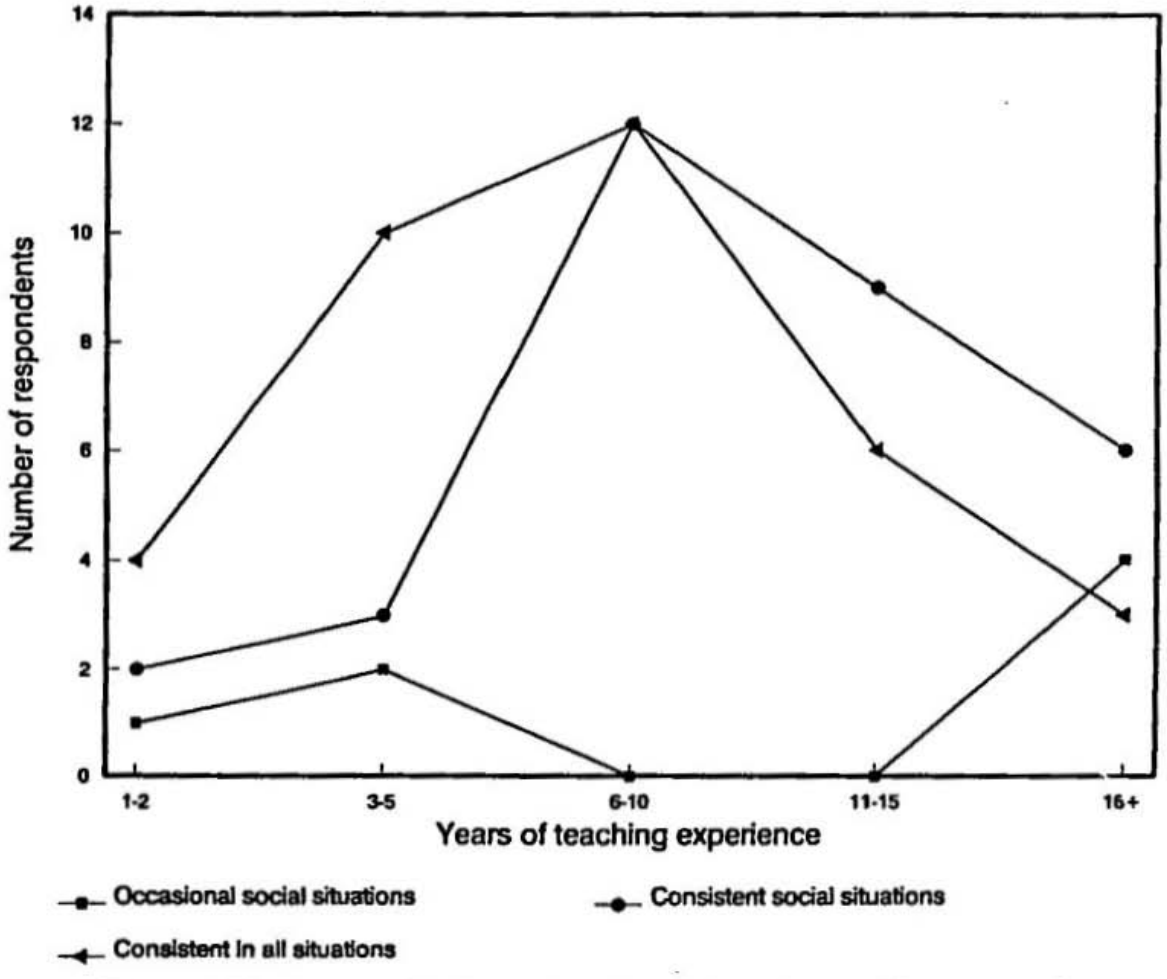

Figure 2: Distribution of years teaching experience by speaking extent of teachers in the French milieu.

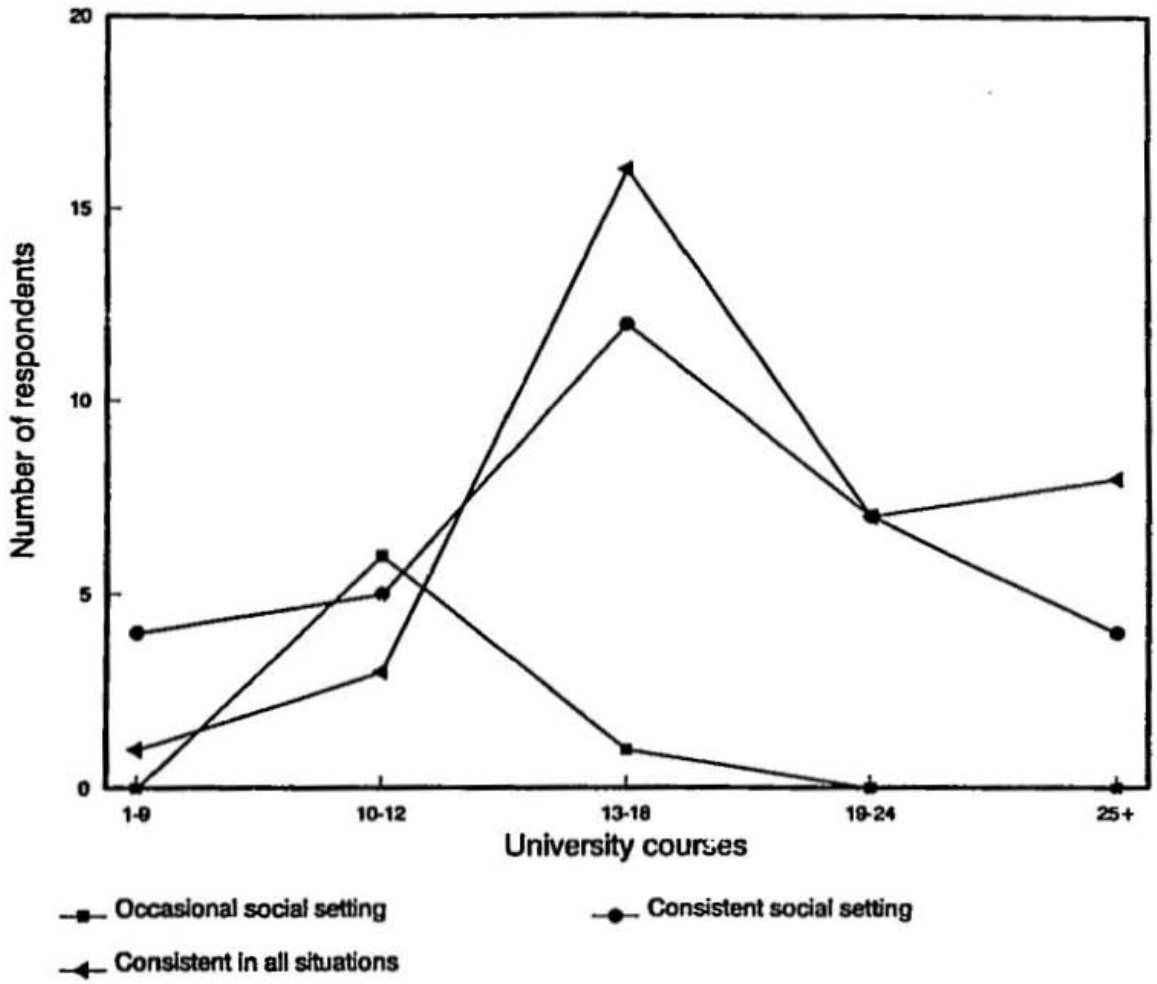

Figure 3: Distribution of university courses by speaking extent of teachers in the French milieu.

Speaking extent at home and with acquaintances.

As indicated in Table 11, despite all the effort expended in speaking French in the milieu, the use of oral French does not seem to have been transferred to the subjects' home and social settings in the province. These findings are not unexpected in the Newfoundland context. Only 13 percent of the respondents (10) speak French fifty percent of the time or more with acquaintances. An even smaller percentage of respondents use French at home. About two thirds of the respondents never use French at home; eighty-five percent use French rarely or occasionally with friends and acquaintances. As indicated in Tables 12 and 13, it is those respondents who have little exposure to the French milieu (4 months or less), less than 13 academic courses in French or more than 16 years teaching experience, who tend to use French least at home and with friends and acquaintances. These findings suggest that it is only those respondents who have studied French beyond the requirements for a French major who use French outside the classroom. Given these findings, the question of transfer of target language use to the second language classroom for the majority of teachers might be raised.

TABLE 11

Speaking Extent at Home and With Acquaintances

Amount of French Spoken	Number of Respondents	
	Home	Acquaintances
Never/rarely	50	6
Occasionally	21	59
Regularly/half time or more	4	10
TOTAL	75	75

TABLE 12

Speaking Extent in a French Milieu Compared With Speaking Extent with Friends and Acquaintances

Extent in Milieu	Number of Respondents			Total Number of Respondents
	Never	Occasional Conversation	Regularly	
Occasional social situations	2(1)*	5(5)	0(1)	7
Consistent social situations	4(3)	27(25)	1(4)	32
Consistent in all situations	0(3)	26(27)	9(5)	35
TOTAL	6	58	10	74

*The numbers in brackets are the expected values

TABLE 13

Speaking Extent in a French Milieu Compared With Speaking Extent in the Home

Extent in Milieu	Speaking Extent in the Home			Total Number of Respondents
	Never or rarely	Half the Occasionally	Time	
Occasional social situation	7(5)*	0(2)	0(1)	7(8)
Consistent social situation	26(22)	5(9)	1(2)	32(33)
Consistent in all situations	17(24)	15(10)	3(2)	35(36)
TOTAL	50(51)	20(21)	4(5)	74(77)

*The numbers in brackets are the expected values

Oral proficiency (self-rated).

Subjects were asked to judge their own proficiency level according to the E.T.S. rating scale ability. As reported in Table 14, one-third of the respondents (35 percent) indicated an 'average' level of speaking proficiency. The term 'average' is employed here to denote those who have a better than limited ability and who can discuss social topics. A 'high' level of speaking ability denotes fluent or unconstrained ability. Only twelve percent of the respondents placed themselves in this category. None of the respondents rated themselves at a very basic or limited level of speaking ability. A little more than half of the respondents (53

percent) rated themselves as somewhat above average, or able to discuss professional topics.

TABLE 14

Level of Oral Proficiency (Self-Rated)

Level of Ability	Number of Respondents	Percentage of Respondents
Basic or limited	0	0
Speak about social topics	26	35
Speak about professional topics	40	53
Fluent or completely fluent	9	12
TOTAL	75	100

However, a comparison of observed and expected values for this variable compared with time spent in a French milieu raises some questions about these ratings. As indicated in Table 15, there is a higher number of respondents than would be expected at the 'social talk level'. This result may be attributable to the fact that there are no respondents indicating a basic level of speaking ability.

There is also a wide difference between expected and observed values for the category of 15 months or more in a French milieu. It may be that respondents who have spent extended periods of time in a French milieu rate themselves somewhat high.

It is also interesting to note that fluency or near complete fluency appears to be unattainable during short séjours in the milieu for Newfoundland teachers. This finding

is not unexpected given the Newfoundland secondary school Core French program and the unilingual anglophone background of the sample population.

TABLE 15

Level of Oral Proficiency Compared With Time Spent in a French Milieu

Ability Level	Number of Respondents					Total Number of Respondents
	0-2 Months	3-4 Months	5-9 Months	10-14 Months	15+ Months	
Basic level	0	0	0	0	0	0
Social talk level	4(3)*	9(5)	9(7)	3(2)	1(8)	26
Professional talk level	5(5)	5(8)	11(11)	4(4)	15(14)	40
Fluent or near complete fluency	0(1)	0(2)	1(3)	0(1)	8(4)	9
TOTAL	9	14	21	7	24	75

*The numbers in brackets are the expected values

As indicated in Table 16, when comparisons are made between the level of speaking ability and the speaking extent in the French milieu, the differences between observed and expected values appears to suggest that those respondents at the extreme upper and lower ends of the continuum may have rated themselves somewhat high.

However, the findings indicate that 88 percent of the subjects surveyed have spent at least three months in a French

milieu. These subjects have indicated that they have taken a high number of French language university courses both at home and in the milieu. Therefore, it is plausible that the time and energy expended in the milieu has resulted in higher than normally expected levels of achievement. Respondents have also indicated an average to high level of oral/aural skills and seem to have made considerable effort to speak French in the milieu. Teachers who have less than 6 years teaching experience

TABLE 16

Level of Speaking Ability Compared With Speaking Extent in the Milieu

Speaking Extent	Level of Speaking Ability			Total Number of Respondents
	Social Talk	Professional Talk	Fluent	
Occasional social situations	6(2)*	1(4)	0(1)	7
Consistent social situations	10(11)	21(17)	1(4)	32
Consistent in all situations	9(12)	18(19)	8(4)	35
TOTAL	25	40	9	74

*The numbers in brackets are the expected values

may have been motivated by the emphasis on oral/aural skills introduced in the curriculum recently. The most experienced teachers may not have been influenced as much by these developments. The most experienced teachers have fewer university courses completed and have spent less time in the milieu, as indicated previously. Therefore, given the above information, it may be reasonable that there was a higher number than expected of social talk level speakers who spoke only occasionally in social situations. In addition, there could be a higher number than expected of respondents who spoke consistently in all situations and consider themselves to be fluent.

Table 17 indicates that there are no great differences between observed and expected counts at any level of speaking ability for respondents when oral proficiency is compared with the number of university courses taken. This finding in addition to previous findings suggests that the more university courses completed the higher the level of speaking ability. Such a finding is not unexpected.

TABLE 17

Level of Speaking Ability Compared With University French Courses

Ability Level	Respondents					Total Number of Respondents
	1-9	10-12	13-18	19-24	25+	
Basic	0	0	0	0	0	0
Social talk	3(2)*	8(5)	9(10)	6(5)	0(4)	26
Professional talk	2(3)	6(8)	18(16)	6(7)	8(6)	40
Fluent or near complete fluency	1(0)	0(2)	2(4)	2(2)	4(1)	9
TOTAL	6	14	29	14	12	75

*The numbers in brackets are the expected values

There is a further interesting point with regard to the gender of respondents. Differences between the observed and expected values as reported in Table 18 may indicate that females have a somewhat higher level of speaking ability than males. This finding may be related to the larger number of females who are in the 1 - 5 years teaching experience range. These teachers may be recent graduates who have spent more time in a French milieu.

TABLE 18

Level of Oral Proficiency Compared With Gender

Ability Level	Number of Respondents		
	Male	Female	Total
Basic level	0	0	0
Social talk	20(15)*	6(11)	26
Professional talk	18(23)	22(17)	40
Fluent or complete fluency	5(5)	4(4)	9
TOTAL	43	32	75

*The numbers in brackets are the expected values

'Can do' statements.

The questionnaire also asked subjects to rate themselves on some 'can do' statements about their speaking ability in French. Table 19 shows the results as indicated by the respondents. Figure 4 shows a histogram comparing the distribution of the level of speaking ability of respondents for the five 'can do' statements. It is interesting to note that the distribution for the statements which are the two most difficult to perform are nearly symmetrical and approximate the normal distribution while the remaining three distributions are positively skewed. This distribution supports the findings in the self ratings of speaking ability as most respondents rated themselves as 'average' to 'high'. What is both interesting and important is that all five of the 'can do' statements, when

crosstabulated with level of speaking ability and speaking extent in the French milieu, are significant at the .05 level. These data are reported in Table 20. This finding suggests that speaking the target language in the milieu is a productive exercise.

TABLE 19

Level of Proficiency for the 'Can Do' Students

Description of Statements	Level of Proficiency				Total Number of Respondents
	Quite Easily	Some Difficulty	Great Difficulty	Not at All	
Talk about favorite hobby	63	12	0	0	75
Talk about future plans	67	8	0	0	75
Argue a case	41	33	1	0	75
Formal presentation	17	36	22	0	75
Argue a controversial topic	12	42	18	3	75

TABLE 20

Significance Levels for 'Can Do' Statements Compared With Speaking Extent and Levels of Oral Proficiency

Description of Statements	Extent	Level
A. Talk about favorite hobby	.02662	.00053
B. Talk about future plans	.00196	.00392
C. Argue a case	.00021	.00027
D. Formal presentation	.00047	.00257
E. Argue a controversial topic	.02506	.00041

Aural Proficiency (Self-rated)

Listening comprehension is believed to be developed prior to acquisition of oral skills, according to Winitz (1981) and Asher (1969). As listening comprehension is believed to be an integral part of oral proficiency, subjects were asked to rate their own level of listening comprehension. The results are reported in Table 21.

The level of aural proficiency of the respondents appears to be not quite as high as their speaking ability, as more respondents rate themselves in the categories expressing difficulty, and fewer in the category 'quite easily' for most statements. However, as reported in Table 22, when these results are cross-tabulated with the two factors, level of speaking French in the milieu and level of speaking ability, all are significant at the .05 level or below.

Figure 5 shows a histogram comparing the distributions of listening performance of interviewers for four listening situations. It is noteworthy that the distribution for Figure 5 is similar to that of Figure 4. This finding is consistent with the self ratings of level of speaking ability and would be expected since the literature indicates that listening comprehension and oral performance are correlated,

TABLE 21

Aural Proficiency (Self-rated)

Descriptions of Situations	Number of Respondents				Total Number of Respondents
	Quite Easily	Some Difficulty	Great Difficulty	Not at All	
Understand native speakers	35	32	8	0	75
Understand telephone	13	43	18	1	75
Understand rapid speech	13	38	24	1	75
Understand movies	13	45	17	1	75

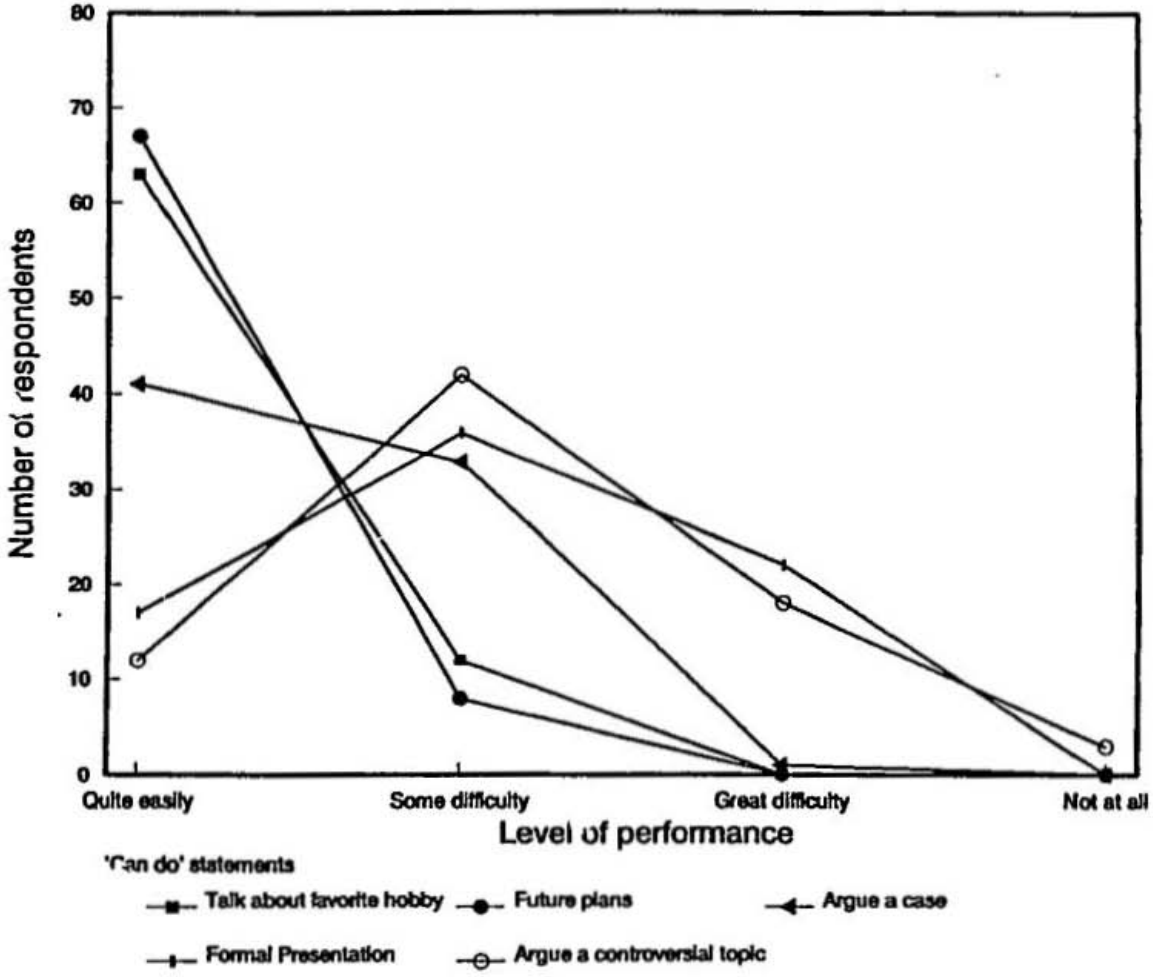

Figure 4: Distribution of oral performance level of teacher/interviewers for five 'can do' statements.

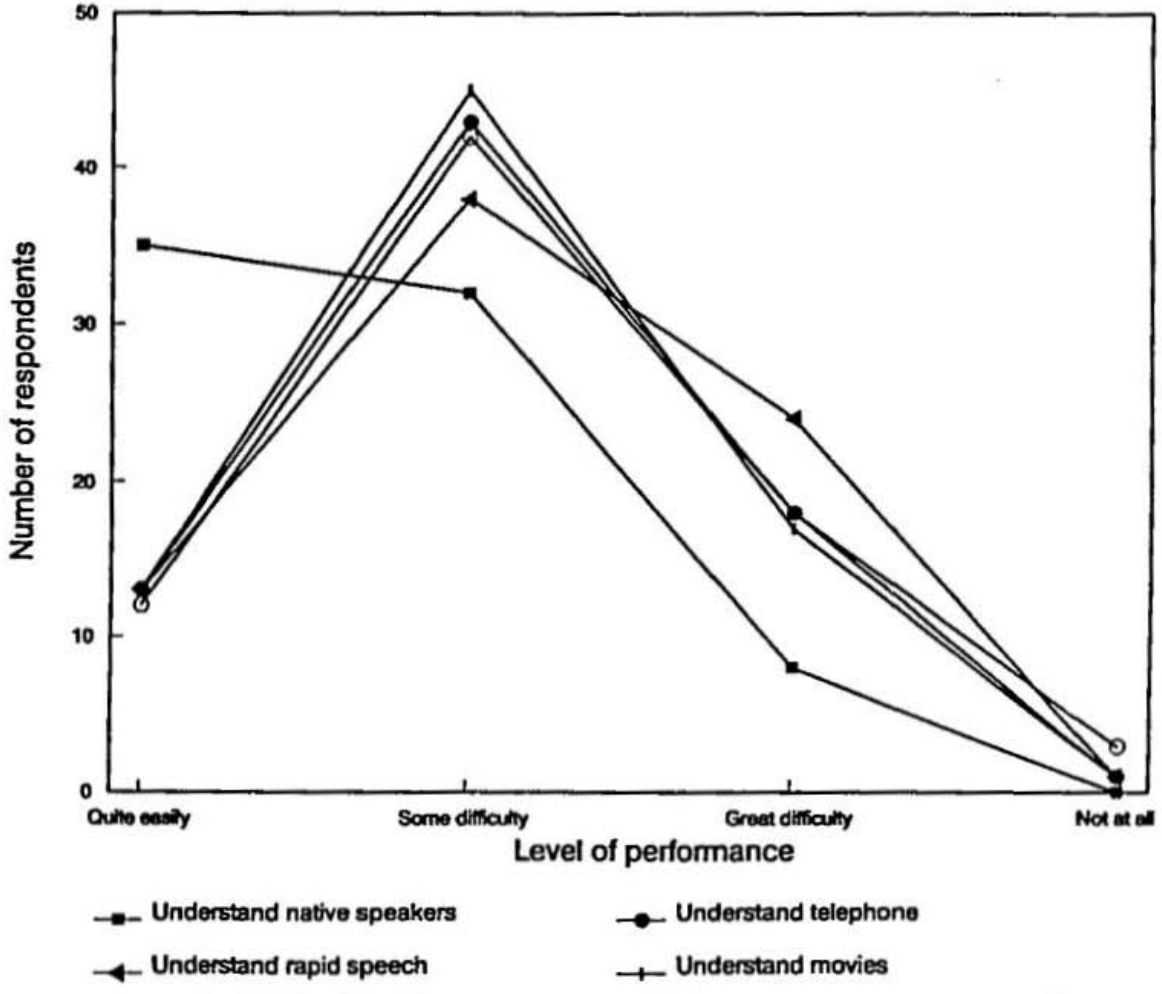

Figure 5: Distribution of aural performance level of teacher/interviewers for listening situations

TABLE 22

Significance Levels for Aural Proficiency Compared With Speaking Extent and Levels of Oral Proficiency

Description of Statements	Significance Levels	
	Extent	Level
Understand face to face conversations	.00061	.00037
Telephone Conversations	.0002	.0000
Understand two	.00061	.0000
Understand movie	.02472	.0000

Profile of a Typical Respondent

An analysis of the data collected in the first section of the questionnaire permits the construction of a profile of the typical respondent. The typical interviewer is a male with 11 to 15 years of teaching experience who is teaching in a rural setting. He has completed 16 to 17 university French courses, and therefore has completed 4 to 5 courses beyond what is required for an academic major in the French language. The typical respondent has spent two to three semesters (5 to 9 months) in a French milieu. Complementing the time in the milieu the respondent has spoken the target language consistently in social situations while in the milieu. However, he rarely speaks French in the home, and has only occasional conversations in French with friends or acquaintances. He rates himself as having an above average speaking proficiency in French. He is able to talk quite well about favourite hobbies

and future plans but he has some difficulty in arguing a case, giving a formal presentation or arguing a controversial topic. The respondent also rates himself as being able to understand native speakers in a face-to face conversation with considerable ease, but as having some difficulty in comprehending movies, conversations on the telephone and rapid speech between two native speakers.

An analysis of the data also indicated that the dispersement for various background characteristics of the typical teacher/interviewer was more widespread than anticipated. While wide variations exist in such categories as number of years teaching experience, number of courses taken, and time spent in a French milieu, there are a very small number of respondents in the categories at either extreme of the continuum. The distribution of responses generally followed a normal pattern with the exception of time spent in a French milieu.

Respondents generally rated themselves as having 'average' to 'high' oral proficiency in French, but somewhat lower aural proficiency. It may be that respondents over-rated somewhat their speaking ability, particularly those at the lower end of the ability level, since the respondents also indicated that they rarely spoke French to friends or acquaintances or in the home.

Other General Information

Table 23 indicates the results of the questions asked of subjects concerning their awareness of the goals of the program as a result of the oral proficiency guideline outlined in French 3200 Oral Testing: A Manual for Interviewers (1986). Results of a second question asking subjects to indicate whether a change in teaching strategies or methodology had occurred as a result of the oral interview are also indicated in the same table. It appears that the guidelines may have ameliorated or enhanced the awareness of the program's goals and have already begun to promote change in teaching strategies.

TABLE 23

Awareness of Goals and Methodological Change

Awareness/Change	Yes	No	No Response	Total
More aware of goals	67	6	2	75
Change in teaching strategy or methodology	9	23	3	75

Subjects were also asked to indicate how comfortable they were with the proficiency scale employed in the French 3200 Oral Interview testing procedure. As may be seen in Table 24, 47 subjects or 63 percent indicated they were comfortable with the scale. Only one subject indicated any displeasure with the scale employed. Such positive reaction may indicate to some measure the validity of the testing procedure.

TABLE 24

Respondents Familiarity with Proficiency Scale

Rating	Number of Respondents	Percentage of Respondents
Very comfortable	47	63
Comfortable	27	36
Not at all comfortable	1	1
TOTAL	75	100

Another question asked subjects to indicate which aids were employed to help determine a level for interviewees. Table 25 shows that 59 percent of the subjects indicated that they employ a combination of scales provided in the Manual for Interviewers and ideas gleaned from training sessions. It may be that teachers employ not only ideas provided in the manual but also an intuitive sense presumably gained from experience in conducting interviews.

TABLE 25

Aids Employed in Determining Proficiency Level

Aids	Number of Respondents	Percentage of Respondents
Levels as outlined in Manual	19	25
Chart as outlined in Manual	7	9
Information gleamed from Manual and training sessions	44	59
TOTAL	70	93

The final question required subjects to indicate other factors which may influence the rating of an interview. All of the factors which were suggested by subjects are listed in Table B1 in Appendix B. The diversity of factors ranged from 'choice of topics for discussion' to 'getting back into the groove of interviewing'. The more common examples of factors included: 'student/teacher fatigue', 'number of interviews per day', and the 'physical setting, i.e. distractions and interruptions during interview'. The relative importance of the suggested variables may be the foundation of a future study centering on the administration of the oral proficiency interview.

CHAPTER FIVE

Analysis of Variance

The following discussion and description is based on the results of a statistical analysis of variance for the independent variables from Part One of the questionnaire and the sets of dependent variables from Part Two.

Table 26 indicates that there were, from the wide range of possibilities for all three dependent variables, relatively few values which were significant. The relationships which were found to be significant are given in Tables 27, 28 and 29.

As may be seen from the tables, the independent variables which did show a significant relationship with the dependent variables were limited in number. The most important were oral proficiency and aural proficiency. Other variables which showed a significant relationship were: speaking extent in the milieu, with friends and acquaintances and in the home; time spent in a French milieu; years of teaching experience; gender; and location.

TABLE 26

Actual Numbers of Significant Values at or Below .05

Dependent Variables	Number of Samples	Number of Traits	Number Possible	Actual Number	Percentage %
Proficiency Factors	5	12	60	2	3.3
Proficiency Levels	20	12	240	21	8.7
Acceptability of Errors	11	12	132	16	12.1
TOTAL	36	36	432	39	9.02

TABLE 27

Independent Variables Showing Significant Relationships With Items Measuring the Importance of Speaking Proficiency Factors

Independent Variable	Proficiency Item	Value
Oral proficiency (a)	Vocabulary	.0280
Aural proficiency (c)	Grammar	.0388

TABLE 28

Independent Variables Showing Significant Relationships With Items Measuring the Acceptability of Errors

Independent Variable	Errors	Significance Values
Oral proficiency (a)	Error 1	.0052
Oral proficiency (c)	Error 1	.0049
Oral proficiency (c)	Error 2	.0367
Oral proficiency (c)	Error 4	.0237
Oral proficiency (d)	Error 6	.0366
Oral proficiency (e)	Error 9	.0307
Oral proficiency (c)	Error 11	.0227
Aural proficiency (d)	Error 11	.0068
Aural proficiency (a)	Error 1	.0159
Aural proficiency (b)	Error 1	.0400
Aural proficiency (d)	Error 2	.0091
Aural proficiency (a)	Error 4	.0216
Aural proficiency (a)	Error 4	.0216
Aural proficiency (d)	Error 8	.0269
Aural proficiency (c)	Error 9	.0045
Aural proficiency (c)	Error 9	.0045

TABLE 29

Independent Variables Showing Significant Relationships With Items Measuring Level of Oral Proficiency

Independent Variable	Proficiency Item	Value
Oral proficiency (c)	Pronoun 2	.0404
Oral proficiency (e)	Fluency 2	.0007
Oral proficiency (d)	Fluency 2	.0321
Oral proficiency (e)	Fluency 3	.0244
Oral proficiency (d)	Grammar combined	.0216
Oral proficiency (d)	Grammar 4	.0031
Oral proficiency (e)	Vocabulary combined	.0369
Oral proficiency (a)	Vocabulary 2	.0031
Gender	Vocabulary 2	.0143
Gender	Vocabulary 3	.0294
Years teaching experience	Vocabulary combined	.0395
Years teaching experience	Vocabulary 4	.0026
Speaking extent in the home	Pronoun 2	.0064
Speaking extent in the home	Vocabulary 1	.0010
Speaking extent with friends	Fluency combined	.0048
Speaking extent with friends	Fluency 2	.0148
Speaking extent with friends	Vocabulary 3	.0147
Speaking extent with friends	Grammar combined	.0135
Speaking extent in milieu	Grammar 1	.0406
Months in milieu	Vocabulary 1	.0243
Location	Grammar 4	.0146

Ranking of Speaking Proficiency Factors and Independent Variables

The first question in Part Two of the questionnaire asked subjects to rank the five factors of speaking proficiency as identified by Harris (1969): pronunciation, grammar, vocabulary, fluency, and comprehension in order of importance with one (1) being the most important and five (5) being the least important.

Table 30 indicates that there were significant differences in the mean rankings between groups for the proficiency factor 'grammar' based on the subjects' self rating of aural skill performance as described in Question 11c of Part One. Those respondents who rated themselves as having the greatest facility in performing the aural skill 'understanding a native speaker on the telephone' placed more importance on grammar than those who had 'some difficulty' or 'great difficulty' in performing the aural skill as described. This finding is consistent with the findings of the Higgs and Clifford (1982) study which indicated that grammar is a more important component of oral proficiency for those speakers who have a higher level of language proficiency.

There was also a significant value result for the proficiency factor 'vocabulary' based on the respondents' self rating of the performance of the oral skill described in Question 10a of Part One. Those teachers who had considerable oral fluency as determined by the self rating 'quite easily' placed less importance on vocabulary than did those teachers

who were less fluent. This finding is also consistent with the results of the Higgs and Clifford profile of oral proficency factors.

Results of the study generally appear to suggest that the level of oral and aural proficiency could create significant differences in ratings in an oral interview. Vocabulary appears to be a more important factor for less proficient interviewers, while grammar appears to be a more important factor for those whose proficiency level is higher. It may be that grammar is underrated by those with more years of teaching experience, given the findings in Table 31. This may be because those teachers do not necessarily fall into the most fluent group.

TABLE 30

Ranking of Speaking Proficiency Factors
Compared With Oral and Aural Proficiency

Proficiency Factor	Level of Oral/Aural Proficiency			Significance Values
	Quite Easily	Some Difficulty	Great Difficulty	
'Vocabulary'	3.04	2.25	-	.0280
'Grammar'	3.00	3.89	3.83	.0388

Table 31 indicates the results of the mean rating of the factors in order of importance for the independent variable years of teaching experience. Included in Table 31, at the bottom, is the overall mean ranked rating in order of importance for all cases in the study. While there were no significant differences in these data, it is interesting to note the pattern of results.

Table 31 shows that subjects overall ranked the proficiency factor, comprehension, as most important, followed by fluency, vocabulary, pronunciation and grammar. However, comprehension was the only factor for which the ranking was agreed upon by all groups. This finding may be due to the idea, expounded in Chapter 1, that the oral competence of an individual is not unitary, and that each language factor is not necessarily developed to the same extent as another. Figure 6 shows the distribution of the ranking of the factors in speaking proficiency by years of experience. It may be noted that as the number of years of experience increases, the uniformity by group of the ranking of the importance of the factors of speaking proficiency also increases. It may be that the less experienced teachers are less structured or perhaps more eclectic in promoting oral proficiency.

TABLE 31

Ranking of Speaking Proficiency Factors
Compared With Years of Teaching Experience

Years Teaching Experience	Rankings in Order of Importance				
	Compre-hension	Fluency	Vocabulary	Pronunciation	Grammar
1 - 2	1	4	2	5	3
3 - 5	1	2	2	3	3
6 - 10	1	2	4	3	5
11 - 15	1	2	3	4	5
16+	1	2	3	4	5
Ranking of all cases	1	2	3	4	5

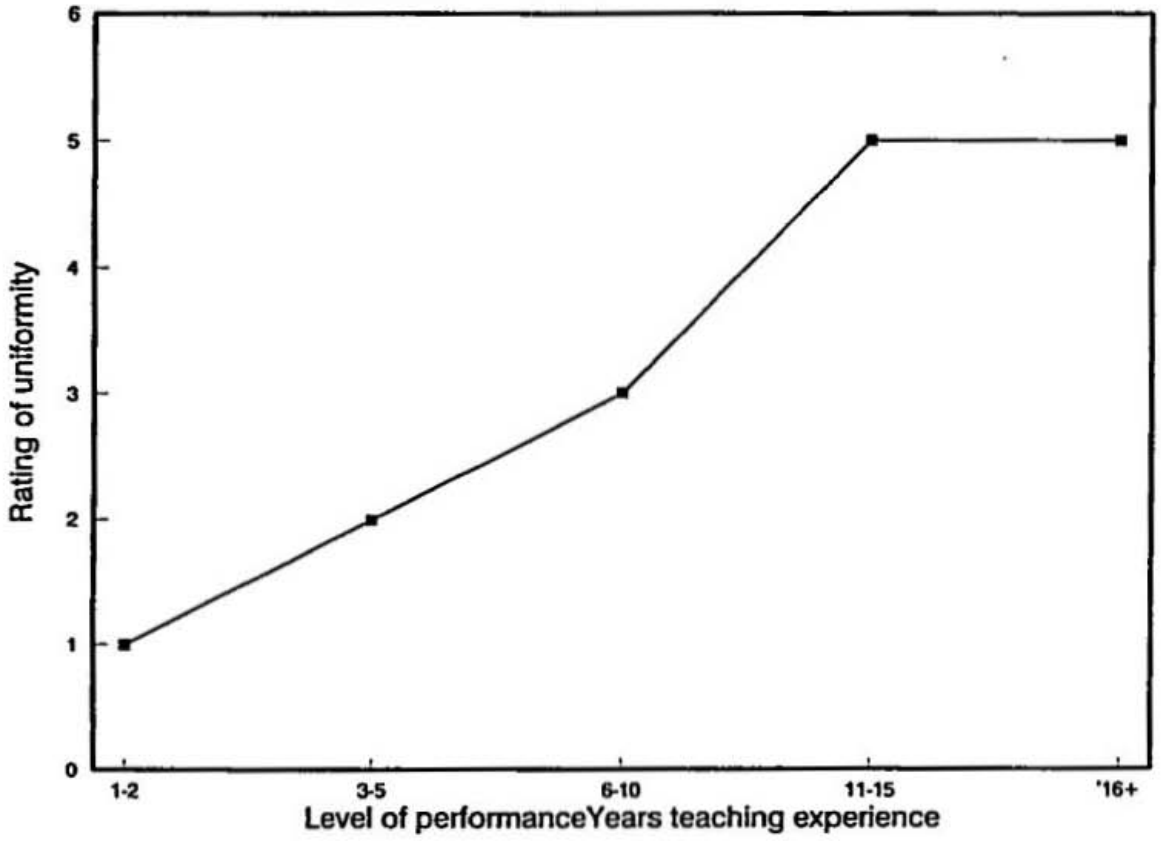

Figure 6: Distribution of the rating of the factors in speaking proficiency for years experience

Acceptability of Errors and Independent Variables

The acceptability of errors question had 11 examples of typical errors selected from the guidelines for the rating of proficiency levels outlined in the booklet, French 3200 Oral Testing: A Manual for Interviewers. Subjects were required to rate the acceptability of selected errors on a scale of 1 to 4, with one (1) being 'not acceptable' and four (4) being 'acceptable'. The mean ratings for the acceptability of errors, as given in Tables 32 and 33, indicate that significant relationships in this area were related only to the levels of oral and aural proficiency of the respondents as measured by Questions 10a to 10e and 11a to 11d in Part Two of the questionnaire. Questions 10a to 10e are five 'can do' statements about the respondents' speaking ability in French. Questions 11a to 11d are four "listening situations" about the respondents' second language listening ability. Tables 32 and 33 indicate a basic linear relationship common to all examples: the greater the interviewers' facility to perform the oral/aural skill, the higher the rating of acceptance of errors. Those respondents who had a higher level of proficiency were more tolerant of the error than those who possessed lower levels of proficiency (oral and aural).

TABLE 32

Mean Rankings for Acceptability of Errors
Compared With Oral Proficiency (Self-Rated)

Error	Oral Proficiency	Self-Rated Levels for Oral Skill				Significance Value
		Quite Easily	Some Difficulty	Great Difficulty	Not at All	
Error 1	Oral Proficiency (a)	2.00	1.63	-	-	.0052
Error 1	Oral Proficiency (c)	*2.05	1.84	*1.00	-	.0049
Error 2	Oral Proficiency (c)	*2.27	*2.03	2.00	-	.0367
Error 4	Oral Proficiency (c)	*2.00	1.81	*1.00	-	.0237
Error 6	Oral Proficiency (d)	*2.05	1.68	*1.61	-	.0366
Error 9	Oral Proficiency (e)	*2.08	1.88	*1.56	2.00	.0307
Error 11	Oral Proficiency (c)	*2.25	2.06	*1.00	-	.0227
Error 11	Oral Proficiency (d)	2.25	*2.28	*1.85	-	.0068

*denotes the pairs of mean ratings significantly different.

TABLE 33

Mean Rankings for Acceptability of Errors
Compared With Aural Proficiency (Self-Rated)

Error	Aural Proficiency	Self-Rated Levels for Aural Skill				Significance Value
		Quite Easily	Some Difficulty	Great Difficulty	Not at All	
Error 1	Aural Proficiency (a)	*2.05	1.90	*1.62	-	.0159
Error 1	Aural Proficiency (b)	2.07	*2.00	*1.70	2.00	.0400
Error 2	Aural Proficiency (d)	2.07	*2.27	*1.94	-	.0091
Error 4	Aural Proficiency (a)	*1.97	1.93	*1.51	-	.0216
Error 4	Aural Proficiency (a)	*1.97	1.93	*1.51	-	.0216
Error 8	Aural Proficiency (d)	2.00	*2.27	*2.00	-	.0269
Error 9	Aural Proficiency (c)	*2.23	*1.72	1.83	-	.0045
Error 9	Aural Proficiency (c)	*2.23	1.72	*1.83	-	.0045

*denotes the pairs of mean ratings significantly different.

Ratings of Levels of Oral Proficiency and Independent Variables

Respondents' ratings of the five levels of proficiency as outlined in the manual French 3200 Oral Testing: A Manual for Interviewers (1986) was another of the three major themes investigated in Part Two of the questionnaire. The oral proficiency level rating question had four main categories: pronunciation, grammar, vocabulary and fluency. Each of the four main categories had four discrete examples describing language characteristics. Respondents were asked to rate the examples in accordance with the proficiency levels outlined in the manual. Mean values for each of the four examples were calculated. Thus, pronunciation had four examples referred to as pronunciation 1, 2, 3, and 4. In addition, an overall combined mean result for a category, called, for example, 'pronunciation combined', was produced. The following is a description of the results of the analysis of variance of the four items for each of the four categories and their relationship to the various independent variables where this relationship was shown to be significant.

Location.

Table 34 indicates that there was a significant difference between the mean rankings of the item 'grammar 4' in Question 3 of Part Two of the questionnaire based on location of the respondents. The urban setting teachers rated the item 'grammar 4' higher than the teachers in rural settings. As the results showed, no other examples with significant differences were

produced to support any trends for the variable location. As was reported above, E.T.S. (1985) indicated that location was not an important factor in self-rating oral proficiency. Given the unfolding of the findings of this study, results seem to be related to oral proficiency of the respondents. Therefore, similarly to the results of the E.T.S. study, as were discussed above, the variable location may not be an important factor.

TABLE 34

Mean Rankings for Proficiency Levels (Grammar) Compared With Location

Proficiency Item	Location		Significance Value
	Rural	Urban	
'Grammar 4'	2.42	3.05	.0146

Gender.

Table 35 indicates that there were differences in the mean ratings for the proficiency items, 'vocabulary 2' and 'vocabulary 3', based on gender. Given that there was evidence in Part One of the study to indicate that females tended to have a higher level of oral proficiency than males, this result may be associated with oral skills more directly than gender. There were no further data to support any trends for the variable gender. It is to be noted that the significant differences occurred with items measuring the factor vocabulary.

TABLE 35

Mean Rankings for Proficiency Level (Vocabulary) Compared With Gender

Proficiency Item	Gender		Significance Value
	Male	Female	
'Vocabulary 2'	3.51	4.00	.0143
'Vocabulary 3'	1.90	1.53	.0294

Time spent in a French milieu.

Table 36 indicates that there were significant differences between groups of subjects having spent varying amounts of time in a French milieu. Although there is no consistent pattern or model, significant differences did occur between those who spent more than two years in a French milieu and some other groups. This finding seems to indicate that there could be significant differences in assessment of proficiency levels based on the oral skills of the interviewer, as those who have spent more time in the French milieu tend to be those who have the highest levels of proficiency in French.

It is to be noted that the differences occurred with items measuring vocabulary and that the most proficient respondents gave less emphasis to vocabulary.

TABLE 36

Mean Rankings for Proficiency Levels (Vocabulary) Compared With Time Spent in a French Milieu

Proficiency Item	Months Spent in a French Milieu						Significance Value
	0-2	3-4	5-9	10-14	15-24	25+	
'Vocabulary 1'	3.11	*3.57	3.42	3.14	3.55	*2.50	.0243
'Vocabulary 1'	3.11	3.57	*3.42	3.14	3.55	*2.50	.0243
'Vocabulary 1'	3.11	3.57	3.42	3.14	*3.55	*2.50	.0243

*denotes pairs of means signficantly different.

Years of teaching experience.

Significant differences resulted between groups of subjects with varying amounts of years teaching experience and the items sampling an interviewee's use of vocabulary. These examples are found in Question 3 in Part Two of the survey. Table 37 shows various pairs of groups which have significant differences between mean scores. In addition to the 'vocabulary 4' item, totals of mean scores for each example of 'vocabulary 1' to 'vocabulary 4' were computed and are indicated as 'vocabulary combined'. The difference in the mean ratings between those respondents with 11 - 15 years of teaching experience and the most experienced group for the item 'vocabulary 4' appears to be associated with oral performance, as the lesser experienced group are probably more fluent than the most experienced group. This finding would then also be consistent with the indications that vocabulary is a less important factor for the more fluent speakers.

Significant differences also resulted between mean ratings for groups of respondents with 3 - 5 years of teaching experience and those with 11 - 15 years for 'vocabulary 4' and 'vocabulary combined'. It appears that those with fewer years teaching experience tend to rate vocabulary items higher than those with 11 - 15 years experience.

TABLE 37

Mean Rankings for Proficiency Levels (Vocabulary) Compared With Years of Teaching Experience

Proficiency Item	Years Teaching Experience					Significance Values
	1-2	3-5	6-10	11-15	15+	
'Vocabulary 4'	2.57	* 3.20	2.70	* 2.26	3.15	.0026
'Vocabulary 4'	2.57	3.20	2.70	* 2.26	*3.15	.0026
Vocabulary Combined	11.57	*12.20	11.79	*10.31	11.76	.0395

*denotes pairs of means signficantly different

Speaking extent in the milieu with friends or acquaintances and in the home.

Tables 38, 39, and 40 all indicate that there were significant differences in pairs of mean rankings of selected items for the proficiency level ratings related to respondents' varying degrees of consistency in speaking French in the milieu, with friends and acquaintances outside the milieu and in the home. It can be seen that those respondents who spoke French only occasionally in the milieu tend to rate grammar differently from those who spoke French consistently in all situations. Those respondents who never or rarely spoke French with friends or acquaintances tended to rate fluency, grammar and some aspects of vocabulary differently from those who spoke occasionally or regularly with friends and acquaintances. Lastly, there were some specific differences which were rated differently by those who spoke French half-time or more in the home. These patterns would seem to suggest that some significant differences in the ratings of oral interviews could be related to the speaking proficiency of the teacher inteviewer. These differences would fall primarily into the categories of vocabulary, grammar, and fluency. These differences are consistent with the previously developed pattern of the more fluent speaker rating grammar items as more important than vocabulary.

TABLE 38

Mean Rankings for Proficiency Levels (Grammar) Compared With Speaking Extent in the Milieu

Proficiency Item	Speaking Extent in the Milieu			Significance Value
	Occasional in social situations	Consistent in social situations	Consistent in all situations	
'Combined Grammar'	*2.57	*1.78	2.02	.0406

*denotes pairs of means significantly different

Oral Proficiency (Self-Rated)

Table 41 indicates that there was a significant difference between a pair of mean rankings for the dependent variable 'grammar 4'. This differences was based on the degree of self-rated oral proficiency of the respondents. A basic linear relationship may be perceived, that is, the better the skill level, the higher the mean rating given to grammar. This finding is consistent with previous findings and with the results of the Higgs and Clifford study.

Table 41 also shows significant differences between pairs of mean ratings for the proficiency level items 'fluency 2' and 'fluency 3' for the category 'not at all' as compared to the other three mean rankings for each dependent variable. The 'not at all' group, however, was less than the arbitrary minimal level of seven cases, or 9 percent, of the total number of respondents. Therefore, this finding may not be informative. Without the group 'not at all' there are no important

TABLE 39

Mean Rankings for Proficiency Levels (Vocabulary, Grammar, Fluency) Compared With Speaking Extent With Friends and Acquaintances

Proficiency Item	Speaking Extent With Friends and Acquaintances			Significance Value
	Never or rarely	Occasionally	Regularly	
'Fluency 2'	*2.16	* 3.03	2.70	.0148
'Vocabulary 3'	1.16	* 1.86	* 1.33	.0147
'Combined Grammar'	*8.33 *8.33	*10.06 10.06	11.30 *11.30	.0135 .0135
'Combined Fluency'	*8.50 *8.50	*11.27 11.27	10.80 *10.80	.0048 .0048

*denotes pairs of means significantly different

TABLE 40

Mean Rankings for Proficiency Levels (Vocabulary, Pronouns) Compared With Speaking Extent in the Home

Proficiency Item	Speaking Extent in the Home			Significance Value
	Never or Rarely	Occasionally	Half Time or more	
'Pronoun 2'	3.12	*2.71	*3.75	.0064
'Vocabulary 1'	*3.56	2.95	*3.25	.0010

*denotes pairs of means significantly different.

TABLE 41

Mean Rankings for Proficiency Levels (Vocabulary, Grammar, Fluency, Pronunciation) Compared With Oral Proficiency (Self-Rated)

Independent Variable	Proficiency Level Items	Self Rated Levels for Oral Skill				Significance Value
		Quite Easily	Some Difficulty	Great Difficulty	Not at All	
Oral proficiency (c)	'Pronunciation 2'	* 2.87	* 3.21	4.00	-	.0404
Oral proficiency (d)	'Grammar 4'	* 3.17	* 2.25	2.61	-	.0031
Oral proficiency (a)	'Vocabulary 2'	* 3.61	* 4.25	-	-	.0192
Oral proficiency (e)	'Fluency 2'	3.08	3.04	*2.76	*1.33	.0007
		3.08	* 3.04	2.76	*1.33	
		* 3.08	3.04	2.76	*1.33	
Oral proficiency (e)	'Fluency 3'	2.25	2.50	*2.62	*1.33	.0244
Oral proficiency (e)	'Vocabulary Combined'	*11.58	*11.50	11.54	-	.0369

*denotes pairs of means significantly different

differences between the mean rankings of the groups for this item. Table 41 also shows a significant difference between pairs of mean ratings for the proficiency item 'vocabulary 2' for the groups 'quite easily' and 'some difficulty ' based on the respondents self-rated oral proficiency. A basic linear relationship may be perceived: the lower the self-reported skill level, the higher the mean rating given to vocabulary. This finding is consistent with previous findings and with the results of the Higgs and Clifford (1982) study.

Aural Proficiency (self-rated).

The results for the proficiency level variables, 'fluency 2' and 'grammar combined', with the aural proficiency (self-rated) of the respondents, as indicated in Table 42, showed that those respondents who possessed a high level of aural ability rated grammar and some aspects of fluency significantly differently from those whose aural skills were not quite so well developed. It seems that those respondents with a high self-rated level have a greater facility with the examples of 'fluency' and 'grammar combined'.

TABLE 42

Mean Rankings for Proficiency Levels (Grammar, Fluency) Compared With Aural Proficiency (Self-Rated)

Independent Variable	Proficiency Level Items	Self Rated Levels for Aural Skill				Significance Value
		Quite Easily	Some Difficulty	Great Difficulty	Not at All	
Aural proficiency (d)	'Fluency 2'	* 3.30	2.93	*2.58	-	.0321
Aural proficiency (d)	'Grammar Combined'	*11.11	*9.52	10.22	-	.0216

*denotes pairs of means significantly different.

Conclusion

The analysis of variance indicates that there were significant differences for all three categories of dependent variables for a small number of the independent variables.

For the category, factors in speaking proficiency, those respondents with a high level of oral competence rated grammar as more important than did those with a lower level of competence. Those with a lower level of aural proficiency rated vocabulary as a significantly more important factor than did those who had a high level of aural competence.

For the second category, acceptability of errors, there was a significant relationship between the aural and oral proficiency of the respondents and the acceptability of certain errors. The relationship is basically linear: the higher the level of aural and oral competence, the greater the acceptance of error. In this category there was a greater homogeneity of results than in the other categories.

For the third category, ratings of levels of oral proficiency, the independent variables, gender, years of teaching experience, speaking extent in the milieu, with friends and acquaintances, and at home, and the measures of aural and oral proficiency, have significant relationships with the dependent variables. As has been indicated previously, the characteristics of gender and years of teaching experience may be related to competence in French. It could also be demonstrated that speaking extent in the milieu, with friends and acquaintances and in the home would ultimately affect oral

and aural proficiency (E.T.S. 1985). Therefore, it appears that those charcteristics of the respondent which could bring about significant differences in their ratings of student interviews are primarily associated with their level of aural and oral proficiency in French. The differences occur particularly in the categories grammar, vocabulary and fluency.

Given the results of the study, it appears that there are some significant relationships between the characteristics of the interviewers pertaining to their level of fluency in French and their assessment of the oral proficiency of the student in an interview situation. The full impact of these findings is discussed in the following chapter.

CHAPTER SIX

Summary and Discussion

Problem

This thesis is essentially a pilot study to investigate the relationship between the characteristics of teacher interviewers and the reliability of the rating of an oral proficiency test. There are three questions which this study addressed: (1) Are there significant variations in the way interviewers rank in importance the five factors of speaking proficiency (vocabulary, pronunciation, grammar, fluency, comprehension)? (2) Do interviewers' scores vary significantly when rating the acceptability of different types of errors? (3) Do interviewers' scores vary significantly when rating proficiency levels? Furthermore, this pilot study addressed the question as to whether any such significant differences that were found could be attributed to specific background characteristics of the respondents.

Methodology

The data were gathered by the means of a questionnaire. Fifty percent of the total population of the trained teacher interviewers in the island portion of the province of Newfoundland were surveyed by mail. The questionnaire focused on two major types of information. The background characteristics of the teacher interviewers provided the independent variables. The rating of items such as factors of speaking proficiency, certain types of errors, and levels of

proficiency provided the dependent variables. An analysis of variance was performed to determine whether any of the ratings were significantly different based on differences in the background characteristics of the respondents.

Results

Results indicated that there were a small number of significant differences for each of the areas investigated.

Higgs and Clifford (1982) proposed the Hypothesized Relative Contribution Model (RCM) to describe tester perception of the relative roles of five identified factors or subskills of speaking proficiency: vocabulary, grammar, pronunciation, fluency and sociolinguistics. The RCM was based on an opinion poll of 50 foreign language teachers representing 17 languages. Similar to the survey results found in this thesis the RCM results were not formulated on actual interviews but rather on impressions of the importance of proficiency factors described. The Higgs-Clifford study proposed that if oral proficiency is the goal, then the emphasis in teaching and testing on the various language subskills should be representative of their relative importance in oral proficiency. However, the study showed that the relative importance of the factors changed according to the proficiency of the respondents. Respondents at the lower level of oral proficiency (Level 1) ranked vocabulary as most important, while respondents at higher levels of proficiency (Level 3) ranked grammar as most important. Only at

the highest level of proficiency (Level 5) were the various factors ranked equally.

The results of this study were similar to those of the Higgs and Clifford study. While the categories used in this study were vocabulary, grammar, pronunciation, fluency and comprehension, there was a tendency for those respondents with lesser proficiency in French to rate vocabulary as more important than grammar, whereas those with greater proficiency in French tended to give more importance to grammar.

There were significant differences in the rating of some errors. These differences were related to the oral and aural proficiency of the respondents. A basic linear relationship was found which indicated that the higher the level of proficiency of the teacher interviewer, the higher the degree of acceptance of the specific examples of error.

There were a modest number of significant differences found related to some background characteristics of respondents. In general, all the background characteristics which showed significant relationships could be attributed to differences in the French language proficiency of the teacher interviewer. Areas where significant differences appeared were in the assessments of vocabulary, grammar and fluency.

Thus, the analysis of variance indicated that significant differences in the rating of specific examples of aspects of oral proficiency did occur. The results also indicated that these differences were related to the background characteristics of the respondents. The background

characteristics which showed significant relationships with ratings were all related to the target language competence of the respondents.

Further analysis of the data indicated that nine of the independent variables were involved in twelve significant pairs of relationships with each other. These results are indicated in Figure 7. All of the independent variables, except location, were paired at least once, producing a significant value at or below the .05 level.

The variable location (urban or rural), when cross-tabulated with each of the other independent variables, did not produce any pairs with significant values. Therefore, it was concluded that there were no significant differences in the rating of oral proficiency based on the geographical location of respondents. Because of this information as indicated earlier, the one finding of a significant difference for the evaluation of pronouns based on differences in the location of the respondents was not felt to be important.

Some gender differences related to oral proficiency in French also appeared. These gender differences, however, may be explained in terms of the years of teaching experience of the respondents. Since there is a disproportionately higher number of females in the lower categories for number of years of teaching experience, it may be suggested that their more recent experience in university and in a French milieu explains the finding of a higher level of oral proficiency for females.

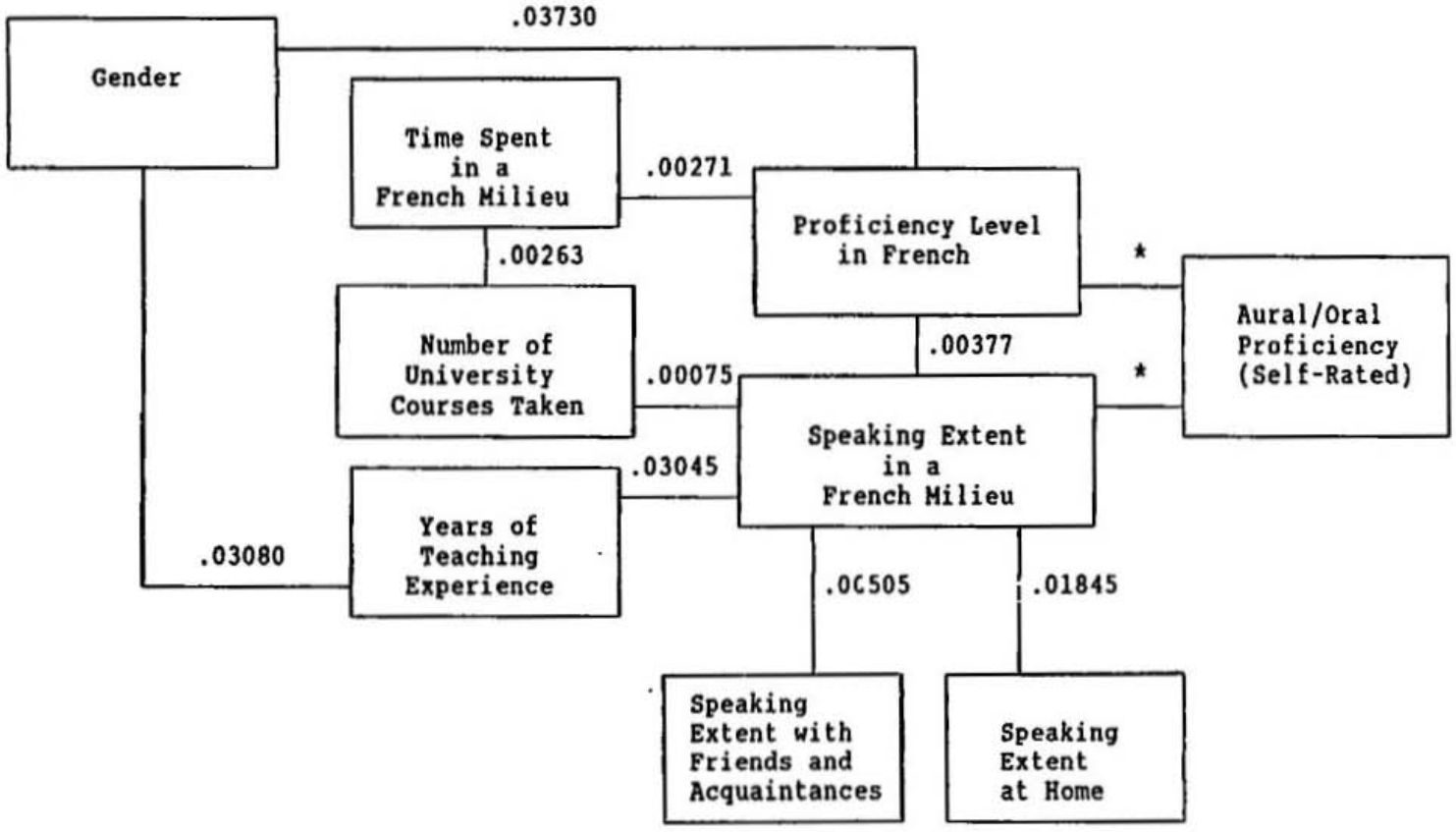

FIGURE 7: Cross-tabulations of pairs of independent variables with significance values below the .05 level.

*See Appendix C for levels of significance for each item of the dependent variables.

It had been reported in the literature (E.T.S. 1985) that there could be a relationship between the extent of the use of French outside the milieu, with friends and acquaintances and at home, and the self-rated proficiency levels. No significant relationship of this type was found in this study. However, the variables speaking extent in the milieu, with friends and acquaintances, and at home, did give an interesting finding for Newfoundland teachers. Results indicated that, while efforts were made by most respondents to speak French consistently in the milieu, few respondents continued to speak French with friends and acquaintances and at home.

A significant relationship was found to exist between the extent of the use of the target language in the milieu and the self-proficiency ratings of aural and oral proficiency. A basic linear relationship was evident between speaking extent in the French milieu and proficiency levels. Consistent use of French in the milieu appears to contribute to proficiency in the target language.

The results of these cross-tabulations further indicate that it is those respondents who spoke French consistently in the French milieu who have regular conversations with friends and acquaintances, and who reported that they spoke French half-time or more at home. These findings would suggest that it is the more fluent teachers who continue to use French when outside the French environment.

Since the E.T. S. study implied that it was only those teachers who used their French outside the milieu who really maintained a high level of proficiency in the language, two questions may be raised. The findings of the study indicated that most respondents rated themselves at an above average level of proficiency. To what extent is this rating an assessment of their current level of proficiency? This assessment may be based on the level of proficiency reached by the respondent when last in a French milieu. A second question may be raised about the quality and quantity of French used in the classroom. To what extent do teachers who do not use French outside the milieu use French inside the classroom, and what is the quality of French used in these communications? In addition, a recent general evaluation of the French 3200 Oral Tests, conducted by the Department of Education on a provincial and at the district level, have indicated, among others, a concern over the use of English during the proficiency interview. Questions may be posed in order to establish the reasons for the apparent use of English during the interview. The answer to all these questions may be of considerable importance in assessing the communication aspects of a given classroom, or of a given interview situation.

The above discussion indicates that, of the background characteristics studied, there are basically five which are of importance in determining the French proficiency of the teacher as indicated in Figure 8. The number of university courses taken, the time spent in a French milieu, and the extent to

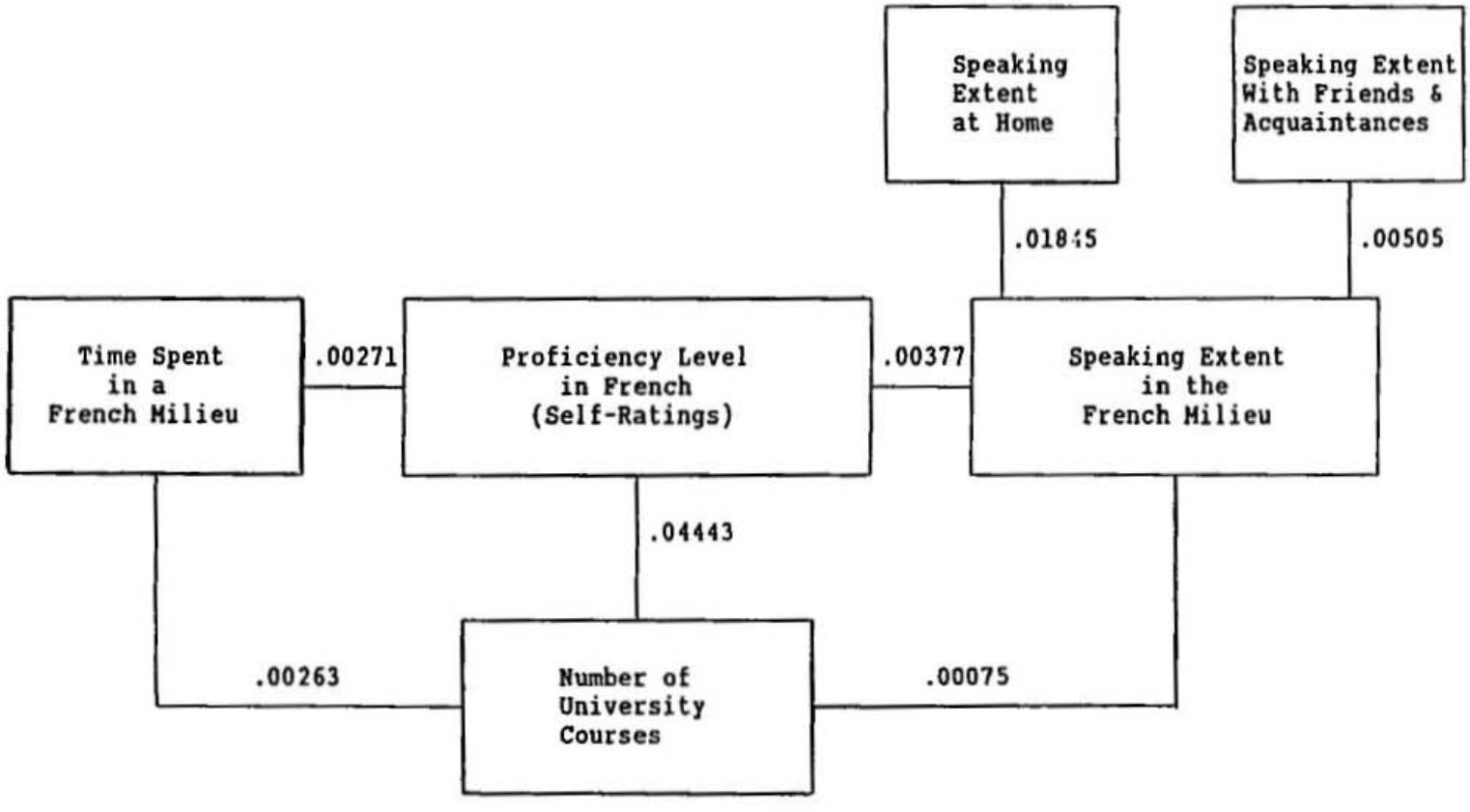

FIGURE 8: Background Characteristics Significantly Related to Proficiency in French

which French is spoken in the milieu appear to have considerable bearing on the proficiency level achieved as determined by the self-ratings of the respondents. There is also a further point to be considered. Speaking extent with friends and acquaintances or at home may have an effect on the level of proficiency maintained outside the milieu. However, this factor may not be reflected in the self-rated oral and aural proficiency of the respondents if there has been a lapse of time since the respondent was in a French milieu. For those respondents who have not recently spent time in a French milieu, and who do not use French regularly outside the milieu (85% of the respondents), the actual level of proficiency in French may be lower than that indicated by the self-ratings.

Conclusion

The conclusions which may be drawn from this study are threefold. They suggest that (1) there were significant differences in the ratings of the oral interview; (2) these differences were related to the oral proficiency of the interviewer; and (3) these differences were in the areas of assessment of vocabulary, grammar, and, to some extent, fluency.

Global ratings verses discrete example ratings.

Results of the mean ratings of the dependent variable proficiency level items indicated significant differences for items pertaining to vocabulary, grammar, and fluency related to the proficiency level of the interviewer. While the data produced some general tendencies in the areas described above, there was no strong relationship appearing regularly in the data for all, or even most, items in the analysis. The number of non-significant relationships was considerably greater than significant ones. Similarly, results indicated that the discrete items for the dependent variables did not produce regular patterns of significant differences which might influence rater reliability in a consistent manner.

The analysis of the data indicated that there were a very limited number of discrete oral proficiency items which were significantly affected by the background characteristics of the respondents. The characteristics of the respondents which were important were unequivocally those associated with the proficiency level of the interviewer. Thus a case might be made that the language proficiency of the interviewer will significantly affect the oral rating given to an interviewee.

However, since the number of significant relationships was relatively small, given the number of possible relationships, and since the relationships tended to be somewhat isolated rather than falling into regular and consistent patterns, it seems appropriate to look at another aspect of this issue.

Communicative competence is the current primary issue in language proficiency. In addition, an overall framework has been developed to describe the characteristics of the concept, and a global rating technique has been developed to evaluate it. Therefore, it may be that the global rating of an interview will give a more accurate evaluation of oral proficiency interviews than the use of a rating based on discrete item evaluation.

The use of a global rating for levels of proficiency rather than discrete item descriptions may lessen the possibility of significant differences between oral ratings of students by teacher/interviewers. The relationship between the independent and dependent variables in this study provide only limited insight into rater reliability. The global rating factor may be a means of overcoming the discrete differences which might bring about differences in rater reliability. The global rating factor was not employed as a variable in this study. An investigation of such scope would involve live interviews with either simultaneous or delayed scoring techniques. As discussed earlier, financial and other constraints rendered such an investigation unfeasible. However, certain general questions asked of the respondents may be useful at this point.

Teacher interviewers may have unknowingly indicated a measure of support for the hypothesis of global proficiency testing and evaluation. Table 31 indicates that in rating the five proficiency factors respondents rated the most globally

viewed factors, comprehension and fluency, as the more important factors. The table also shows that subjects rated the other three discrete type proficiency factors as of less importance.

The discrete items, such as vocabulary, grammar and pronunciation, used as the basis for the dependent variables in this study, play a role in proficiency testing. These items are employed as strategies by the interviewers to guide the interview process and to help establish the proficiency levels. However, according to the anecdotal reports of respondents, few use the discrete item guides as an actual evaluative measure.

Recommendations

Outlined below are suggestions or ideas for further investigations of the issues presented in this pilot study. The first recommendation calls for a further study utilizing the same independent variables but employing live or taped interviews and the global technique of scoring or rating the interview. As suggested above, such a study may prove to yield interesting information about rater reliability.

The second major recommendation for consideration is to discover how teachers employ the discrete points such as those outlined in the proficiency levels. A similar study to this pilot investigation, but using a population chosen at the two extremes of proficiency in French, might yield more definite information on whether oral competence of the interviewers affects consistently discrete items in scoring.

Background characteristics of the respondents produced several interesting issues which may be worthy of further study. In the area of speaking extent, it would be valuable to learn why respondents made such great efforts to speak French in the milieu but did not continue the trend with friends and acquaintances and at home. It is important to determine whether this lack of use outside the milieu affects oral and aural proficiency. Furthermore, is the same minimal effort being made by teachers and students in speaking French in the classroom?

In addition, results from a study of background characteristics might be employed in the development of professional standards for teacher certification.

It also appears that time in the milieu during which academic courses are studied is more effective in achieving higher levels of oral proficiency than just time spent in a milieu. Investigators may wish to study more closely this relationship and make suggestions for the development of a course of action other than a strictly academic route for the development of oral proficiency in the milieu.

Conclusion

Oral proficiency testing is an extremely complicated process. It is difficult to attain both validity and reliability for instruments assessing communicative competence. The present study depicts some aspects of these difficulties. One such example is the dilemma of establishing the degree of influence of the background of the teacher interviewer on the

rating of an oral interview. It may be that this influence cannot be determined solely by rating a variety of discrete items or proficiency factors. While discrete item testing is reliable and valid from a psychometric point of view, it has little validity from the point of view of communicative competence. In order to test oral language proficiency, an instrument must be employed that exhibits actual language use and permits an evaluation of the process. The oral proficiency interview appears to be such an instrument.

Carroll (1979,1961) suggested that a combination of discrete item testing and integrative skill or global testing was needed. The oral proficiency interview allows for the multifaceted communicative approach by the use of a global rating factor. In addition, the oral proficiency interview employs discrete-point items. The literature, (Magnan 1987), indicates that these discrete items may be employed as a secondary factor rating system. This writer maintains that teacher/interviewers employ these discrete point items as strategic items to assist in finding the level of the student during the interview process, but that they do not use them to rate the overall level of proficiency of the student. Therefore, while this study has indicated that there appears to be significant differences in the ratings of discrete items of vocabulary, grammar, and fluency based on the oral proficiency of the interviewer, these differences may not affect significantly the overall global rating given to an interviewee.

However, given the findings of this study, there may be a need for professional standards for second language teachers, or at least for those teachers who will be giving the oral interview. At the present time, the oral interview training sessions sponsored by the Newfoundland and Labrador Department of Education are being used by some school boards as an unofficial qualification for teaching positions at the senior high school level. However, it may be that professional standards based on the current or actual proficiency level of a teacher should be established if reliable scoring of the oral interview for all students is to be attained.

BIBLOGRAPHY

ACTFL Provisional Proficiency Guidelines (1986). Hastings-on-Hudson NY: American Council on the Teaching of Foreign Languages Materials Center.

_____ Provisional Proficiency Guidelines (1982). Hastings-on-Hudson, NY: American Council on the Teaching of Foreign Languages Materials Center.

Adams, Marianne (1978). "Measuring Foreign Language Speaking Proficiency: A Study of Agreement Among Raters." In Direct Testing of Speaking Proficiency: Theory and Application. Ed. John L.D. Clark pp. 131-149. Princeton: Educational Testing Service.

Allen, J.P.B. (1983). "A Three-Level Curriculum Model for Second-Language Education." The Canadian Modern Language Review, 40(1), 23-43.

Asher, J. (1969). "The Total Physical Response Approach to Second Language Learning." The Modern Language Journal, 53(1):3-17.

Bachman, L.F. and Sandra Savignon (1986). "The Evaluation of Communicative Language Proficiency: A Critique of the ACTFL Oral Interview." The Modern Language Journal. 70(4):380-390.

_____ and Adrian S. Palmer (1984). "Some Comments on the Terminology of Language Testing." In Communicative Competence Approaches to Language Proficiency Assessment: Research and Application. Ed. C. Rivera. Clevedon Avon England Multilingual Matters Ltd.

_____ and Adrian S. Palmer (1981). "The Construct Validation of the FSI Oral Interview." Language Learning 31:67-86.

Bernhardt, Elizabeth and Joann Hammadou (1982). "A Decade of Research in Foreign Language Teacher Education" The Modern Language Journal 71( ):289-299.

Bragger, Jeannette D. (1985). "The Development of Oral Proficiency" In Proficiency, Curriculum, Articulation: The Ties that Bind. Ed. Alic Omaggio. Middlebury, VT. Northeast Conference.

_____ (1985). "Materials Development for the Proficency-Oriented Classroom." In Foreign Language Proficiency in the Classroom and Beyond. Ed. Charles J. James. Lincolnwood, Illinois. National Textbook Company.

Briere, E.J. (1971). "Are We Really Measuring Proficiency With Our Foreign Language Tests?" Foreign Language Annals, 4(4): 385-391.

Byrnes, Heidi (1987). "Proficiency as a Framework for Research in Second Language Acquisition." The Modern Language Journal 71(1):44-49.

_____ and Michael Canale (1987). Defining and Developing Proficiency: Guidelines, Implementations and Concepts. National Textbook Company. Lincolnwood, Illinois.

_____ (1985). "Teaching Toward Proficiency: The Receptive Skills" In Proficiency, Curriculum, Articulation: The Ties that Bind. Ed. Alice Omaggio, Middlebury, VT. Northeast Conference.

Canale, M. (1984). "A Communicative Approach to Language Proficiency Assessment in a Minority Setting" In Communicative Competence Approaches to Language Proficiency Assessment: Research and Application. Ed. C. Rivera. Clevedon Avon England Multilingual Matters Ltd.

_____ and Swain, M. (1980). "Theoretical Bases of Communicative Approaches to Second Language Teaching and Testing." Applied Linguistics, 1(10), 1-47.

Carroll, J.B. (1979). "Psychometric Approaches to the Study of Language Abilities." In Individual Differences in Language Ability and Language Behavior. Eds. C.J. Fillmore, D. Kempler, and S.Y. Wang. New York: Academic Press.

_____ (1978). "Language Proficiency Tests Developed for IEA International Study of Achievement in French as a Foreign Language." In Some Major Tests. Ed. B. Spolsky. Arlington, Virginia: Center for Applied Linguistics.

_____ (1968). "The Psychology of Language Testing." In Language Testing Symposium: A Psycholinguistic Approach. Ed. Alan Davies (pp.46-69). London: Oxford University Press.

_____ (1967). "Foreign Language Proficiency Levels Attained by Language Majors Near Graduation From College," Foreign Language Annals 1:131-151.

_____ (1961). "Fundamental Considerations in testing for English language proficiency of foreign students." In Testing the English Proficiency of Foreign Students. Washington, D.C.: Center for Applied Linguistics.

Chastain, Kenneth (1980). "Native Speaker Reaction to Instructor - Identified Student Second Language Errors." Modern Language Journal, 64(2), 210-215.

_____. (1976). Developing Second-Language Skills: Theory to Practice. Chicago: Rand McNally.

Clark, John L.D. (1978). "Interview Testing Research at Education Testing Service." Direct Testing of Speaking Proficiency Theory and Application. Ed. J.L.D. Clark. Princeton, N.J., Educational Testing Service.

_____ (1975). "Theoretical and Technical Considerations in Oral Proficiency Testing?" In Testing Language Proficiency. Eds. R.L. Jones and B. Spolsky. Washington, D.C., Center for Applied Linguistics.

_____ (1972). "FSI Language Proficiency Interview." In Foreign Language Testing: Theory and Practice, Ed. John L.D. Clark. Philadelphia: Center for Curriculum Development pp. 121-129.

Clifford, Ray T. (1980). "Foreign Service Institute Factor Scores and Global Ratings." In Measuring Spoken Language Proficiency. Ed. James R. Firth, Washington, DC. Georgetown University Press. pp. 27-30.

Curran, C.A. (1976). Counseling-Learning in Second Languages. Apple River Ill.: Apple River Press.

Dandonoli, Patricia (1987). "ACTFL's Current Research in Proficiency Testing." In Defining and Developing Proficiency: Guidelines, Implementations and Concepts. (Ed) Heidi Byrnes and Michael Canale. Licolnwood, Illinois, National Textbook Company pp. 25-74.

d'Anglejan, Alison; Birgit Harley; and Stan Shapson (1985). "Student Evaluation in a Multidimensional Core French Curriculum." The Canadian Modern Language Review, 47(1):106-124.

Davies, A. (1968). Language Testing Symposium: A Psycholinguistic Approach. London: Oxford University Press.

Delisle, Helga (1982). "Native Speaker Judgement and the Evaluation of Errors in German. Modern Language Journal, 66(1), 39-48.

Educational Testing Service (1985). The Oral Language Proficiency of Teachers in the United States in the 1980's. An Empirical Study. Self-Assessment of Oral Language Proficiency. Final Report. Eds. Hilton Thomas L., and Others. Princeton, N.J., Educational Testing Service.

Ensz, Kathleen Y. (1982). "French Attitudes Toward Speech Errors" Modern Language Journal, 66(2), 133-139.

Farhady, H. (1979). "The Disjunctive Fallacy Between Discrete Point and Integrative Tests". TESOL Quarterly, 13(3):347.

Freed, Barbara F. (1987). "Preliminary Impressions of the Effects of a Proficiency-Based Language Requirement," Foreign Language Annals 20(2):139-146.

Galloway, Vicki (1987). "From Defining to Developing Proficiency: A Look at the Decision" In Defining and Developing Proficiency: Guidelines, Implementations, and Concepts. (Ed) Heidi Byrnes and Michael Canale. Lincolnwood, Illinois, National Textbook Company pp. 25-74.

_____ (1980). "Perceptions of the Communicative Efforts of American Students of Spanish." Modern Language Journal, 64 ( ), 428-433.

Gattegno, C. (1972). Teaching Foreign Languages in Schools: The Silent Way. New York: Educational Solutions.

Gonzalez, Pino Barbara (1989). "Prochievement Testing of Speaking." Foreign Language Annals, 22(5):487-96.

Groves, Herbert J. (1985). A Guide to Curriculum Planning in Foreign Language Madison, WI: Wisconsin Department of Public Instruction.

Guntermann, Gail (1978). "A Study of the Frequency and Communicative Effects of Errors in Spanish." Modern Language Journal 62(5-6), 249-253.

Gynan, Shaw N. (1985). "Comprehension, Irritation and Error Hierarchies" Hispania 68(1):160-165.

Hammerly, Hector (1985). An Integrated Theory of Language Teaching and It's Practical Consequences. N. Burnaby, B.C. Second Language Publications.

Hanzeli, U.E. (1977). "The effectiveness of cloze tests in measuring the competence of students of French in an academic setting." The French Review, 50(6), 865-874.

Harlow, Linda L (1990). "Oral Testing of Beginning Language Students at Large Universities: Is It Worth the Trouble?" Foreign Language Annals 23(6):489-501.

Harris, D.P. (1969). Testing English as a Second Language. New York: McGraw Hill.

Heilenman, Laura K. (1983). "The Use of Cloze Procedure in Foreign Language Placement" The Modern Language Journal 67(2):121-126.

_____ and Isabelle M. Kaplan (1985). "Proficiency in Practice: The Foreign Language Curriculum." In Foreign Language Proficiency in the Classroom and Beyond. Ed. Charles J. James. Lincolnwood, Illinois, National Textbook Company, pp. 55-78.

Higgs, Theodore V. (1984). "Language Teaching and the Quest for the Holy Grail." In Teaching for Proficiency, the Organizing Principle. Ed. Theordore V. Higgs. Lincolnwood, Illinois: National Textbook Company, pp. 1-9.

_____ and Ray Clifford (1982). "The Push Toward Communication." In Curriculum, Competence, and the Foreign Language Teacher. Ed. Charles James. Lincolnwood, Illinois. National Textbook Company, pp. 57-78.

Hilton, Thomas L, and Others (1985). The Oral Language Proficiency of Teachers in the United States in the 1980's. An Empirical Study. Self-Assessment of Oral Language Proficiency. Final Report. Princeton, N.J. Educational Testing Service.

Hiple, David V. and Joan H. Manley (1987). "Testing How Well Foreign Language Teachers Speak: A State Mandate." Foreign Language Annals 20(2):147-153.

Hirsch, Bette (1985). "A Proficiency Based French Conversation Course." French Review 59(2):210-218.

Hymes, D. (1972). "On communicative competence". In J.B. Pride and J. Holmes (Eds) Sociolinguistics: Selected Readings. Harmondsworth: Penguin Books.

Ingram, E. (1978). "The Psycholinguistic Basis" In Papers in Applied Linguistics. Advances in Language Testing Series 2 Approaches to Language Testing. Ed. Bernard Spolsky. Arlington, Virginia Center for Applied Linguistics.

_____ (1968). "Attainment and diagnostic testing." In Language Testing Symposium: A Psycholinguistic Approach. Ed. Alan Davies. London: Oxford University Press.

James, Charles J. (1985). Ed. Foreign Language Proficiency in the Classroom and Beyond. Lincolnwood, Illinois. National Textbook Company.

Johnson, Carl and Bobby LaBouve (1984). "A Status Report on the Testing of Prospective Language Teachers for Initial State Certification." Foreign Language Annals 17(5):461-472.

Kaplin, Isabelle M. (1984). "Oral Proficiency Testing and the Language Curriculum: Two Experiments in Curricular Design for Conversational Courses." Foreign Language Annals 17(5):491-497.

Kelly, Louis G (1969). 25 Centuries of Language Teaching. Rowley, M.A., Newbury House.

Khalil, Aziz (1985). "Methodological Issues in Communicative Error Evaluation: A Case Study" TESOL Quarterly.

Kramsch, Claire (1986). "From Language Proficiency to Interactional Competence." The Modern Language Journal 70(4):366-372.

Lado, R. (1978). "Scope and Limitations of Interview - Based Language Testing: Are We Asking Too Much of the Interview?" In Direct Testing of Speaking Proficiency: Theory and Application. Ed. J.L.D. Clark. Princeton, N.J., Educational Testing Service.

_____ (1961). Language testing: The construction and use of foreign language tests. New York McGraw-Hill Book Company.

Lalande, John F. II (1985). "Proficiency: A Logical Outgrowth of Communicative Competence." Peals 23(2):1-4.

Lantolf, James P. and William Frawley (1985). "Oral-Proficiency Testing: A Critical Analysis. The Modern Language Journal 69(4):337-345.

Larson, Jerry W. and Randall L. Jones (1984). "Proficiency Testing for the Other Language Modalities. In Teaching for Proficiency, the Organizing Principle. Ed. Theodore V. Higgs. Lincolnwood, Illinois, National Textbook Company, pp. 113-138.

_____ (1984). Testing Speaking Ability in the Classroom: The Semi-Direct Alternative." Foreign Language Annals 17(5):499-507.

Leblanc, Leona B. (1990). "Raising the Entrance Requirements: The Foreign Language Educator's Role in Implementing Florida's New Model for Teacher Certification." Foreign Language Annals, 23(6):511-515.

Linder, Cathy (1977). Oral Communication Testing. Lincolnwood, Illinois, National Textbook Company.

Liskin-Gasparro, Judith E. (1984). "The ACTFL Proficiency Guidelines: A Historical Perspective." In Teaching for Proficiency, the Organizing Principle: Ed. Theodore V. Higgs. Lincolnwood, Illinois, National Textbook Co. pp. 11-42.

_____ (1984). "Foreign Language Oral Proficiency Testing: An Organizing Principle for Instruction." Paper Presented at the Annual Meeting of the National Council on Measurement in Education in New Orleans, LA. Education Testing Service Prinston N.J.

Lowe, Pardee Jr. (1986). "Proficiency: Panacea, Framework, Process? A Reply to Kramsch, Schulz and Particularly to Bachman and Avignon." The Modern Language Journal 70 (4).

_____ and Judith E. Liskin-Gasparro (1982). Testing Speaking Proficiency: The Oral Interview. Washinton, D.C. Center for Applied Linguistics.

Lozanov, G. (1979). Suggestology and Outlines of Suggestopedy. New York: Gordon and Breach, Science Publishers.

Ludwig, Jeannette (1982). "Native Speakers Judgement of Second-Language Learners' Efforts at Communication" A Review." Modern Language Journal 66(3), 274-283.

Magnam, Sally Sieloff (1988). "Grammar and the ACTFL Oral Proficiency Interview: Discussion and Data." The Modern Language Journal 72(3):266-276.

_____ (1987). "Rater Reliability of the ACTFL Oral Proficiency Interview." The Canadian Modern Language Review 4393):525-537.

_____ (1987). "Proficiency in Perspective in the Foreign Language Classroom." In Proficiency, Policy, and Professionalism in Foreign Language Education. Ed. Birckbichler, Diane W. Selected Papers from the 1987 Central States Conference.

_____ (1986). "Assessing Speaking Proficiency in the Undergradute Curriculum: Data from French." Foreign Language Annals 19(5):429-438.

_____ (1985). "From Achievement Toward Proficiency Through Multi-Sequency Evaluation." In Foreign Language Proficiency in the Classroom and Beyond. Ed. Charles J. James. Lincolnwood, Illinois, National Textbook Company.

Manley, Joan (1985). "The Push for Proficiency: Helping Present Foreign Language Teachers Keep Up." Report at the AATF/AATG/AATSP/AATI/ACTFL Joint Annual Meeting. New York.

Medley, Frank W. Jr. (1985). "Designing the Proficiency-Based Curriculum." In Proficiency, Curriculum, Articulation: The Ties that Bind. Ed. Alice Omaggio. Middlebury, UT., Northeast Conference.

_____ (1984). "Motivating the Language Learner Through Real Communication." ACTFL Master Lecture Series. Defense Language Institute. Monterey, CA., Foreign Language Center.

Mullen, K.A. (1980). "Rater Reliability and Oral Proficiency Evaluations" In Research In Language Testing. Eds. J.W. Oller, J. and K. Perkins. Rowley, Mass. Newbury House Publishers, Inc. 91-101.

New Brunswick Department of Education (1984). New Brunswick Middle School Scale 1984 for the Assessment of Oral Proficiency. Princeton, N.J.: Educational Testing Service.

_____ (1974). Manual for Interviewers of French. Princeton, N.J.: Educational Testing Service.

Newfoundland and Labrador Department of Education (March 1991). Education Statistics Elementary/Secondary.

_____ (Revised July 1990). French 3200 Oral Interview Testing: A Manual for Interviewers.

_____ (1986). French 3200 Oral Testing: A Manual for Interviewers.

Oller, J.W., Jr. (1978). Pragmatics and language testing. In Advances in Language Testing Series: 2: approaches to Language Testing. Ed. Bernard Spolsky. Arlington VA. Center for Applied Linguistics.

_____ (1975). "Discrete-point tests versus tests of integrative skills." In Focus on the Learner: Pragmatic Perspectives for the Language Teacher. Eds. J.W. Oller and Jack C. Richards. Rowley, Mass. Newbury House Publishers, Inc.

_____ and U. Streiff (1975). Dictation: A test of grammar based expectancies. In Testing Language Proficiency. Eds. R.L. Jones and B. Spolsky. Arlington Virginia Center for Applied Linguistics.

_____ (1972). "Contrastive Analysis, Difficulty, and Predictability." Foreign Language Annals 6(1), 95-106.

Omaggio, Alice C. (1984). The Proficiency-Oriented Classroom." In Teaching for Proficiency, the Organizing Principle. Ed. Theodore V. Higgs. Lincolnwood, Illinois, National Textbook Company pp. 43-84.

_____ (1983). "Methodology in Transition: The New Focus on Proficiency." The Modern Language Journal 67(4):330-341.

Perren, G.E. (1968). Testing spoken language: Some unsolved problems. In Language Testing Symposium: A Psycholinguistic Approach. Ed. Alan Davies. London. Oxford University Press.

Piazza, Linda Gaylord (1980). "French Tolerence of Grammatical Errors Made by Americans." Modern Language Journal 64(5):422-427.

Savignon, Sandra J. (1985). "Evaluation of Communicative Competence: The ACTFL Provisional Proficiency Guidelines." The Modern Language Journal 69(2):129-134.

_____ (1983). Communicative competence. Theory and classroom practice: Texts and contexts in second language learning. Reading, MA: Addison-Wesley.

Schulz, Renate A. (1986). "From Achievement to Proficiency Through Classroom Instruction: Some Caveats." The Modern Language Journal 70(4):373-379.

Scott, Virginia M. (1989). "An Empirical Study of Explicit and Implicit Teaching Strategies in French." The Modern Language Journal 73(1):14-22.

Shapson, Stan (1988). "An Agenda for Evaluating a Core French Curriculum." Research and Evaluation Task Force, National Core French Study, Canadian Association of Second Language Teachers.

Spolsky, B. (1978).: "Linguistics and Language Testers." In Advances in Language Testing Series 2 Approaches to Language Testing. Ed. B. Spolsky. Arlington Virginia. Center for Applied Linguistics.

Stern, H.H. (1983). Fundamental Concepts of Language Teaching. Oxford, Oxford University Press.

_____ (1983). "French Core Programs Across Canada: How Can We Improve Them?" The Modern Language Journal 67(3):235-244.

Stevick, Earl W. (1984). "Curriculum Development at the Foreign Service Institute." In Teaching for Proficiency, the Organizing Principle. Ed. Theodore V. Higgs. Lincolnwood, Illinois, National Textbook Company.

Terrell, Trace (1986). "Acquisition in the Natural Approach: The Binding/Access Framework." Modern Language Journal 70(3):213-227.

Underhill, N. (1982). "Common sense in oral testing: Reliability, Validity and Affective Factors." In On TESOL '82. Pacific Perspectives on Language Learning and Teaching. Eds. M.A. Clarke and J. Hardscombe. Washington, D.C., Teachers of English to Speakers of Other Languages.

Upshur, J.A. (1975). "Productive Communication Testing: Progress Report." In Focus on the Learner: Pragmatic Perspectives for the Language Teacher. Eds. J.W. Oller and Jack C. Richards. Rowley Mass. Newbury House Publishers, Inc.

Van Patten, Bill (1986). "The ACTFL Proficiency Guidelines: Implications for Grammatical Accuracy in the Classroom?" Studies in Second Language Acquisition 8( ):56-67.

Walz, Joel (1986). "Is Oral Proficiency Possible with Today's French Textbooks?" The Modern Lanaguage Journal 70(1):13-19.

Wilds, C.P. (1975). "The Oral Interview Test." In R.L. Jones and B. Spolsky (Eds). Testing Language Proficiency. Washington, D.C. Center for Applied Linguistics.

Wilkins, D.A. (1976). National Syllabuses. London: Oxford University Press.

Winitz, H. (1981). The Comprehension Approach to Foreign Language Instruction. Rowley Mass.: Newbury House Publishers, Inc.

APPENDIX A

Survey of Background Characteristics of Interviewer and Rating of Oral Interview

PART I Background - Education and Language Experience

1. Location. Respondents' location was determined according to specific criteria set by researcher.

2. Please indicate your sex.

   ____ (1) Male

   ____ (2) Female

3. How many university courses did you complete in French language or literature?

   _____
   (courses)

4. Including this year, how many years have you taught French?

   _____
   (years)

5. Altogether, how many MONTHS have you spent in a French-speaking milieu? Please try to be as accurate as possible, adding together all of your visits or periods of residence.

   _____
   (months)

This questionnaire is confidential and will be used only for research purposes.

6. Which of the following statements best describes the extent to which you spoke French during your stay(s) in French-speaking areas.

____(1) I was never in a French-speaking milieu.

____(2) I spoke only in English.

____(3) I used a few words of French.

____(4) I spoke French occasionally in social situations (greeting people, ordering a meal, asking directions, etc.) but otherwise used English,

____(5) I spoke French consistently in social situations, but not in situations requiring complicated, abstract, or specialized language usage.

____(6) I spoke French consistently in all situations, including work and study.

7. How frequently do you speak French in your home?

____(1) Never

____(2) Rarely, or only a few words

____(3) Occasionally

____(4) About half the time

____(5) More than half the time

8. Which statement best describes your use of French with friends or acquaintances?

____(1) I never speak French with friends or acquaintances.

____(2) I occasionally exchange a few words or short sentences in French with friends or acquaintances.

____(3) I have occasional short conversations in French with friends or acquaintances.

____(4) I speak French regularly with one or more friends or acquaintances.

This questionnaire is confidential and will be used only for research purposes.

9. This question asks you to judge your own level of speaking ability in French, Please read each one of the six paragraphs below and decide which paragraph best describes your ability to speak and to understand spoken French. Check the space next to the number preceding the <u>one</u> paragraph that best describes your speaking ability in French. Please be as honest and as accurate as possible. If you believe that your speaking ability in French is between levels, choose the lower level (e.g. the lower numbered paragraph).

_____(1) My speech in French is very limited, and I have great difficulty understanding the spoken language, even when it is spoken slowly and clearly. I cannot really communicate much information in the language.

_____(2) I can ask and answer questions about very familiar subjects and can understand simple questions and statements if they are spoken slowly, and sometimes repeated or paraphrased. My vocabulary is limited to basic needs (food, asking directions, greeting people, and so forth). I make many grammatical mistakes but can usually be understood by French-speakers who are used to dealing with foreigners. I can order food in a restaurant, get a room in a hotel, introduce myself to people, ask directions on the street, and understand the reply, if it is not too complicated.

_____(3) I can talk with native speakers of French about myself and my family, my job, studies, or hobbies. I can narrate and describe with grammatical accuracy events in the past, present, and future. I can understand most simple conversations about concrete topics. Though I may occasionally need help, I can handle limited work requirements. My grammar is fairly good but I make mistakes with particular thought or object, I can usually describe it by using other, easier words. My accent, though far from native, is understandable.

_____(4) I can talk about professional topics with ease, and am able to state and support my opinions. I can understand almost everything spoken at a normal rate of speech by native French speakers. My vocabulary is good enough so that I usually know most or all of the words for what I want to say. My control of grammar is good and any mistakes I make are usually with the more complicated constructions.

This questionnaire is confidential and will be used only for research purposes.

____(5) I can speak fluently and accurately about almost any subject with which I am familiar, including professional, abstract, or controversial topics. I can always understand native French speakers, even when they are speaking quickly and using sophisticated or colloquial expressions. My vocabulary is very extensive and I make very few grammatical errors.

____(6) I have complete fluency in French. My speech on all levels is fully accepted, in all of its features, by educated native speakers. This includes vocabulary and idiom, colloquialisms, and cultural references.

10. Listed below are a number of "can do" statements about a person's speaking ability in French. Read each description carefully and indicate by circling a number in one of the four columns, whether you would be able--at the present time--to carry out this task "quite easily," "with some difficulty," "with great difficulty," or "not at all." Assume in each case that you have the necessary facts. Base your response on whether you can perform the linguistic task indicated.

	Quite Easily	With Some Difficulty	With Great Difficulty	Not at all
A. Talk about my favourite hobby at some length, using appropriate vocabulary	1	2	3	4
B. Tell what I plan to be doing 5 years from now, using appropriate future tenses. . . .	1	2	3	4
C. Argue your case with the principal for having a student suspended from school. . . .	1	2	3	4

This questionnaire is confidential and will be used only for research purposes.

	Quite Easily	With Some Difficulty	With Great Difficulty	Not at all
D. Give a prepared half-hour formal presentation in a professional topic of interest (e.g. individualized instruction, teaching multi-level classes, organizing a Foreign language Week Program). You may use notes but not read from a prepared text .	1	2	3	4
E. State and support with examples and reasons, a position on a controversial topic (for example, birth control, nuclear safety, environmental pollution . . .	1	2	3	4

This questionnaire is confidential and will be used only for research purposes.

11. Regardless of how well you currently speak French, please answer each of the following in terms of your present level of listening comprehension in French. Circle the number that best describes your level of listening comprehension of each of the following:

		Quite Easily	With Some Difficulty	With Great Difficulty	Not at all
A.	In face-to-face conversation, understand native speakers who are speaking to me as quickly as they would to another native speakers .	1	2	3	4
B.	On the telephone, understand a native speaker who is talking as quickly and as colloquially as he or she would to a native speaker of the language . . .	1	2	3	4
C.	Understand two native speakers when they are talking rapidly with one another . . .	1	2	3	4
D.	Understand movies without subtitles . .	1	2	3	4

This questionnaire is confidential and will be used only for research purposes.

PART II

I. General Questions

1. How do you rate the importance of each of the factors in speaking proficiency? Rate in order of importance (1 - 5).

(1) most important
(5) least important

_____ pronunciation _____ fluency

_____ grammar _____ comprehension

_____ vocabulary

2. How comfortable do you feel you are with the proficiency scale for the French 3200 oral exam? (check [_] one).

_____ very comfortable _____ not at all comfortable

_____ comfortable

3. In rating speaking proficiency which of the following do you use to aid you in determining a level for the interviewee. (check [_] one).

_____ 1. Levels of Proficiency scale as outlined in Manual for Interviewers pp 3-5.

_____ 2. Chart for "Factors in Speaking Proficiency" as outlined in Manual for Interviewers p. 8.

_____ 3. A set of criteria based on information in the "Manual for Interviewers" and gleaned from training sessions.

4. Do you feel there are other factors which may influence the interview rating? please specify in space below.

Comments e.g. number of interviews to be conducted per day.

_____

_____

_____

_____

This questionnaire is confidential and will be used only for research purposes.

5. Are you more aware of the goals of the program because of the guidelines? (check [_] one).

Yes _____ No _____

Comment _____

_____

_____

_____

_____

_____

6. As a result of the proficiency levels, have you or will you change your method of teaching?

Yes _____ No _____

Comment _____

_____

_____

_____

_____

_____

This questionnaire is confidential and will be used only for research purposes.

II Rate the acceptability for the following errors on a scale of 1 to 4. Use the scale outlined below and circle your choice.

Scale

1. not at all acceptable
2. rarely acceptable
3. occasionally acceptable
4. acceptable

1.	lack of subjective-verb agreement	1	2	3	4
2.	lack of noun-adjective agreement	1	2	3	4
3.	misuse of partitive articles	1	2	3	4
4.	misuse of most forms of negative	1	2	3	4
5.	inaccurate employment of direct/indirect object pronouns	1	2	3	4
6.	gross inaccurate use of articles	1	2	3	4
7.	inaccurate use of irregular verb forms	1	2	3	4
8.	inaccurate word order of adjectives	1	2	3	4
9.	inaccurate use of vocabulary items	1	2	3	4
10.	inaccurate use of verb tenses	1	2	3	4
11.	inaccurate use of idiomatic expressions	1	2	3	4

This questionnaire is confidential and will be only used for research purposes.

III Rating for proficiency Levels

At what level would you rate an interviewee who displays the following language characteristics.

Circle the level of proficiency. 1 2 3 4 5 according to your interpretation of the proficiency levels outlined for the French 3200 oral exam.

Pronunciation

1. Frequent gross errors impeding initial comprehension but message is generally communicated after repetition.

1 2 3 4 5

2. Good attempt at native French accent, but with mispronunciation which require repetition to be clearly understood.

1 2 3 4 5

3. Accent requires concentrated listening and mispronunciations lead to occasional misunderstanding and apparent error in grammar or vocabulary.

1 2 3 4 5

4. Accurate use of vocabulary and grammar but accent is markedly not French.

1 2 3 4 5

Grammar

1. Frequently employs the infinitive form of irregular verbs.

1 2 3 4 5

2. Consistently employs present tense while responding to questions with reference to past or future tense.

1 2 3 4 5

This questionnaire is confidential and will be used only for research purposes.

3. Accurately (subject-verb agreement) employs only selective past tense verb forms.

1 2 3 4 5

4. Inaccurate use of articles indicating limited concept of gender.

1 2 3 4 5

Vocabulary

1. Extensive range of vocabulary but sentence structure is fragmented, however, not hampering communication.

1 2 3 4 5

2. Circumlocutes only on topics of student interest and special field of competence.

1 2 3 4 5

3. Inconsistent use of vocabulary in identifying objects, food and personal interests.

1 2 3 4 5

4. Understands simple speech on common social and familiar topics requiring some repetition and rephrasing.

1 2 3 4 5

Fluency

1. Initiated conversation is slow and uneven except in routine high frequency phrases.

1 2 3 4 5

2. Speech is occasionally hesitant with some unevenness caused by rephrasing and groping for accurate vocabulary.

1 2 3 4 5

This questionnaire is confidential and will be used only for research purposes.

3. Speech is frequently hesitant and jerky; sentences may be left incomplete when interviewees feels his point of view has been comprehended by interviewer.

1 2 3 4 5

4. Speech is characterized by an initiated rhythmic flow which sustains conversation but vocabulary and grammar are inaccurate.

1 2 3 4 5

This questionnaire is confidential and will be used only for research purposes.

Thank you for completing the questionnaire.
Please return it by mail in the envelope provided to:

Kevin Flynn
41 Virginia Place
St. John's, Newfoundland
A1A 3G5

Telephone: 722-2930

P.S. If you wish to have the results of the survey, please complete the form and enclose it with the questionnaire.

Name _____

Address _____

_____

Postal Code _____

APPENDIX B

TABLE B1

Other Factors Influencing Interview Rating

1. Nervousness
2. Teacher and student fatigue
3. Timing of interview period
4. Number of interviews per day
5. Choice of topics for discussion
6. Amount of second language exposure outside the classroom,
7. Physical setting i.e. distractions interruptions
8. Student/teacher rapport
9. Teacher expectations
10. Student's expectations
11. Interviewees personality
12. Getting back into the groove of interviewing
13. Possible interference from knowing the student
14. Experince in interviewing
15. Tendency to measure personality of student
16. Speaking proficency of interviewer
17. Use of equipment i.e. tape recorder
18. Number of practice interviews
19. Difficulty of having pre-determined proficiency level for students especially in small size classes
20. Student's attitude
21. Teacher workload

APPENDIX C

TABLE C1

Oral Proficiency Compared With Speaking Extent and Proficiency Levels

	Description of Statements	Significance Levels	
		Extent	Level
A.	Talk about favorite hobby	.02662	.00053
B.	Talk about future plans	.00196	.00392
C.	Argue a case	.00021	.00027
D.	Formal presentation	.00047	.00257
E.	Argue a controversial topic	.02506	.00041

TABLE C2

Aural Comprehension Compared With Speaking Extent and Proficiency Levels

	Description of Statements	Significance Levels	
		Extent	Level
A.	Understand face to face conversations	.00061	.00037
B.	Telephone conversations	.0002	.0000
C.	Understand two native Speakers	.00061	.0000
D.	Understand movie	.02472	.0000