



Ontario
College of
Teachers

Ordre des
enseignantes et
des enseignants
de l'Ontario

Case Study

The Ontario College of Teachers French Language Proficiency Assessment Project

This case study has been made possible by a financial contribution from the Official Languages Support Branch of the Department of Canadian Heritage, under the *Promotion of Linguistic Duality* component of the *Enhancement of Official Languages* Program.

Table of Contents

Table of Contents	2
1. Context.....	3
1.1 The Ontario College of Teachers.....	3
1.2 Issues.....	4
1.3 Plan for the Improved Delivery of French-Language Services.....	4
1.4 Rationale and Scope of the Case Study.....	5
2. Challenges.....	5
3. Strategies Implemented.....	7
3.1 Selection of Reference Framework for Job Classifications	7
3.2 Application of the Framework to the Bilingual Positions.....	8
3.3 Proficiency Assessment.....	9
4. Challenges and Solutions.....	10
4.1 Fostering a Supportive Climate for Change	10
4.2 Realistic and Achievable Timetable	12
4.3 Building from within – training and retention.....	12
4.4 Resources.....	12
4.5 Future Challenges	13
5. Key Success Factors	13
6. Conclusion	14

Appendices

1	Designated Bilingual Positions
2	Existing Reference Frameworks Analysis
3	Groupings of Designated Bilingual Occupations
4	Final Report of the Consultant: French Language Proficiency Assessment Project

1. Context

1.1 The Ontario College of Teachers

The Ontario College of Teachers (the “College”) was established in 1997 by the Government of Ontario, under the *Ontario College of Teachers Act* (the “Act”), which was passed on July 5, 1996.

The College currently has 206,859 members, 11,817 of whom—or approximately six per cent—communicate with the College in French. The College’s members are mainly elementary and secondary school teachers, principals, supervisory officers and members of faculties of education. It is the largest professional body in Canada, and has 154 full-time employees.

Most members of the College are teachers, but principals, vice-principals, supervisory officers and directors of education must also belong to the College in order to hold their positions. The College is funded by its members’ annual fees, which are currently set at \$104.

The College has the following main objects:

- To regulate the profession of teaching and to govern its members
- To develop, establish and maintain qualifications for membership in the College, issue certificates of qualification
- To issue, renew, amend, suspend, cancel, revoke and reinstate certificates of qualification
- To accredit teacher education and professional learning programs
- To receive and investigate complaints against members of the College and to deal with discipline and fitness-to-practise issues
- To establish and enforce professional standards and ethical standards applicable to members of the College.

The College is governed by a 37-member Council—23 members of the College elected by their peers (three of whom are francophones) and 14 members (including one francophone) appointed by the Lieutenant Governor of Ontario.

1.2 Issues

The Act provides that a person has the right to use French in all communications with the College, at all times, and that the Council shall take all reasonable measures and make all reasonable plans to ensure that persons may use French in all communications with the College.

Thus, out of a concern for equity and in order to provide quality services in both French and English, the College strives constantly to improve both the quality of French-language services and access thereto for its members who are part of Ontario's francophone community. The objective is to be able to provide all of its services in both official languages to all of its members, its partners and the community.

Accordingly, and on the basis of the Registrar's consultations with the College's francophone partners, a new French Language Services Co-ordinator position was created. The Co-ordinator, who has been in the position since the fall of 2005, has developed a Plan for the Improved Delivery of French-Language Services. It is designed to enhance the provision of French-language services at every level of the College's operations and thus ensure that *all services* are available in French *at all times*.

1.3 Plan for the Improved Delivery of French-Language Services

The Plan for the Improved Delivery of French-Language Services is based on three main action components. The first two concern community relations and communications, while the third (the Language Proficiency Assessment Project) concerns human resources and is aimed essentially at increasing the number of bilingual positions at the College. The College received funding from Canadian Heritage¹ to assist with establishing this project.

¹ *Promotion of Linguistic Duality* component of the *Enhancement of Official Languages* Program, Official Languages Support Branch

1.4 Rationale and Scope of the Case Study

As part of the commitment under the funding from Canadian Heritage, the College undertook to write a case study describing the project set-up in the following terms: context, objectives, issues and key stakeholders, the processes implemented and the approach used, the results of the tool-seeking process, the costs involved and the outcomes.

This case study is made public in order to support and promote the transfer of the learning acquired during the process and, ultimately, to encourage a broader exchange of information between learning communities and organizations or companies that are considering similar programs.

The case study focuses on the strategies funded by Canadian Heritage and implemented by the College through its French Language Proficiency Assessment Project.

2. Challenges

The Plan for the Improved Delivery of French-Language Services consisted primarily of enhancing French-language services by increasing the number of bilingual positions. Working closely with senior management, the human resources manager and the other College co-ordinators, the French Language Services Co-ordinator began by carefully studying the nature of the existing bilingual positions.

The purpose of this exercise was to assess the actual capacity of each unit in the College to provide all French-language services whenever they were required. This initial analysis was a team effort involving the active participation and support of senior management in each unit. Two key observations emerged:

Observation 1: All of the units in the College offered French-language services.

Observation 2: These services were not sufficiently staffed to deliver these services to everyone wishing to make use of them at all times, as required by the Act.

Two strategies helped mitigate the deficiencies of the second observation:

- **Hiring of additional bilingual staff** – In some cases, all that was needed was to hire more bilingual staff for positions that already existed. This was especially the case with the contact centre in the Membership Services Department, where there were not enough bilingual employees to allow for members of the College to receive French-language services during all business hours.
- **Increase in the number of bilingual positions** – In other cases, it was necessary to add new types of bilingual positions. For example, there was a need to newly designate some College leadership positions (among the Registrar, Deputy Registrar, Co-ordinator (director level) and manager roles) to provide high level communication capability, and to establish a bilingual library assistant position to provide French-speaking members with this important service.

As a result of this comprehensive exercise aimed at improving the provision of French-language services in every unit in the College, the Plan for the Improved Delivery of French-Language Services foresaw moving to a minimum of 55 bilingual positions by 2009, or 29 more than the 26 positions that existed in 2005) covering 38 different occupations (Appendix 1).

The planning exercise raised certain implementation challenges:

Linguistic Challenges:

- What reference framework should be used to develop the appropriate operational definitions for each required language skill area (speaking, listening, reading and writing in French)?
- What proficiency levels should be required for each bilingual position identified? How would the College ensure that future employees, in each bilingual position, achieve the required proficiency level?

Organizational and Human Resources Challenges:

- How to ensure that the exercise is participatory, inclusive and respectful of everyone, and that it gains maximum commitment of staff and managers.
- How to ensure that the implementation of the French Language Proficiency Assessment Project does not create undue insecurity for current staff members who are unilingual anglophone or do not have the French language proficiency levels that are required or will eventually be required
- How to avoid negatively affecting the career plans of current employees through the new language proficiency requirements

The College thus had to:

- Identify and develop a reference framework to use in assessing the language proficiency of College staff
- Assess the language skill levels required for each bilingual position, according to the reference framework developed in Step 1
- Develop assessment tools to determine actual language skill levels
- Communicate effectively with College staff and managers.

3. Strategies Implemented

The College decided to hire an outside consultant specialized in the provision of linguistic services. The consultant was commissioned to perform several basic steps.

3.1 Selection of Reference Framework for Job Classifications

What reference framework should be used to develop the appropriate operational definitions for each required skill level in speaking, listening, reading and writing in French?

Step 1 consisted of identifying existing reference frameworks for the classification of the College's bilingual positions. Could existing frameworks be used or adapted for the College's specific needs? Several frameworks were analyzed by the French Language Services Co-ordinator, working in close cooperation with the human resources manager, specifically those used by:

- The Government of Canada
- The Government of Ontario
- The City of Ottawa
- The Centre for Canadian Language Benchmarks

- The American Council for the Teaching of Foreign Languages/Oral Proficiency Interviews (OPI)
- The Common European Framework of Reference for Languages
- Educational Testing Service

Please refer to Appendix 2 for the complete report of the existing reference frameworks analysis. The College selected the framework of the Centre for Canadian Language Benchmarks because it corresponded most closely to its needs:

- The 12 Benchmarks for the four language skills (speaking, listening, reading and writing – see Appendix 3) used by the Centre for Canadian Language Benchmarks are sufficiently precise at the intermediate and advanced levels to enable accurate and appropriately differentiated classifications of the College's varied bilingual positions
- The descriptors for each Benchmark are well adapted to the Canadian context and the types of jobs offered at the College.

It should be noted here that the outside company, working closely with the College's human resources manager, adapted the descriptions of each of the 12 Benchmarks for the four language skills to the College's needs (Appendix 3).

3.2 Application of the Framework to the Bilingual Positions

The four Canadian language skill areas (speaking, listening, reading and writing) and their 12 respective Benchmarks had to be applied to each of the College's 38 bilingual occupations. What proficiency levels should be required for each of the College's 38 bilingual occupations?

Addressing this step comprised a three-part process:

- Development of a draft rating grid by the external consultant with the help of the human resources manager, indicating the benchmark required (from one to 12) for each of the 38 bilingual occupations and each of the four language skill areas (speaking, listening, reading and writing). A sample from the grid below is given by way of illustration:

Canadian Language Benchmarks	Speaking	Listening	Reading	Writing
Membership Services Department				
Client Services Assistant	9	9	7	7
Senior Evaluator	10	10	11	11
Membership Records Assistant	8	8	9	8

- The draft rating grid was subsequently validated and revised as required in consultation with the departmental co-ordinators and managers of each unit
- The human resources manager and the French Language Services Co-ordinator then validated the revised changes and confirmed the ratings. The College's senior management approved the final version of the the resulting grid in March 2007.

It is important to note that the calibre of the documentation provided to the consultant with respect to job descriptions, and specifically the nature of the French language communication requirements of each occupation was instrumental in facilitating the process, and that this is a crucial prerequisite to the implementation of an analytical process of this nature.

3.3 Proficiency Assessment

How to ensure that the required proficiency levels are in fact achieved for each bilingual position?

To answer this question, we evaluated the same language proficiency systems described earlier, but this time from the standpoint of their appropriateness with respect to available tools for proficiency assessment. None of them, however, could satisfy the College's specific occupational assessment needs.

- The Canadian Language Benchmarks offered the best classification in terms of language profiles, but they didn't offer a test to assess the skills.
- Among the other assessment tools considered, only the language proficiency tests of the American Council for the Teaching of Foreign Languages (ACTFL) offered standardized assessments of functional speaking ability that the College considered pertinent. In the end, however, with the assistance of the external consultant, it was concluded that the available tests associated with this framework were too limited in scope and focussed more on achievement of targets, rather than providing broad occupationally relevant profiles of language proficiencies that could be applied to the 12-point CFB benchmarks.

The College therefore gave the outside consultant the additional task of developing tools to assess French-language proficiency in the four language skill areas (speaking, listening, reading and writing) and the required benchmarks adapted by the College for each of the bilingual occupations.

Rather than developing a series of tests adapted uniquely to each of the 38 bilingual occupations – which would have been costly, very labour-intensive and lead to extensive delays in implementation – the College adopted a more practical approach, organizing the 38 occupations into seven groups of positions: investigations, evaluation, secretarial services, client services, public relations, communications and leadership and/or teaching (Appendix 4).

The seven series of tests to be developed relate to the College's unique environment and to the education community. They cover the following aspects all of which focus on content that is occupationally relevant to the different groups of positions:

- Interview and role-playing questions to assess speaking proficiency
- Structured oral statements with questions to assess listening comprehension proficiency
- Texts and questions to assess reading comprehension proficiency
- Written tests to assess writing proficiency.

The College then plans to validate the assessment process with volunteer staff and managers. If needed, the tests and/or benchmarks will be modified prior to use in future recruitment events. We plan to treat the first year of implementation as a pilot with a review the appropriateness of the tests and benchmarks in 2008.

4. Challenges and Solutions

4.1 Fostering a Supportive Climate for Change

Organizational changes as far-reaching as the College's Plan for the Improved Delivery of French Language Services encompass many challenges. Resistance to change, very often the result of the uncertainty and personal insecurity of current staff and managers facing large-scale change, is a challenge the College was able to address in a simple yet effective manner. In this regard, the College leadership established a systematic communications program with clear rules and monitoring procedures that fostered engagement in the planning and addressed the security and career interests of current staff.

- College leadership ensured a participatory and inclusive process right from the beginning of the project with extensive consultation and face-to-face communication meetings. The French Language Services Co-ordinator and the human resources manager took the time to educate and inform staff who were affected by the policies.
- Implementation was based in respect and care for employees affected in various ways by the new policies:
 - 1) Jobs newly designated as bilingual are affected by the plan only when a non-bilingual incumbent chooses to move on. Thus, no non-French speaking staff members are affected. Natural attrition opportunities afford the College the opportunities to increase the bilingual complement as required.
 - 2) Employees who are bilingual (or identified as such prior to the setting-up of the Language Proficiency Assessment Project) are not required to complete proficiency tests in order to keep their positions. The College also selected a “portfolio” approach, which enables employees (or potential employees) with a high level of French language fluency to demonstrate their language proficiency without resort to proficiency testing.(Appendix 5)
 - 3) Implementation continues some career progress for non-French speaking staff and managers through two mechanisms.
 - Where there is an objective of newly establishing as bilingual one of several leadership positions in a department, the first and second vacancies that occur during the plan implementation result in French language proficiency posted as an asset. If a fully qualified candidate with the French language proficiency is not found, an appointment may be made among qualified non-French-speaking candidates. Should a bilingual appointment not be made in the first two leadership opportunities, the third posting includes the French language proficiency as a requirement.
 - For multi-incumbent occupations where the plan intends the additions of several further bilingual jobs, there is rotation of bilingual and unilingual English postings as vacancies occur.

Through these procedures, the plan moves forward while continuing to afford career opportunities for non-French speaking staff members.

4.2 Realistic and Achievable Timetable

- The College chose to implement the project over a three- year period instead of introducing the changes in one fell swoop. This was based on a detailed assessment of the historic rate of attrition for various departments affected by the change.

4.3 Building from within – training and retention

- The College established a new intensive French-language training program to enable staff with an existing foundation in French language proficiency to achieve the levels required under the new benchmarks. Through immersion studies, intensive writing courses and job exchanges College staff will have opportunities during the plan implementation and beyond to strengthen their French-language proficiency with a focus on increased quality of French language services.
- This investment in current staff training has garnered a strong response with 20 individuals requesting and approved for training in the first year of the new program availability.
- This investment in people is expected to support the retention of highly qualified staff with French-language proficiency.

4.4 Resources

For some organizations, the setting-up of such a process can be challenging from a budget perspective. The strong commitment of the Registrar and the College Council was essential for this project to proceed with the resources required. Here is the breakdown:

- Initial studies (partially funded by Canadian Heritage): \$ 85,000
- Implementation of the language-training program: \$127,000 per year (training, accommodation, travel)
- Cost of assessing applicants to bilingual positions: \$20,000 per year
- Development of language proficiency tests: \$6,000

4.5 Future Challenges

Despite the amount of painstaking work devoted to the setting-up of its French-Language Proficiency Assessment Project, the College still has certain questions:

- Will we be able to fill all the bilingual positions, at the required high proficiency levels, from the pool of candidates available in Toronto?
- Can the project be fully completed by 2009?
- What assessment approaches will be appropriate for assessing the value of the annual training investment ?
- How will the project affect the communication assessment processes established for unilingual English positions?

5. Key Success Factors

The College identified several key factors that will maximize the project's chances of succeeding:

- **Genuine political will conveyed by clear and effective leadership:** the College's chief administrators (the Council and the Registrar/Chief Executive Officer) are the real leaders in setting up the project. With their unwavering support, the College's senior managers were able to work in close cooperation, keeping in mind the primary objective of improving the delivery of French-language services.
- **Respect for individuals:** the solutions are implemented with respect for individuals through the introduction of non-threatening rules designed to preserve a feeling of security among College staff. The attrition-based tactic and the close cooperation among management staff (French Language Services, Human Resources, department co-ordinators, the Registrar's Office) are illustrations of this respect.

- **Communication and transparency:** at each step of the process, the various decisions were presented to each unit by the French Language Services Coordinator and the human resources manager, along with the relevant management staff in charge of each department in a transparent and straightforward manner. Several information sessions were organized and those employees who were most affected by the new measures were able to express their concerns freely.
- **The time factor:** the College had the good judgment not to rush matters, but to carry out the changes over a three-year period.
- **Consistent budgets:** the entire process was achievable because the College Council allocated the appropriate required funding.
- **Support and advice from Canadian Heritage:** Canadian Heritage's initial recommendations, specifically those concerning potential consultant resources, enabled the College to clearly identify and provide a framework for the implementation of the operational strategies for the Language Proficiency Assessment Project.

6. Conclusion

The setting-up of the French-Language Proficiency Assessment Project required considerable work by the College. This work could not have been accomplished without a cooperative approach on the part of all the internal partners. The College's healthy corporate culture also substantially facilitated the implementation of the new policies: high standards of quality, the professionalism of human resources, transparent communications and results-based actions are some of the factors that made it possible to meet the challenges that arose.

Appendix 1: Designated Bilingual Positions

The bilingual staffing plan includes one or more positions for each occupation listed below, for a total of 38 bilingual occupations:

Membership Services Department

Co-ordinator
Manager
Senior Evaluator
Evaluator
Membership Analyst
Evaluation Assistant
Client Services Assistant
Client Services Researcher
Client Services Officer
Correspondence Assistant
Membership Records Officer
Membership Records Assistant
Membership Records Researcher
Secretary
Receptionist

French Language Services Department

Co-ordinator
External Relations Officer
External Relations Assistant
Executive Secretary

Investigations and Hearings Department

Co-ordinator
Manager
Investigator
Intake Assistant

Professional Affairs Department

Co-ordinator
Manager
Program Officer, Accreditation

Accreditation Assistant
Program Officer, Standards of Practice and Education
Executive Secretary
Secretary

Executive Department

Leadership positions (Registrar, Deputy Registrar, Executive Co-ordinator and/or managers)
Translator
Communications Officer
Production Officer
Human Resources Assistant
Executive Secretary
Library Assistant
Help Desk Specialist

Appendix 2: Linguistic Profiles for Bilingual Positions at the Ontario College of Teachers

TABLE OF CONTENTS

	Page
Introduction	3
Project goals	3
Methodology	3
Definition of language proficiency	4
Required nature of a language proficiency test	4
Description of proficiency levels	6
- Government of Canada	6
- Government of Ontario	7
- City of Ottawa	9
- Canadian Language Benchmarks	9
- ACTFL Proficiency Guidelines	11
- Common European Framework	13
- Educational Testing Service	14
Assessment tools	15
- Government of Canada	15
- Government of Ontario	19
- City of Ottawa	22
- Canadian Language Benchmarks	26
- ACTFL Proficiency Guidelines	27
- Test de connaissance du français	30
- Test de français international	33
- Test d'évaluation du français	35
Organization of the test information for each test	
• Overview	
• Background	
• Test design and theoretical foundation	
• Components	
• Reliability and validity	
• Scoring	
• Test administration	

- Administration time
- Cost

Recommendations	37
Tests for occupational purposes	40
Remarks on reliability and validity	40
Proposed model	40
Proposed assessment tool	41
Conclusion	42
References	43
Appendices	
Appendix A: ACTFL 1989 Proficiency Guidelines	
Appendix B: Canadian Language Benchmarks, 2000	
Appendix C: Rating Scales Checklist (per model)	
Appendix D: Test Selection Checklist (per test)	

INTRODUCTION

Testing for language proficiency has long been the subject of research and debate. Educational institutions have mostly been concerned with testing for accurate placement in their language training courses or requirements for academic purposes (assessing second language knowledge to ensure adequate participation in a university curriculum). Over the last few years, testing for occupational purposes has become more widespread. Unfortunately, Human Resources departments have often used and are still often using tools that were not designed to assess level of language proficiency in the work place. The ensuing unreliability of the process is also cause for heated discussions among test developers and marketers.

PROJECT GOALS

The purpose of this document is to identify systems in order to establish the necessary French language speaking, writing, oral comprehension and reading comprehension competency levels for various Ontario College of Teachers bilingual jobs providing services to French language stakeholders and community.

It is to recommend a system for the College that is appropriate to the range of designated College jobs.

It is to research, compile and analyze existing models and terms of reference with respect to French language proficiency assessment, particularly those used by governments of Canada, Ontario and Ottawa and other relevant jurisdictions and organizations.

Information will be provided about testing systems in use in the federal government, in the Ontario provincial government, at the municipal level in Ottawa. It will describe benchmarks used in Canada, a widely-used system in the United States and the system now in use in Europe and an American-born system now in use in Canada. It will then describe the assessment tools that have been linked to these systems.

It is also in the purview of this project to suggest possible formats for a proficiency test for use by the Ontario College of Teachers.

METHODOLOGY

We will first take a look at language proficiency and its nature. Then, we will delineate rating scales models; study assessment tools using these rating scales. They will be summarized and recommendations will be made on options, including use of currently available systems, and assessment tools and adapting through modifications of current systems, and assessment tools.

First, rating scales or proficiency guidelines will be shared. Assessment tools do not exist in a vacuum. They need to be linked to a rating scale or proficiency description.

Though proficiency tests used at the University of Toronto and University of Ottawa were first considered, they were left aside as irrelevant to the College's situation. The College seeks to find an occupationally-related assessment tool for its employees. It needs to be applicable to hiring purposes as opposed to academic purposes.

DEFINITION OF LANGUAGE PROFICIENCY

The concept of language proficiency has long been the source of numerous debates among linguists. Indeed, varying opinions continue to be found. It is however crucial for this project to choose a definition that will form the basis of our reasoning as we strive to study existing French language proficiency tests and examine if their theoretical framework actually supports the proposed test items.

Language is more than the sum of pronunciation, grammar and vocabulary. It does not function in isolation. It exists for an individual to be able to communicate with another or a group of others. Language is therefore dynamic, contextually-dependent (it varies according to the situation, the status of the speakers, the topic); it is discursive (it requires connected speech) and it requires the use of integrative skills to achieve communicative competence as Canales explains in “Linking language assessment to classroom practices”.

For the purpose of this project, we rely on A.C. Omaggio’s definition of a proficiency test: a tool “to measure an individual’s general competence in a second language, independent of any particular curriculum or course of study.” We further adhere to J.L.D. Clark’s definition of a proficiency test as being: “any measurement procedure aimed at determining the examinee’s ability to receive or transmit information in the test language *for some pragmatically useful purpose within a real-life setting.*” The latter is our operational definition.

Four language skills will be considered here: speaking or the ability to use language appropriately and effectively orally; listening – the ability to understand the oral language of others and eventually act upon it.; reading – the ability to comprehend and interpret texts of differing natures; writing – the ability to produce written text with content and format appropriate to the communication purpose.

REQUIRED NATURE OF A LANGUAGE PROFICIENCY TEST

The nature of a proficiency test should reflect the purpose for which it will be used. That is why placement tests (for language training purposes) and tests for academic purposes have been left aside. Language testing should reflect the kind of contextualized language the test taker will use.

We are considering **communicative proficiency** as a goal. In second language testing for communicative proficiency in the College’s context, a level of proficiency needs to be determined for individuals to be deemed bilingual according to the various tasks and context they will be called upon to interact.

Ideally, test components would reflect the various tasks test takers have to accomplish in their work milieu.

Canale and Swain (1980)’s model of communicative competence proposes four major components: grammatical competence, sociolinguistic competence, discourse competence and strategic competence. The Canadian Language Benchmarks, for their

part, delineate linguistic competence, textual competence, functional competence, socio-cultural competence and strategic competence. The American Council for the Teaching of Foreign Languages' guidelines consider global tasks or functions; context and content; accuracy; text types. The Common European Framework considers twelve qualitative categories relevant to oral assessment: turn taking strategies, co-operating strategies, asking for clarification, fluency, flexibility, coherence, thematic development, precision, sociolinguistic competence, general range, grammatical accuracy, vocabulary control, phonological control.

The Canadian Language Benchmarks distinction is both realistic and practical when it comes to actually using an assessment tool. Linguistic competence is the knowledge of grammar and vocabulary at a sentence level. Textual competence is defined as the knowledge and application of cohesion and coherence rules and devices in building larger texts/discourse. Functional competence is competence to convey and interpret communicative intent (or function) behind a sentence, utterance or text. Socio-cultural competence focuses on appropriateness in producing and understanding utterances. Strategic competence manages the integration and application of all the other language competence components to the specific context and situation of language use. It is therefore clear that a language proficiency test should incorporate more than linguistic elements such as phonology, morphology, lexicon and syntax.

A test also needs to demonstrate **validity**. Validity is the most important criterion to evaluate a test. It purports to knowing whether and to what extent a test measures what it claims to measure. Several criteria can be taken into account to answer the validity question: face validity, content validity, construct validity, predictive validity and native speaker performance.

Face validity answers the question: Does it look like a reasonable test?

Content validity asks: Do the items or tasks in the test match what the test as a whole is supposed to assess?

Construct validity's concern is: Do the items or tasks in the test match the theory behind it?

Predictive validity is the degree to which a test can predict how well an individual will do in a future situation.

The native speaker performance criterion asks: Would examinees test at a superior level if tested orally in their native language?

A test should also be **reliable**.

Intra-rater reliability asks: "Is the test consistent in the scores it gives on second or subsequent applications to the same candidates?"

Inter-rater reliability is the degree to which an instrument yields the same results for the same individual at the same time with more than one assessor.

Scorability is a factor here as the College's Human Resources department needs to use a test that will be easy to score.

The **representativity** factor is also important: Is the test representative on both content and scope levels? Hence, the necessity to consider similar purpose tests.

Finally, the **practicality** aspect of the test is also considered: Is the test easy to administer?

DESCRIPTION OF PROFICIENCY LEVELS

Because the next component of this project refers to the determination of levels of proficiency for employees to be considered **bilingual**, only the high end levels of proficiency of the various systems studied will be incorporated here for each of the four skills: oral expression, oral comprehension, reading comprehension and written expression. This does not mean the following would be the only levels considered in the determination of linguistic profiles. Providing the whole of all scales would have been cumbersome. Once a system is selected, the determination of linguistic profiles will look at all levels within the system.

GOVERNMENT OF CANADA

Here are the descriptions for the highest level determined by the Government of Canada:

Oral interaction (includes oral expression and oral comprehension) - level C

“Minimum level of second language ability in oral interaction that should be identified for positions which require handling sensitive situations where the understanding and expression of subtle, abstract, or complicated ideas is required or where unfamiliar work-related topics must be dealt with. A person at this level can support opinions or understand and express hypothetical and conditional ideas. However, the ease and fluency of a native speaker is not required or expected. There may be errors and deficiencies in pronunciation, grammar and vocabulary but such errors rarely interfere with communication.”

Reading Proficiency – level C

“Is the minimum level of second language ability in reading that should be identified for positions requiring the comprehension of texts dealing with a wide variety of work-related topics? At this level most complex details, inferences and fine points of meaning are understood. Specialized or less familiar material can also be read with good comprehension. Some seldom-used expressions may be missed, however, and there may be some difficulty with very complex grammatical structures.”

Writing Proficiency – level C

“Is the minimum level of second language ability in writing that should be identified for positions that require a person to write explanations or descriptions in a variety of informal and formal work-related situations? At this level, a person can write texts in

which the ideas are developed and presented in a coherent manner. The style of presentation and use of vocabulary, grammar and spelling are generally appropriate and require few corrections.”

GOVERNMENT OF ONTARIO

Among the four main levels of proficiency the Government of Ontario recognizes, only the advanced and superior levels are considered functional.

Oral proficiency (includes oral expression and oral comprehension) – Advanced

“Candidates can provide a lot of detailed information about their work, their ministry and its mandate/programs/services. They can narrate, describe an incident or event in the three major simple time frames (past, present and future) and they are able to handle an unusual or unexpected request or transaction without complications... At this level, candidates are comfortable in familiar and routine work settings ... they speak coherently linking sentences together smoothly so that their response is more than just a collection of sentences. They can verbalize their ideas and thoughts clearly and logically. Their command of the past, present and future verb tenses allow them to relate events in good chronological order. They exhibit a good, although not perfect, command of the most frequently used grammatical constructions. Errors that do not hamper communication are acceptable. Their range of everyday work vocabulary allows them to speak comfortably about all routine and concrete aspects of their job. They are able to paraphrase or use a synonym when lacking technical or specialized terms. They can provide clear instructions and explain the steps involved in following a procedure or carrying out a task. They ... have the language skills to present ideas in clear and logical order so that the listener understands exactly what to do. They can cope with an unusual request, a misunderstanding, or an unexpected turn of events. They exhibit the fluency of someone who is comfortable with the language. Pronunciation ... never hampers communication although some influence from their first language is likely.

Advanced level speakers are understood by people who are unaccustomed to dealing with non-native speakers.

Oral proficiency – Advanced plus

At the Advanced Plus level, knowledge of general and work-related vocabulary broadens substantially and candidates demonstrate some control of complex structures. They have the ability to speak at length about a topic using both main and supporting arguments. They can provide detailed explanations, descriptions and narration. They can discuss a wide range of interests and explain their special fields of expertise and competence. They can sometimes present arguments or points of view and propose and defend a course of action quite effectively. However, the quality and precision of language required to accomplish these tasks are not sustained, and patterns of errors and/or language breakdown will typically occur.

Oral proficiency – Superior

Speakers can discuss in depth a wide range of professional topics. They can explain complex or technical processes and procedures in their area of expertise. They can effectively express their views, present arguments, negotiate, persuade and advise. They can also linguistically manage an unfamiliar situation easily and deal with a complex transaction. Superior level speakers can participate effectively and accurately in most formal and informal conversations on a wide range of professional topics. Their speech is highly organized and they have the grammar and vocabulary (including very precise and abstract terms) to verbalize ideas and thoughts ranging from the concrete to the conceptual. They can discuss in depth a variety of job tasks or topics using more impersonal or formal language when appropriate. They can present arguments convincingly and explain complicated or technical procedures and processes. They possess the language skills to negotiate, advise, persuade and propose and defend a course of action effectively. Sporadic errors may occur, but they never disturb the native speaker.

Written proficiency – Advanced

Candidates are able to join sentences in simple discourse of several paragraphs in length. They can handle uncomplicated, routine business correspondence as well as narrative and descriptions of a factual nature on familiar topics. Revision by a superior writer is necessary. They demonstrate sufficient control of the most frequently used structures so that their message is understood. Errors in complex grammar are common. Vocabulary relates to the everyday, concrete (not conceptual) aspects of their work. Their text displays a logical sequencing of ideas and their message will be clear to readers unaccustomed to the writing of non-natives.

Written proficiency – Advanced plus

Candidates are able to write with greater precision and detail than at the Advanced level. They can handle routine business correspondence with greater accuracy and they begin to use appropriate tone. A broader knowledge of general and work-related vocabulary enables them to expand into topics relating to special fields of competence. They can express views and present arguments quite effectively. However, the quality and precision of the message conveyed will be lacking at times and may result in occasional miscommunication.

Written proficiency – Superior

Superior writers are able to express themselves effectively in most formal and informal writing on practical, social and professional topics. They exhibit good control of a full range of structures, as well as a broad knowledge of general and work-related vocabulary. Good chronological and logical ordering, cause and effect, comparison and thematic development are evident. They can present arguments and handle concepts effectively. Sporadic errors may occur, but never cause miscommunication.”

Reading proficiency is not described.

CITY OF OTTAWA

The City of Ottawa proficiency level structure describes its most advanced level as follows:

Oral expression - Advanced 3

“Uses complex structures in all types of situations with occasional errors.

Oral comprehension – Advanced 3

Understands all rates of speech and slurred speech on general and work-related topics in all situations.

Reading comprehension – Advanced 3

Reads and understands all types of texts.

Written expression – Advanced 3

Writes general and work-related texts with few errors”.

CANADIAN LANGUAGE BENCHMARKS

The global performance descriptors of the Canadian Language Benchmarks read as follows for their highest level of proficiency:

Speaking – Benchmark 12

“Learner can create and co-create oral discourse, formal and informal, general or technical, in own field of study or work, in a broad range of complex situations. Satisfies most academic and work-related expectations for competent communication. Can deliver public presentations to audiences. Can lead formal group discussions, meetings and workshops. Can communicate to explain complex ideas to diverse groups, to debate arguments on complex matters, to teach, to negotiate, and to resolve conflict in a variety of situations. Discourse is fluent and “natural” (native-like in phrasing). Language is complex, accurate and flexible in the manipulation of structure of information in clauses to express emphasis, comment, attitude. Content, organization, format, delivery, tone and conversational style of discussion or presentation are appropriate to purpose and audience.

Listening – Benchmark 12

Learner can competently and fluently interpret all spoken discourse, formal and informal, general and technical, in own field or work, in a broad variety of demanding contexts, live and audio/video recorded. Can follow long stretches of oral discourse, monologic or multi-speaker exchanges, with complex abstract and conceptual language, to obtain complex, detailed, and specialized information, Canadian cultural references; figurative, symbolic and idiomatic language; irony; sarcasm; verbal humour. Can critically evaluate most aspects of oral discourse. Has adequate listening/interpreting skills to satisfy all academic and work-related expectations for competent understanding of communication.

Reading – Benchmark 12

Learner can read a full variety of general and literary texts, and specialized or technical (academic and professional) texts in own field. Can get information, ideas and opinions on familiar and unfamiliar abstract and conceptual topics from propositionally, linguistically, stylistically, and culturally complex texts in demanding contexts of study and work. Can read critically and with appreciation for aesthetic qualities of text, register, stylistic and rhetorical nuance, tone (i.e. humour, irony, sarcasm), genre awareness, writer's bias, and point of view. Can understand almost all idiomatic and figurative language, and socio-cultural references. Can search through complex displays of information, and use high-level inferences, extensive background, and specialized knowledge to locate and integrate multiple pieces of abstract information across various multiple complex and dense texts. Can interpret, compare, and evaluate both the content and the form of written text. Reads fluently and accurately, adjusting speed and strategies to task.

Writing – Benchmark 12

Learner can write complex, original, formal texts needed for very complex, technical or specialized tasks in demanding contexts of language use. Texts are often for public consumption and for various purposes: reporting, projecting, evaluating, promoting, expounding an argument, or appealing to an unfamiliar audience. Can synthesize and evaluate extensive complex information and ideas from multiple sources as a coherent whole (i.e. as an evaluative report of desired length and detail level). Can take notes/write minutes and commentaries at symposia, public/stakeholders' meetings, and consultations on complex topics. Can write highly specialized, complex, external correspondence, proposals, news releases. Formal and public reports, publication reviews. Can write highly specialized internal documents (i.e. procedures, policy manuals). Can develop innovative complex forms and other formats of information display. Can produce effective and stylistically polished essays, documents, articles, theses (over 20 double-spaced typed pages or other volume, appropriate to purpose and audience). Demonstrates excellent control over grammar, vocabulary, cohesion, thematization, organization, tone and style; errors are rare and minimal. Can effectively proofread, revise, and edit all aspects of texts, using own resources.”

ACTFL PROFICIENCY GUIDELINES

ACTFL proficiency guidelines generic descriptions delineate superior and distinguished levels as their highest levels on their scale.

Speaking - Superior Level (1999)

“Speakers at the Superior level are able to communicate in the language with accuracy and fluency in order to participate fully and effectively in conversations on a variety of topics in formal and informal settings from both concrete and abstract perspectives. They discuss their interests and special fields of competence, explain complex matters in detail, and provide lengthy and coherent narrations, all with ease, fluency and accuracy. They explain their opinions on a number of topics of importance to them, such as social and political issues, and provide structured argument to support their opinions. They are able to construct and develop hypotheses to explore alternative possibilities. When appropriate, they use extended discourse without unnaturally lengthy hesitation to make their point, even when engaged in abstract elaborations. Such discourse, while coherent, may still be influenced by the Superior speakers own language patterns, rather than those of the target language.

Superior speakers command a variety of interactive and discourse strategies, such as turn-taking and separating main ideas from supporting information through the use of syntactic and lexical devices, as well as intonational features such as pitch, stress and tone. They demonstrate virtually no pattern of error in the use of basic structures. However, they may make sporadic errors, particularly in low-frequency structures and in some complex high-frequency structures more common to formal speech and writing. Such errors, if they do occur, do not distract the native interlocutor or interfere with communication.

Listening - Superior level

Able to understand the main ideas of all speech in a standard dialect, including technical discussion in a field of specialization. Can follow the essentials of extended discourse which is propositionally and linguistically complex, as in academic/professional settings, in lectures, speeches, and reports. Listener shows some appreciation of aesthetic norms of target language, of idioms, colloquialisms, and register shifting. Able to make inferences within the cultural framework of the target language. Understanding is aided by an awareness of the underlying organizational structure of the oral test and includes sensitivity for its social and cultural references and its affective overtones. Rarely misunderstands but may not understand excessively rapid, highly colloquial speech or speech that has strong cultural references.

Listening - Distinguished level

Able to understand all forms and styles of speech pertinent to personal, social and professional needs tailored to different audiences. Shows strong sensitivity to social and cultural references and aesthetic norms by processing language from within the cultural framework. Texts include theatre plays, screen productions, editorials, symposia, academic debates, public policy statements, literary readings, and most jokes and puns. May have difficulty with some dialects and slang.

These guidelines assume all reading texts to be authentic and legible.

Reading - Superior level

Able to read with almost complete comprehension and at normal speed expository prose on unfamiliar subjects and a variety of literary texts. Reading ability is not dependent on subject matter knowledge, although the reader is not expected to comprehend thoroughly texts which are highly dependent on knowledge of the target culture. Reads easily for pleasure. Superior-level texts feature hypotheses, argumentation and supported opinions and include grammatical patterns and vocabulary ordinarily encountered in academic/professional reading. At this level, due to the control of general vocabulary and structure, the reader is almost always able to match the meanings derived from extra-linguistic knowledge with meanings derived from knowledge of the language, allowing for smooth and efficient reading of diverse texts. Occasional misunderstandings may still occur; for example, the reader may experience some difficulty with unusually complex structures and low-frequency idioms. At the Superior level the reader can match strategies, top-down or bottom-up, which are most appropriate to the text. (Top-down strategies rely on real-world knowledge and prediction based on genre and organizational scheme of the text. Bottom-up strategies rely on actual linguistic knowledge.) Material at this level will include a variety of literary texts, editorials, correspondence, general reports and technical material in professional fields. Rereading is rarely necessary, and misreading is rare.

Reading - Distinguished level

Able to read fluently and accurately most styles and forms of the language pertinent to academic and professional needs. Able to relate inferences in the text to real-world knowledge and understand almost all socio-linguistic and cultural references by processing language from within the cultural framework. Able to understand a writer's use of nuance and subtlety. Can readily follow unpredictable turns of thought and author intent in such materials as sophisticated editorials, specialized journal articles, and literary texts such as novels, plays, poems, as well as in any subject matter area directed to the general reader.

Writing - Superior level (2001)

Writers at the Superior level are able to produce most kinds of formal and informal correspondence, complex summaries, précis, reports and research papers on a variety of practical, social, academic, or professional topics treated both abstractly and concretely. They use a variety of sentence structures, syntax, and vocabulary to direct their writing to specific audiences, and they demonstrate an ability to alter style, tone and format according to the specific requirements of the discourse. These writers demonstrate a strong awareness of writing for the other and not for the self.

Writers at the Superior level demonstrate the ability to explain complex matters, provide detailed narrations in all time frame and aspects, present and support opinions by developing cogent arguments and hypotheses. They can organize and prioritize ideas and maintain the thrust of a topic through convincing structure and lexicon and skilful use of writing protocols, especially those that differ from oral protocols, to convey to the reader what is significant. Their writing is characterized by smooth transitions between subtopics and clear distinctions made between principal and secondary ideas. The relationship among ideas is consistently clear, evidencing organizational and developmental principles such as cause and effect, comparison, chronology, or other orderings appropriate to the target language culture. These writers are capable of extended treatment of a topic which typically requires at least a series of paragraphs but can encompass a number of pages.

Writers at the Superior level demonstrate a high degree of control of grammar and syntax, both general, and specialized/professional vocabulary, spelling or symbol production, cohesive devices, and punctuation. Their vocabulary is precise and varied with textured use of synonyms, instead of mere repetition of key words and phrases. Their writing expresses subtlety and nuances and is at times provocative. Their fluency eases the reader's task.

Writers at the baseline of the Superior level will not demonstrate the full range of the functional abilities of educated native writers. For example, their writing may not totally reflect target language cultural, organizational, syntactic, or stylistic patterns. At the baseline Superior level, occasional errors may occur, particularly in low-frequency structures, but there is no pattern. Errors do not interfere with comprehension and they rarely distract the native reader.

COMMON EUROPEAN FRAMEWORK

The highest level of proficiency of the Common European Framework illustrate its scales as follows:

Oral Production – C2

“Can produce clear, smoothly flowing well-structured speech with an effective logical structure which helps the recipient to notice and remember significant points.

Listening comprehension – C2

Has no difficulty in understanding any kind of spoken language, whether live or broadcast, delivered at fast native speed.

Reading Comprehension – C2

Can understand and interpret critically virtually all forms of the written language including abstract, structurally complex, or highly colloquial and non-literary writing. Can understand a wide range of long and complex texts, appreciating subtle distinctions of style and implicit as well as explicit meaning.

Written production – C2

Can write clear, smoothly flowing, complex texts in an appropriate and effective style and a logical structure which helps the reader to find significant points.”

EDUCATIONAL TESTING SERVICE (TFI)

We are considering here the scale used for le Test de français international (TFI). Based on six levels with a numerical scale going from 5 to 125 for beginning proficiency to 455 to 495 for professional proficiency in each of two categories: listening section (includes speaking and listening) and reading section (includes reading and writing), the TFI’s highest level describes the four skills as follows:

Speaking – 455-495

“Can conduct meetings with native speakers of French; performs all of the lower levels (beginning proficiency, elementary proficiency, advanced elementary proficiency, basic working proficiency and working proficiency) with a greater degree of ease....

Listening – 455-495

Can understand native speakers of French in meetings; functions in all of the situations described in the lower levels whether professional or social, concerning concrete or abstract subjects.

Reading – 455-495

Can read without difficulty for most professional needs; read highly technical manuals in own area; perform all of the reading tasks of the lower levels.

Writing – 455-495

Can write effectively, both formally and informally; however, work for publication will still require review; produce the documents described in the lower levels without undue difficulty.”

ASSESSMENT TOOLS

The limited timeframe of the project constrained our research to a few well-known systems and tools in use in Canada, the United States and Europe. Namely, we examined the Public Service Commission of Canada’s assessment tool; the Government of Ontario’s tool; the City of Ottawa tests; the widely-used Oral Proficiency Interview; the Test de connaissance du français; the Test de français international and the Test d’évaluation du français..

Within the limit of available information, we looked at test purpose; test design and theoretical foundation for the test; reliability and validity; test administration, test components, scoring, administration time and cost, delineating 14 criteria to form the basis of a checklist applied to every test.

GOVERNMENT OF CANADA – PUBLIC SERVICE COMMISSION TEST

Purpose

In the Public Service, all jobs requiring both official languages are termed “bilingual” positions. The Commission developed a Second Language Evaluation System, composed of three tests (oral interaction, reading and writing) and second language standards to determine whether candidates meet the second language requirements of these positions.

Background

In 1969, following through on recommendations made in the report of the Royal Commission of Inquiry on Bilingualism and Biculturalism, Canada’s Parliament passed

the first Official Languages Act, which recognized French and English as the official languages of all federal institutions in Canada. This led to the design of a complex system of language designation of public service positions, provision of French language training programs and assessment to measure the civil servants' proficiency in their second language.

Test design and theoretical foundation

R.A.C. Goodison once stated: "In government offices, institutional memories are often short or non-existent" (1990:2). Government workers feel more compelled to write internal memos, manuals and reports than to publish scholarly articles. Internal papers tend to disappear with personnel turnover." This is what seemed to have happened for the Government of Canada and the Government of Ontario for which the rationale and theoretical foundation for their test design seem to no longer exist.

Despite these shortcomings, we know for sure that second language standards (content and performance)¹ had to be designed. Managers prepared lists of language tasks in each skill (oral interaction, reading and writing) required to carry out the duties and responsibilities of specific jobs. They outlined communicative tasks and the contexts in which they take place as well as describing linguistic proficiency that reflects the accuracy levels and other requirements for employees in bilingual positions to function effectively. They included details such as affecting language performance (i.e. context: informal meeting), performance conditions (i.e. face-to-face or telephone), and constraints (i.e. may make many errors...). These standards were developed from lists prepared by managers.

The rationale behind the choice of levels of proficiency that were to be used nationally seems to have disappeared. Whatever the reason, one fact remains: These standards are divided into three levels: A (lowest) to C (highest). Each level is sub-divided by skill (oral interaction, reading, writing). These second language standards form the link between the requirements for bilingual positions and the three tests: jobs are classified according to these three levels and results on the tests are expressed in terms of these same proficiency levels. This allows a clear understanding of what is required for designated positions by both candidates and employers.

Test components

Second language proficiency is divided into three components: oral interaction, which encompasses, speaking and listening, reading and writing proficiency.

Oral interaction

The oral interaction test consists of three phases: a warm-up during which the candidate is not assessed. This phase is an adaptation phase for the candidate to become more comfortable with speaking their second language, adjusting to the examiner and environment.

The second phase or interview consists of recorded question-and-answer exchanges and one or more dialogue situations on work-related topics. The candidates are called upon to talk about their past and present work experiences or studies, they may be asked to give explanations or express and support an opinion. The examiner also leads a dialogue in a simulated situation the candidates might encounter in their work.

The last phase, or wind down, brings the test to a close. Candidates are not assessed during the wind down.

This phase can be conducted face to face or over the phone.

The criteria used in assessing the test are fluency, grammar, vocabulary and pronunciation.

Specific topics vary but the subject matter remains work-related.

The interview lasts thirty minutes.

Reading test

The reading test is composed of 65 multiple-choice questions worth one point each. A candidate's proficiency level is determined by his/her test score. Candidates can also take that portion of the test online.

The reading test lasts two hours.

Writing test

The writing test is comprised of 55 multiple-choice questions worth one point each. A candidate's proficiency level is determined by his/her test score. Candidates can also take that portion of the test online.

The writing test lasts an hour and a half.

Reliability and validity

Reliability

Whenever subjectivity enters in an equation, reliability becomes more difficult to maintain.

Because the oral interaction part of the test is scored by assessors, however well trained, there is always a risk of bias or divergence of opinion in the scoring of one or more of the criteria. Indeed, it has been admitted by assessors that some are stricter than others.

Consequently, the same candidate taking the test with different assessors might obtain a different score.

This risk disappears with the modality of the reading and writing tests which are only multiple-choice.

Validity

When one wants to measure speaking and listening, one has to place candidates in situations where they have to speak and listen. Furthermore, the Public Service

Commission assesses its candidates to know which level of proficiency they have attained as per their position.

The oral interaction test demonstrates validity as it strives to entice candidates to speak about their work and place them in realistic situations candidates could encounter.

The listening or oral comprehension of the component of the interaction is assessed throughout the conversation as candidates answer the questions they have been asked or interact with the assessor on a situation they have been given.

The reading component of the test incorporates overall and detailed comprehension through a variety of tasks and does measure reading comprehension.

The writing component of the test is based on translation and grammatical accuracy recognition through multiple choice. Candidates are not required to write anything. It can therefore hardly be called a writing test. It does strive to check candidates' syntactical and lexical knowledge. It cannot determine candidate's ability to write in their second language since no output is required, therefore jeopardizing the validity criterion of the test.

Scoring

The scoring range for the oral interaction was not made available.

Level A reading ranges from 20 to 37 inclusive; level B from 38 to 50; level C from 51 to 58. A higher score earns the candidate an exemption from future testing in reading.

Level A writing ranges from 18 to 30 inclusive; level B from 32 to 43; level C from 44 to 50. A score of 51 or higher earns the candidate an exemption from future testing in writing.

Multiple choice items are easy and quick to score so the scorability for the reading and writing parts rates high.

It is however more difficult to score the oral interaction. However, since assessors are trained in the conduct of these tests, scorability does not represent a major problem except that these tests require a trained human assessor.

Test administration

If we favour the use of technology to rate tests, the test's oral interaction component does not rate highly in practicality since it requires a human to conduct the assessment.

However, it remains difficult to proceed with a reliable interactive tool without having the candidate speak in a realistic situation when testing for hiring purposes.

Administration time

The duration of the test is short enough to require a minimal time commitment for the assessment itself. The scoring of the test takes an additional thirty minutes if the tape has to be listened to, which means the time commitment of the assessor is about an hour per test, a reasonable amount of time considering the stakes of the result.

Cost

The Public Service Commission's assessment tool is not available to the general public. It is only intended for public service employees or candidates to a position and is not marketed.

GOVERNMENT OF ONTARIO

Overview

Attempts at contacting provincial civil servants have been made to understand the theoretical foundation and choice of the provincial second language testing system. The persons contacted could not offer any justification or refer us to someone who could.

Background

With the onset of the *French Language Services Act, 1986 (FLSA)*, French-speaking Ontarians have the right to receive provincial government services in French from central and head offices as well as offices located in or serving designated areas. Managers of program areas with obligations under the *FLSA* must ensure that their program areas have a sufficient number of designated positions filled with bilingual staff to ensure that French-speaking Ontarians have access to services of equal quality to those offered in English. The government then had to establish a language proficiency system and use an assessment tool that could indicate if their staff was bilingual or not.

An office was created with the goal of assessing provincial civil servants' French as a second language skills. Its staff decided to resort to a system in existence in the United States: the American Council for the Teaching of Foreign Languages' rating scale which goes from novice low to superior or even distinguished. It also decided to use its assessment tool: the Oral Proficiency Interview to rate verbal skills.

For all intents and purposes, it was also determined that the Advanced level and at times, the Superior level would qualify employees for a bilingual position.

Test design and theoretical foundation for test

There was no specific design or theoretical foundation for this system since it was adapted from the ACTFL Oral Proficiency Interview.

Test components

Oral skills

Oral proficiency is assessed through a 6-step interview during which the test taker is led through a warm up period, a series of questions, descriptions, narrations and detailed explanations, opinion giving and a situation. The latter is accomplished through the giving of a coloured card depending on the level the assessor has determined through the first five steps. The colour of the card changes depending on the level the assessor has

determined during the conversation. This stems directly from the ACTFL proficiency interview.

Reading test

No information was found on this component.

Writing test

The candidate is given a writing task related to what he/she might be called upon to accomplish in their work. The length of the assignment varies from a minimum of 15 lines to a minimum of 2 pages.

Reliability and validity

Reliability

Reliability might be akin to that of the Oral Proficiency Interview which uses the ACTFL proficiency guidelines but this cannot be ascertained due to the scarcity of available information. Neither provincial civil servants nor present government of Ontario assessors were forthcoming with information on test design, scales, assessment tool. When one reads the Assessor's manual, one cannot help but be struck with the high probability of lack of inter-rater reliability. The level of training of provincial assessors is also in doubt and therefore renders an informed opinion on reliability factors impossible.

Validity

We know and we will see later that the OPI is weak in content validity. When it comes to the provincial government tool, face, construct and predictive validity are also in doubt. Confidence is greatest in the native speaker performance criterion of this tool. The assessment of listening skills is wrapped up in the interview and remains superficial. While some might think this is justifiable in general testing, it becomes more difficult to defend for occupationally-related testing. Unless there is a specific correction grid, which we have not seen, to correct the writing task, inter-rater reliability is not likely. However, this test has the merit of actually testing writing as opposed to structure or error recognition, which often passes for a writing test.

Scoring

The criteria used in the oral interview are vocabulary, discourse, grammar, fluency and pronunciation. Special consideration is given for the situation. Short descriptions accompany the five identified levels: no proficiency; survival competency; limited competency; functional competency and professional competency.

Test administration

The oral assessment requires the presence of an assessor. It can be conducted face to face or through the phone.

Administration time

The oral interview takes about twenty minutes.

Cost

This information was not sought through client ministries of the government of Ontario but through private suppliers. It was however not available through the private suppliers. One supplier quickly hung up after a few introductory questions; the others never returned calls.

CITY OF OTTAWA

Overview

The City of Ottawa provides a broad base of services in the areas of environment, culture and recreation, homes for the aged, planning, public health, social services, policing and transportation to more than 750,000 residents of Ottawa. It boasts a workforce of approximately 17,000 employees.

Background

In 2001, twelve former municipalities were joined through amalgamation to make the new City of Ottawa. In May of that year, a Bilingualism Policy was adopted. It was the opportunity for the City to set up a system of language proficiency levels and tests. In 2003, the Human Resources Department and French Language Services Division of the Corporate Service Department of the City of Ottawa identified a requirement for the development of a functionally defined language proficiency level structure for both the English and French languages.

The consultant hired for that project was to develop a language proficiency level structure with clear, functional descriptions indicating the distinction between levels of proficiency, as a first step.

Test design and theoretical foundation for test

The former City of Ottawa had used a 0 (Beginners) to 5 (Superior) scale. Now, the new City wanted an integrated approach for its second language training program. The proficiency structure would also form the basis for linguistic profiles (the determination of language proficiency for employees) and for the design of occupationally- related assessment tools.

The desire to link the structure to language training is crucial for the comprehension of the design. There was to be no mention made of native level proficiency as the structure and ensuing tests were going to be used to place people in the proper level of language training classes as well as being, eventually and quite secondarily, used for hiring purposes.

The original proposal ranged from levels 1 (no proficiency) to 6+ (native-like proficiency) – a twelve (12) level range² but the language proficiency level structure which had to be designed was not to reflect the broad proficiency people can demonstrate in using language but strictly reflect the result of language training instead. Thus, there was to be no beginner or superior level of proficiency and the chosen scale high-end level was to reflect only the top level reached within a language training program (an advanced level) as opposed to native-like fluency, which may be found in society.

A twelve-level range was deemed too broad by the customer and thus, the scale was reduced to a more manageable nine (9) level range³ from Elementary 1 to Advanced 3. The City staff rationale for that decision was that these levels would be better understood by staff and corresponded to the language training levels offered on City premises for City personnel.

Test components

Four skills were examined: oral expression, oral comprehension treated as separate entities, reading comprehension and written expression.

Oral interaction

Four criteria were chosen to assess verbal skills: pronunciation, grammar, vocabulary and fluency. Socio-linguistic competence, originally part of the proposal, was discarded as deemed irrelevant by City Language Training Centre staff.

Oral comprehension was to be scored within an oral proficiency interview conducted to assess second language skills. Reading and writing were assessed separately but still using an Elementary to Advanced 3 scale.

Oral expression and oral comprehension are assessed through a forty minute face to face recorded interview during which the assessor asks increasingly difficult questions from a linguistic competence standpoint. The first part of the interview is composed of general questions aimed at both relaxing candidates, causing them to talk about their own life (without invading their privacy), moves onto general questions and intersperses work-related questions, first on the reason why the testee chose the profession he/she is being assessed for, what their main professional activities are (if they are already in a similar position), what aspect of their work they like best or least, what challenges they face in their work, which direction their occupation is taking, what changes they anticipate etc. The interview moves then onto an aural comprehension component, whereby candidates listen to pre-recorded occupationally-related out of context sentences. The candidate then reformulates what he/she has heard in the other official language with as many details as possible. To avoid the memory aspect of that component, sentences are kept short enough for memory not to come into play. After this admittedly more complex exercise, the candidate is called upon to play his/her role as a professional while the assessor

proposes a role play in which he/she plays the role of a client, patient, student depending on the testee's work milieu.

This phase is usually conducted face to face. It could also be conducted over the phone. Specific topics vary but the subject matter is heavily slanted towards work-related situations.

Reading test

The reading test is composed of a text followed by 10 open-ended related questions worth one point each or of a lesser number of questions worth more than one point. In all cases, the maximum number of points is 10.

There is a different text for each profession category and the tester chooses the most relevant text for the candidate.

A candidate's reading comprehension is determined by his/her test score.

The reading test lasts forty-five minutes maximum.

Writing test

The writing test is made up of a request for a 350 to 400 words on a topic that is provided by the assessor and relates to the candidate's position category.

Candidates are allowed to use an English-French dictionary, a French dictionary and the Bescherelle "L'Art de conjuguer" (for the French language tests).

Writing expression is rated on an Elementary 1 to Advanced 3 scale.

The writing part of the test lasts forty-five minutes maximum.

Reliability and validity

Reliability

The verbal skills assessment demands the presence of a human assessor. Inter-rater reliability is consequently more difficult to achieve 100 % of the time as bias or divergence of opinion in the scoring of one or more criteria can come into play.

Intra rater reliability could also vary depending on the candidate's condition and familiarity with the tool.

The reading test rates high in as much as the candidate does not become too familiar with this component. To date, one text has been designed per position category. In case of re-test, the candidate's memory might come into play.

The writing test rates highly on inter-rater reliability (the assessors are provided with a correction grid with the number of points to be deducted per error type) and they are all linguists whose first language is French and have already demonstrated high level proficiency in that aspect of the test.

Validity

The oral expression and oral comprehension components of the test rate highly in face, content, construct and predictive validity as they look like a reasonable test to measure what is expected and do test oral expression and comprehension, are easy to match with the level descriptions and do predict what a candidate is going to be able to do in his/her work milieu.

Scoring

Oral expression, oral comprehension, reading comprehension and written expression range from Elementary 1 to Advanced 3.

In oral expression, elementary levels 1 to 3 range from 0 to 18; Intermediate levels 1 to 3 range from 19 to 33 and Advanced levels 1 to 3 range from 34 to 45.

In oral comprehension, elementary levels range from 0 to 9; Intermediate levels range from 10 to 18 and Advanced levels range from 19 to 25.

Elementary level reading ranges from 1 to 4.9; Intermediate 1 to 3 ranges from 5 to 7.9 and Advanced 1 to 3 ranges from 8 to 10.

Written expression levels range from 0 to 14 for Elementary levels; from 15 to 29 for Intermediate levels and from 30 to 45 for Advanced levels.

Scoring for all four components require manual scoring.

Reading test items are easy to correct as the assessors are provided with the answer key and the number of points for each answer. They then correlate the result to the scoring grid and find the level of proficiency.

Written expression is a little more difficult to correct. Candidates are asked to write legibly and cleanly. However, assessors may still be faced with legibility issues.

Oral expression and oral comprehension require familiarity with the scoring grid.

However, since assessors are trained in the conduct of these tests, scorability does not represent a major problem except that these tests require a trained human bilingual assessor.

Test administration

The test's oral expression and comprehension components do not rate highly in practicality as they require a human to conduct the assessment.

Reading comprehension and written expression require minimal intervention from the tester. The testee can work independently in a testing room with no telephone available.

Administration time

The duration of the test is short enough to require a minimal time commitment for the oral assessment itself. The scoring of the test takes an additional forty minutes which means the time commitment of the assessor is eighty minutes per test, a reasonable amount of time considering the stakes of the result.

Cost

The City of Ottawa assessment tools belong to the City and are not marketed.

CANADIAN LANGUAGE BENCHMARKS (NIVEAUX DE COMPÉTENCE LINGUISTIQUE CANADIENS)

While there are no French language proficiency tests on the market developed as per the Canadian Language Benchmarks, careful consideration needs to be given to the Niveaux de compétence linguistique canadiens for their importance in the second language world.

The French as a second language Canadian benchmarks were issued in September 2006. They are intended for adult immigrants learning French as a second language and for teachers, program administrators, employers etc. who work with immigrants. They are aimed at defining and determining the language skills required to facilitate and accelerate integration into Canadian society. They describe language proficiency over 12 levels in each of the four skills. They can be used as a common reference framework for learning, teaching and assessing French as a second language. They are not a description of essential knowledge of the language; they do not provide instructions to design curricula; they do not constitute a second language teaching method. They are not an assessment tool either.

The Canadian Language Benchmarks authors indicate that employers can refer to the benchmarks to better understand communication tasks linked to various positions, to be better informed on the level of mastery demanded for some occupations and to better understand what the adult immigrant can accomplish in French. They will assuredly become, in their French version, in the French as a second language field, the reference they already are in their English version in the English as a second language world.

AMERICAN COUNCIL FOR THE TEACHING OF FOREIGN LANGUAGES/ORAL PROFICIENCY INTERVIEW (ACTFL/OPI)

Overview

The OPI is used to predict a speaker's performance in a situation where a particular level of functional language use is required, such as a job, a language class, or living in a country where the language is spoken. It is used by government agencies, academic institutions, and in the private sector. It can be used for diagnostic, placement, evaluation and research purposes.

Background

In the 1950's, Frank Rice and Claudia P. Wilds devised rating scale and proficiency test, based on the concept of the structured interview. In the 1970's, the American Foreign

Service Institute entered into close relationship with the Interagency Language Roundtable and Georgetown University, the Center for Applied Linguistics and the Educational Testing Service, which eventually took part in developing the ACTFL/ETS (American Council for the Teaching of Foreign Languages/Educational Testing Service) proficiency guidelines widely used in schools and colleges. From there, the scope of the system extended to companies in the United States and abroad. To date, the Oral Proficiency Interview (OPI) is a well-known and recognized model.

Test design and theoretical foundation

The ACTFL Oral Proficiency Interview (OPI) is a standardized instrument since, to assure reliability in assessing different speech samples, a prescribed procedure must be observed. The OPI does not strive to measure when, where, why, and under what conditions speakers learned the language. “The OPI is not an achievement test assessing a speaker’s acquisition of various specific aspects of course and curriculum content. The OPI assesses language performance in terms of the ability to use the language effectively and appropriately in real-life situations.”

The rating scale has been designed to assess oral proficiency globally. It seeks to determine the consistent level of use and upper limitations of a speaker’s ability to communicate in his/her second language.

The rating scale ranges from Novice to Superior with descriptors for novice-low, novice-mid, novice high, intermediate-low, intermediate mid, intermediate high, advanced low, advanced mid, advanced high and superior; distinguished levels in a few cases.

The rating scale is built upon a hierarchy of global tasks—asking and answering simple questions, narrating, describing, etc; contexts or set of circumstances—linguistic or situational—in which these tasks are performed, and content areas or topics that relate to these contexts. To assess the accuracy with which the tasks are performed, the following factors are taken into account: grammar, vocabulary, and pronunciation. In addition, OPI assessment recognizes the importance of fluency—rate of delivery and coherence of message—as well as the sociolinguistic appropriateness or acceptability of what is being said within a certain setting, and use of appropriate strategies, such as indicating the intention to continue speaking, for discourse management. This implies the use of discrete words and phrases, sentences, paragraphs, or extended discourse.

Components

Oral interaction

A trained interviewer meets the interviewee in a one-on-one situation. The recorded interview involves a mandatory structure of four phases—warm-up, level checks, probes and wind-down. The OPI should resemble, to the greatest extent possible, a natural conversation.

The structure of the OPI is standardized, but since the interview is based on as natural a conversation as possible between the two conversational partners, its content is unique to each interview and to the interviewee since his or her responses reflect individual background, life experiences, interests and opinions. In this process, the interviewer’s

line of questioning and task-posing is determined by the responses of the interviewee. Although there are standard question types relative to proficiency level, the specific content of the OPI is determined in large part through conversational negotiation, depending on information offered in response to the interviewer's lines of questioning and the tasks posed.

Reading test

The reading proficiency guidelines are to be applied to reading assessment tools test designers will develop but there is no actual reading test designed by ACTFL.

Writing test

The ACTFL Business Writing Test is advertised as being a standardized test for the global assessment of functional business writing ability. It has 5 requests for a written response dealing with general business situations.

Reliability and validity

Reliability

While the OPI training guide states that it is important that the OPI evidence test/re-test reliability in that a speaker tested two or more times will be assigned the same rating in the re-test as in the initial test, the aspect of reliability depends on the tester.

Concurrently, inter-rater reliability is not guaranteed as two testers listening to the same OPI should assign it the same rating but may have differing opinions of the importance and hence rating of mistakes, however well trained and certified.

The OPI is a criterion-referenced, rather than a norm-referenced assessment. Each speech sample is rated solely according to the criteria of the rating scale rather than being compared to performances of other speakers. Because of the global nature of the assessment procedure, there will be a variety of individual performances within the same rating level. However, each individual performance must evidence certain required features to be rated at a given level.

Although the OPI is not a fixed series of questions, the prescribed structure targeting the same global criteria in each OPI ensures comparability from one test to another. Specific tasks vary from OPI to OPI, but the types of tasks posed remain the same. It is, in fact, a critical feature of the design of the OPI that the specific questions vary from interview to interview; neither interviewers nor speakers can prepare for the OPI in the traditional sense.

Validity

When it comes to validity, we can state that the OPI demonstrates face, content, construct, and predictive validity. The four assessment criteria of global tasks, formal and informal contexts, content areas, accuracy features and text type are designed to predict

the level of consistent functional ability in other real-life situations of test takers. The OPI also meets the native-speaker performance criterion.

It is however weak in content validity as it does not treat listening as a specific skill. While this is justifiable in general testing, it becomes more difficult to do so for occupationally-related testing.

Scoring

There are 4 categories of criteria: global tasks or functions performed with the language; social contexts and content areas in which the language is used; accuracy features which define how well the speaker performs the tasks pertinent to those contexts and content areas, and the oral text types—from individual words to extended discourse—produced. The assessor listens to an audiotape of the entire interview before assigning a final rating. Features of the speech sample are compared to the criteria for each major level (Novice, Intermediate, Advanced, Superior) of the rating scale, and then assigned a sublevel rating (Low, Mid, High) by carefully comparing the sample with the appropriate sublevel descriptions in the ACTFL Guidelines.

Test administration

The OPI requires the presence of a trained and certified interviewer and the use of a tape recorder. It can be done face to face or over the phone. The writing test can be done via test booklet or internet but somebody must stay in the testing room to ensure that candidate does not use any resources (dictionaries, grammar texts, notes) or call for assistance.

Administration time

The OPI takes the form of a 10- to 30-minute conversation between a trained interviewer and the interviewee. It is both time and labor intensive. Its applications may be limited when dealing with larger numbers of speakers. The writing test takes 80 minutes.

Cost

The Oral Proficiency Interview costs \$ 129.00 if conducted over the phone; the ACTFL Writing Proficiency Test costs \$ 50.00.

TEST DE CONNAISSANCE DU FRANÇAIS (TCF)

Overview

TCF is the French Education Ministry's test of French language ability. It is a language placement test for non-native speakers of French who wish to obtain a simple, effective and precise evaluation of their general language skills for professional or personal reasons.

Background

TCF indicates it is structured as a comprehensive test of French language ability and that it is fully standardised and calibrated since the test is built around a rigorous test methodology in order to provide a precise grading of proficiency. By standardised, the web site means a test which is administered and marked according to a uniform standard. By calibrated, the web site means that all the items are statistically graded according to a common difficulty scale in order to provide a precise benchmark of “ability in French!”. It states that it provides an accurate benchmark of the candidate’s language ability using the six-level proficiency scale of the Council of Europe.

Test design and theoretical foundation

The TCF relies on the Common European Framework, which is now the basis for testing in Europe. Its theoretical foundation is therefore the same as the Council of Europe’s. Each item is pre-tested with representative samples from the TCG target population according to a standardised procedure; results are analysed statistically using a procedure which eliminates too difficult or too simple items that are unsuitable in distinguishing between different levels of language ability; contain inappropriate distractors; cause scores to fluctuate due to irrelevant factors (i.e. gender, nationality, etc.). The items that are retained in the live test are statistically calibrated to ensure the reliability and accuracy of the grading system.

Components

There are 3 required tests, which contain 80 items and last 1.5 hours: listening comprehension – 30 items, 25 minutes; language structures – 20 items; 20 minutes; reading comprehension – 30 items; 45 minutes and 2 optional tests which last 2 hours: a 15-minute speaking test and a 1.45 hours written expression test. In the required tests, test items take the form of multi-choice questions with several options from which candidates must select the correct answer. Items are presented in ascending order of difficulty.

Oral interaction

Listening comprehension is a required 30-item test.

This exercise tests candidates’ ability to understand spoken French, in particular: common words and expressions in an ordinary, everyday context (i.e. conversations, interviews, discussions, telephone conversations, etc.); raw information contained in simple, straightforward messages and announcements; basic information about people and sequences of facts or events (i.e. a radio news bulletin, or concerning daily life or work); presentations on theoretical or practical subjects; spoken discourse, where person is speaking quickly.

This test is divided into four parts: Part 1 tests candidates’ ability to interpret statements dealing with pictures, images or illustrations depicting everyday scenes; part 2 tests

candidates' ability to understand recorded extracts of brief real-life exchanges (greetings, questions, phone conversations, inquiries, travel announcements, weather bulletins, etc.); part 3 tests candidates' ability to understand a series of conversation snippets (two or three conversations); part 4 tests candidates' ability to understand a presentation, a radio programme, brief extracts from radio newscasts. Candidates may have to cope with several questions on a single extract or rapid delivery by speakers, or background noise.

The speaking test is an optional test. It consists of an individual face-to-face recorded interview taking approximately 15 minutes. Interview questions are graded according to the six-level Common European Framework scale (elementary to proficiency) and assess candidates on their command of the skills required at each level.

Candidates are evaluated on their ability to: describe the place in which they live and the people they know; describe other people, living conditions, work and studies, both in the present and in the past; relate stories, experiences, ideas, events and objectives or describe the plot of a story/film and give a personal reaction; present an opinion, explain the advantages or drawbacks of an idea, express agreement and/or disagreement; describe or argue a case in a clear and structured manner and contextually appropriate style; provide detailed summaries, commentaries and conclusions on complex topics.

Reading test

This test incorporates 30 items. Candidates are tested on their ability to understand: familiar names, easy words and phrases used in communication situations (informal messages and letters, administrative inquiries); language used in ordinary documents, (i.e. classified advertisements, brochures, menus, timetables); information about people, facts and events (i.e. informal letter); routine texts written in ordinary language about everyday life or work-related issues; magazine and newspaper articles in which writers give opinions on concrete or abstract subjects; long and/or complex factual or literary texts and specialist articles; texts dealing with abstract or complex subjects, extracts, specialist articles, works of literature.

Writing test

The required 20-item language structures test assesses candidates' ability to: identify mistakes; select the correct formulation from a series of syntactic or lexical structures. Grammatical structures are systematically related to communication situations and contexts which may be either familiar or abstract according to the testing level.

In the actual written expression component of the test, candidates have to complete all six of the set writing tasks within 1 ¾ hrs: draft a simple, informal message of around 40 words in length for a postcard or a short note; write a friendly letter of around 60 words in length on a topic from everyday life which shows a command of informal and familiar language levels; compose an 80-word summary of an event, presentation, account or opinion; present and argue a case in a text of around 100 words; draft a clear, detailed and convincing argument of around 100 to 125 words based on a social issue debate; summarise a text or a series of documents in approx. 100 words.

This test comprises six writing exercises, all of which are mandatory. Candidates are not allowed to use reference grammars, dictionaries or electronic dictionaries.

Candidates are evaluated on their ability to: present their message in a coherent way; provide the required information; structure their ideas in a clear way by presenting them in sequence; express and argue their opinions; use a wide range of vocabulary which is appropriate for the required task; use complex grammatical structures correctly; use language accurately and reformulate ideas in their own words.

Reliability and validity

Reliability

The TCF demonstrates high inter-rater and intra-rater reliability due to the way it is built and corrected within the official testing centres.

Validity

The TCF is also a valid test though the face validity of the speaking test is a bit weak. However, the listening comprehension, reading and written expression components are quite strong. There is a debate on the relevance of mistakes identification nowadays and the TCF incorporates error identification in its use of grammatical structures component. This is far less of a problem than certain traits or lack thereof found in other tests.

Scoring

Marking of examination papers is completely centralised; required test papers are marked by scanner while optional test papers are corrected by the same team of markers using a standardised marking scale. Papers are marked at the Centre international d'études pédagogiques by a team of accredited markers. The examiner's opinion is also taken into consideration during marking.

Each mark is double-checked.

The speaking test and the written expression tests are scored out of 20.

Test administration

In fact, official testing centres administer the TCF test in compliance with their precise standard procedures. There are two testing centres in Toronto.

Administration time

Speaking:	15 minutes
Listening comprehension:	25 minutes
Reading comprehension:	45 minutes

Language structures: 20 minutes
Written expression: 1 ¾ hrs.

Cost

Awaiting information.

TEST DE FRANÇAIS INTERNATIONAL – TFI

Overview

The Test de français international is a product of Educational Testing Service (ETS), developer of tests such as TOEFL (Test of English as a Foreign Language) and TOEIC (Test of English for International Communication). ETS is the leading private educational assessment organization in the world. The TFI is designed as a management tool for Human Resources to use when making decisions about an individual's French-language ability. It is used by businesses, language programs, government organizations and educational organizations.

According to the TFI website, it is used for recruiting, hiring and determining promotions; selecting employees for training to be conducted in French; defining goals and required language levels; measuring progress and proficiency level following French-language training; establishing a reliable standard for French proficiency levels required within a given organization.

Background

In the early sixties, about thirty public and private organizations got together to develop an assessment tool that would enable decision-makers in the educational and business worlds to assess the knowledge of English of non-native speakers. The Test of English as a Foreign Language (TOEFL) was born. It was later followed by the Test of English for International Communication (TOEIC) which is business-based and administered to more than four million test takers each year. The TOEFL was later introduced as a computer-based test in July 1998. The Test de français international was launched in the late 1990s.

Test design and theoretical foundation

Since 1997, teams of researchers have been working extensively on the design of a new fully communicative TOEFL to replace the existing one which is based on models of language measurement no longer in favour in the linguistics world. The 1960s' TOEFL was more structural in scope. Communication was not its main intent. This is evident when one examines the Building Skills for the TOEFL practice exercises. Moreover, it is unclear to this consultant how the test designers arrived at the point rationale the TOEFL uses in its six proficiency levels nor on the disparity of point spread between the various

levels (points range from 240 at the Beginning proficiency level to a 85 point differential at the Professional proficiency level.)

Apart from issues of reliability and validity, the international use of the TOEFL raised issues of practicality. Multiple choice test assesses a candidate's ability to understand, speak, read French as it is used in the international workplace. The TFI follows the same format and therefore suffers from the same shortcomings.

Components

The TFI contains 180 questions (divided into two sections: listening and reading). Each section is composed of six parts. This is a multiple choice test.

The listening section is composed of three sub-sections: Question-Answer with 40 questions; short dialogues with 30 questions and short conversations with 20 questions. Different French accents are used in the recording for the listening section.

Section ii – Reading – is made up of Error identification (25 questions); incomplete sentences (25 questions) and Comprehension (40 questions).

There are no speaking or writing components to the TFI.

Reliability and validity

Reliability

Reliability is the TFI's strong point as a multiple choice test. The candidate chooses the answer he/she believes is correct. As long as guessing does not enter into his/her responses, intra-rater reliability should be very high. Multiple choice tests do not leave room for disparities in correction, thereby guaranteeing inter-rater reliability.

Validity

Validity is another matter. There has long been critical discussion of the validity of the TOEFL and consequently any test that is designed on the same premise. We can make a case for face validity and construct validity but cannot grant content validity to a test that claims to assess speaking and writing when it does not attempt to measure those two skills.

Scoring

Scorability is high as the test is multiple choice, the only way in fact to guarantee consistent, easy and fast results.

Test administration

The TFI test has to be administered through accredited centres on specific dates. There were eight dates in 2006. There is one accredited centre in Toronto (Alliance française). Candidates have to register in advance.

Administration time

The listening test lasts 42 minutes. The reading test lasts 68 minutes for a total test duration of 1 hour and 50 minutes.

Cost

As of October 2006, the test cost was \$ 52.16 + GST

TEST D'ÉVALUATION DE FRANÇAIS (TEF)

Overview

This test was created in 1998. This is one of the tests Citizenship and Immigration Canada lists as one of their approved language proficiency test, the results of which are used by CIC as a conclusive proof of the immigrant's language skills. The Paris Chamber of Commerce and Industry administers this test.

Background

The Chamber of Commerce and Industry states the TEF is a fast, valid and objective measure administered in 300 centres worldwide.

Test design and theoretical foundation

The TEF was designed and developed by the Paris Chamber of Commerce but it has not been possible to unearth information on its design and theoretical foundation to date.

Components

The Test d'évaluation de français incorporates oral expression, oral comprehension, reading comprehension and written expression. Three components are compulsory: vocabulary and structure; oral comprehension and written expression. Two components are elective tests: written expression and oral expression.

Reliability and validity

Reliability

The likelihood of reliability is fairly high due to the standardized nature of test taking except for speaking and writing.

Validity

Validity is fairly high in as much as this test components actually measure what they intend to measure (the speaking and writing components do relate to actual speaking and writing). There is also in this test a small error identification component, the legitimacy of which is in doubt but otherwise, the exercise types used in the compulsory components of the test show face validity. There is however another disturbing factor in the oral comprehension of the test as it incorporates a memory aspect which should not be included. The candidates end up being tested on their memory rather than on their comprehension of French. This is a regrettable pitfall in an otherwise well constructed test. The blatant European context is also a problem for our purpose here.

Scoring

The TEF parallels the Common European Framework in as much as it incorporates 6 levels. The A levels of the Council of Europe are divided into three elementary levels for the TEF (0+: 0 to 68 points; 1: 69 to 203 points; 2: 204 to 360 points). The B levels from the Council of Europe become 2 intermediate levels (3: 361 to 540 points; 4: 541 to 698 points). The Council of Europe's two C levels become 2 Superior levels with the TEF (5: 699 to 833 points; 6: 834 to 900 points).

Test administration

The test is administered in Alliance française de Toronto. There now is an electronic version of the test. Test dates and cost have been required but not obtained.

Administration time

Vocabulary and structure:	30 minutes
Oral comprehension:	40 minutes
Reading comprehension:	1 hour
Oral expression:	35 minutes
Written expression:	1 hour

Cost

Awaiting information.

RECOMMENDATIONS

It is important for an institution to avail itself of a reliable, valid but also easily understandable tool so the candidates who will have to take the test have access to concise and clear information. Human resources administrators, for their part, must be

able to defend their decision making process. They need to rely on transparency of assessment, form, content, evaluation method, results and meaning.

Government of Canada

Neither the Government of Canada rating scale, nor its assessment tools correspond to the Ontario's College of Teachers needs and expectations when it comes to proficiency assessment. We do not recommend a three step system, which we find too vast and imprecise. We do not totally recommend the assessment tools. We agree with the premise of an oral interview but consider the Government does not focus enough on the listening comprehension aspect of the language. While we agree with context-appropriate reading tasks, the lack of written expression components and the claim that the written expression component tests writing when it clearly does not, cause us to suggest these system and tools be disregarded. Furthermore, on a strictly practical plane, the assessment tool is reserved strictly to the federal public service. Using it for the College's purposes is therefore not possible.

Government of Ontario

The lack of transparency on the part of the Ontario provincial government when it comes to its system and testing tools cause us to steer clear of recommending it be used for the College. The definition of the proficiency levels seems to have been adopted from an existing American system. We will therefore comment on the original system and not an adopted or possibly adapted one. The tool used to assess speaking skills is akin to that of the Government of Canada. It suffers from the same shortcomings in design, reliability and validity. We do not know how reading and writing are assessed. The individuals charged with applying the assessment tools are not willing to share information. We would recommend the College look elsewhere for its language proficiency assessments.

City of Ottawa

If Human Resources professionals seek a correlation between an existing language training program and a second language proficiency level structure, a structure that incorporates the same levels as those of the language training program is understandable and logical. The descriptors for each level become the target objectives for the language training program. The prospective training participants and the candidates to a position then know where they rate upon testing, can be placed in the appropriate level class and progress from there until they have reached the required level of second language proficiency for their position.

Because the avowed aim of the College is to “identify alternative systems for establishing appropriate oral comprehension, speaking, written comprehension and writing French language competency levels appropriate for Ontario College of Teachers bilingual jobs providing services to French language stakeholders and community”, the City of Ottawa scale is deemed too restrictive and consequently not useful for College use.

The language proficiency level structure used there has a different purpose from that of the College. Consequently, we recommend the College disregards this structure as

irrelevant to its stated aim. The assessment tools used are the closest in design to what the College would want to consider since it takes into account occupationally-specific tasks for all four skills. The interview incorporates occupationally-related elements, the listening comprehension tape is entirely job specific, the reading components are all occupation-dependent and the writing tasks, which actually test writing, also refer to tasks accomplished in the workplace. While desirable in design, it is impractical in fact as it is time and labor-consuming. Moreover, the tests are City-specific and owned and not for market consumption. However, we recommend the College takes the occupationally-related concept into further consideration. We have a case here where the level proficiency structure is irrelevant but the test design interesting and adaptable.

Canadian Language Benchmarks

We have here a very interesting system. While the Benchmarks have not been created for language testing, the twelve level descriptions for each of the four skills provide invaluable information as to what each level encompasses. The companion tables to the English language Benchmarks describe speaking, listening, reading and writing tables with features of social tasks, interaction tasks, instructional tasks, suasion tasks, information tasks etc. The existence of twelve levels might appear cumbersome when other systems have delineated 10 (ACTFL) or 6 (CEF, TCF, TFI) levels. This is a Canadian system, which may be an added advantage. It is called upon to be widely used within our borders. While the system offers a lot of promise, there is no existing French language assessment tool. In this case we have a highly researched system but no practical tool to match.

ACTFL Proficiency Guidelines

The system itself is known throughout the world and widely used for the testing of numerous languages. The consideration of a distinguished level goes beyond College's language needs. The Oral Proficiency Interview is what seems to have been adapted or adopted by the Government of Ontario. The scope of described language from novice to superior seems appropriate and exists for the four skills. The related assessment tools however seem to be incomplete. The Oral Proficiency does test speaking, is weak on listening as previously stated and can only relate to occupation-specific circumstances if the assessor makes a definite attempt at asking job-specific questions in the interview. Related reading and writing tests could not be found. However, this consultant will be attending an ACTFL workshop later in November on writing proficiency and expects to gather useful information. She still does not anticipate finding a ready-made tool that would meet the needs of the College. We are therefore faced with a definite shortcoming. The scale could be useful. The tools need adaptation or design.

Common European Framework (CEF)

Europe is feverishly working on a system and tools that will meet testing needs in about twenty-five languages. The reduction of language proficiency to six levels seems to us a bit excessive. It does match the ETS six level scale though and may have been chosen

for practical reasons. The potential for test takers in Europe is astounding. ETS with its TOEFL and TOEIC tools has a huge market as well. The description of the CEF levels is still sketchy and do not do justice to the rich character of language. We would therefore propose to leave it aside and concentrate instead on the American or Canadian system. Assessment tools are being marketed and there is no doubt that with time a number of reliable and valid assessment tools will flood the market. Content, however, will be a problem for us on two counts: the tools are culturally and linguistically inappropriate: they refer to concepts and places unknown here; the language used does not always reflect language in use in Canada. For these reasons, we would leave the system and tools aside though the design of the latter is certainly inspiring should assessment tools need to be created.

Educational Testing Service (TFI)

This is the best-known organization worldwide. Creators of the TOEFL and TOEIC, ETS has recently launched TFI, with six levels of proficiency. TOEFL scores are required for the purpose of admission by more than 2,400 colleges and universities in the United States and Canada alone. It is widely used by regulatory bodies for the purpose of certification and organizations for hiring purposes. Our concern here is with the TFI tool, which does not boast such widespread use since it was launched in the late 1990s as opposed to 1964 for the TOEFL. Again, the system choice of rating scales seems to us too restrictive: the description of the levels is too sketchy and not necessarily reflective of the scope of language. And more importantly from our standpoint, it does not measure what it claims to measure. For the ETS to say the TFI assesses a candidate's ability to speak and write French is astounding, given the components of the assessment tool. For these reasons, we would disregard the TFI for the College's purposes.

TCF and TEF

The apparent forthcoming information on these tests theoretical foundation are problematic. Their correlation to the European-wide Common European Framework might be dubious. The content of their test, while very interesting in construct, is not culturally appropriate for the College's purpose.

TESTS FOR OCCUPATIONAL PURPOSES

These tests have not been analyzed for lack of relevance from a content point of view. A professions specific tool is developed for a specific purpose. To determine relevant competencies, the profession needs to be investigated to determine the language in use and the test designers need to work closely with practitioners.

The face and content validity of these tests are normally high. As they are tailor made, Human Resources administrators, for example, can determine the length and type of test administration they require and the emphasis they want placed on the skills they deem most important. A profession specific tool should replicate the context of language use,

test both the technical and non-technical language that is needed for the profession, be reliable, valid, fair, relevant, representative of authentic communication. It needs to be practical and needs to be a test of language ability, not professional ability. Profession specific tools are costly to develop and require a high level of expertise. They are expensive to maintain, time-consuming to administer, difficult to keep secure, and a significant responsibility for the institution. However, they are the only solution to ensuring candidates will be able to function in their second language **in their workplace**.

REMARKS ON RELIABILITY AND VALIDITY

Interestingly, tests which actually test what they are supposed to test do not rate very high on a reliability scale as there always is an element of subjectivity whenever one has to deal with development type test questions (speaking and writing). This is usually alleviated by a tight correction grid in the number of points to be deducted for the type of mistakes read or heard. Inter-rater reliability is further enhanced by proper assessor's training. In practice, a beginning assessor usually needs to rely on a second opinion (if only to allow him/her to build self-confidence). It becomes obvious with time that several assessors listening to the same taped test or correcting a written assignment show agreement in the correction of the test.

Tests which rate on the average in reliability often rate higher on the validity factor. This is because if they test what they are supposed to test, they are inherently valid from a content standpoint.

PROPOSED MODEL

The rating scales which we favour are the Canadian Language Benchmarks in their French version - Standards linguistiques canadiens and the ACTFL proficiency guidelines.

Both systems are North American, with the Benchmarks having the added advantage of being a national system.

The Benchmarks are not aimed at being the basis for testing. The French version of the Benchmarks just came out. No doubt they will be used as such in the near future, regardless. The ACTFL proficiency guidelines, on the other hand, have been used as the basis of proficiency measurement for several years. They have been adopted or adapted by the government of Ontario. We prefer direct reference to the guidelines. The number of levels and their description seem appropriate for the College's purpose.

PROPOSED ASSESSMENT TOOL

That is where the problem lies. Neither the Benchmarks nor the ACTFL come with a ready-made test, which could be used or modified for occupational purposes.

To test oral expression, it is the opinion of this writer that you do need to measure speaking with a human being as a counterpart. The ACTFL **oral proficiency interview** meets that need. Therefore, it might be possible to find an agency with experienced

assessors in the conduct of oral proficiency assessment as per the ACTFL guidelines. The drawback in this is that there is no **oral comprehension** component that tests the actual comprehension of a sample of language used in the test taker's milieu. Speaking proficiency will therefore be assessed but not oral comprehension per se. This can be remediated by creating occupationally-related oral comprehension sentences but doing so requires development time and finances.

To test vocabulary and structure, incomplete sentences, word substitution, structure recognition can be used.

To test **reading comprehension**, one needs to have a reading test. This, too, should be work-related so a general test of the kind that can be found on the market with the various tools reviewed does not meet the need if one wants to assess the reading comprehension of the texts employees are called upon to use in their workplace.

There are, however, some interesting type of exercises that designers can be inspired by: recognition of familiar names, easy words and phrases used in communicative situations; language used in ordinary documents such as brochures, timetables; information about people, facts and events under the guise of an informal letter; routine texts written about work-related issues; complex factual texts and specialized articles such as the TCF uses. Test takers can also be asked to define the context of a document (where, who, to whom, when, for what purpose); verify detailed comprehension through multiple choice or texts that need to be sequenced in a coherent and logical fashion to make sense.

If we are looking for ease and rapidity in scoring, multiple choice are easier to use than development answers. They are however more time-consuming to create than development questions are.

To test **written expression**, we are of the opinion that the test taker should be required to actually write a relevant text that the test taker will be called upon to write in his/her work milieu. There is no easy way to correct such a component but the time needed to correct it is compensated by the validity of the required task.

CONCLUSION

Gail Stewart justly indicates that "key administrative factors for institutions in selecting tests include cost efficiency, time efficiency and convenience." Unfortunately, when administrators want to assess the occupationally-related second language competence of their employees, they cannot rely on tools that have been created with a different purpose in mind (i.e. placement tests or tests for admission in an academic institution).

Regulatory bodies are beginning to seek occupationally-relevant tests but the creation of same is specific to their purpose and cannot be used for another organization with a need to assess different professions. Occupationally-related testing is notoriously more relevant and valid than general tests but they are also far more restrictive in their use and expensive to create.

For this project, seven rating scales and seven language proficiency assessment tools were superficially reviewed. Time did not permit an in-depth review of each or in-depth research of information that did not appear forthcoming be it in research articles or by people involved in the milieu of interest.

Each system and tool were briefly described. A checklist was used which included information related to the administration of the test, the time it takes to administer, test cost, test components, test scoring and the purpose for each test. Information was also provided on the theoretical framework of the tools and the technical standards for the test when available. Information on each test depended on the amount of information provided in public domain documentation. Therefore, the descriptions of each test varied. Recommendations were made on preferred rating system and type of assessment tools. Judgement was made as pertains to the needs of the College with the aim of correlating the test and the expressed needs of the College.

¹ Content standards: an analysis of target language use situations and the performance(s) which it is intended to predict.

Performance standards: a descriptive taxonomy of how well candidates know and are able to what is set out in the content standards.

² twelve-step numerical scale: 1; 1+; 2, 2+; 3, 3+; 4,4+; 5, 5+; 6, 6+

³ nine-step scale: elementary 1, elementary 2, elementary 3; intermediate 1, intermediate 2, intermediate 3; advanced 1, advanced 2, advanced 3.

REFERENCES

- Bolton, S. (1987). *Évaluation de la compétence communicative en langue étrangère*. Paris, France : Hatier-Crédif.
- Brière E.J. & F.B. Hinofotis eds. (1979). *Concepts in Language Testing – Some Recent Studies*. Washington DC: Teachers of English to Speakers of Other Languages.
- Byrnes H. & I. Thompson. (1989). *The ACTFL Oral Proficiency Interview. Tester Training Manual*. Kathryn Buck, Editor
- Canale M. & M. Swain. (1981). A theoretical framework for communicative competence in *The Construct Validation of Tests of Communicative Competence* ed. By A.S. Palmer, P.J.M. Groot & G.A. Trostler. Washington DC: Teachers of English to Speakers of Other Languages
- Canales, J.A. (1994). *Linking Language Assessment to Classroom Practices*. In R. Rodriguez, N. Ramos & J.A. Ruiz-Escalante (eds.) *Compendium of Readings in Bilingual Education: Issues and Practices*. Austin, TX: Texas Association for Bilingual Education.
- Carroll B.J. & P.J. Hall.(1985). *Make your own language tests*. Oxford, England: Pergamon Press.
- Centre for Canadian Language Benchmarks: *Background Information. Benchmarks. Canadian Language Benchmarks 2000. CLB 2000: Companion Tables. Guidelines for Test Development. CanTest. CELBAN – Niveaux de compétence linguistique canadiens 2006 - online*
- Centre international d'études pédagogiques. *TCF – Test de connaissance du français – online*
- Chambre de commerce et d'industrie de Paris. *e-tef. Test d'évaluation de français - online*
- City of Ottawa (2003). *Language Proficiency Level Structure*.
- Clark, J.L.D. (1972) *Foreign Language Testing: Theory and Practice*. Philadelphia: Center for Curriculum Development.
- College of Pharmacists of British Columbia web site.
- Council of Europe. (2001). *Common European Framework of Reference for Languages: Learning, teaching, assessment*. Cambridge, England: Cambridge University Press.
- De Jong J.H.A.L. ed. (1990). *Standardization in Language Testing – AILA Review*.

Del Vecchio, A. & M. Guerrero (1995). *Handbook of English Language Proficiency Tests*. Evaluation Assistance Center – Western Region. Albuquerque, New Mexico: New Mexico Highlands University.

Educational Testing Service Canada Inc. – Test de français international – website

Government of Ontario (1998). Request for Proposal for French Language Proficiency Evaluations Services and Language Skills Training.

Hart-Gonzalez, L. (1994). Raters and scales in oral proficiency testing: the FSI experience. Presented at the Language Testing Research colloquium, Washington D.C.

Hudon, Marie-Eve. (2006). *Official Languages in the Public Service: From 1973 to the Present*. Ottawa: Library of Parliament Parliamentary Information and Research Service.

MacFarlane, A. (2004). *French Second Language Proficiency Test II*, Canadian Association of Second Language Teachers.

McNamara, T. (2001). The challenge of speaking: research on the testing of speaking for the new TOEFL. In *JALT Testing & Evaluation SIG Newsletter*, vol. 5, no. 1.

Moloney, B. (2000). Book review: Alderson, J.C. *Assessing Reading*. Cambridge: Cambridge University Press.

Morrow, K. ed. (2004). *Insights from the Common European Framework* Oxford: Oxford University Press.

National Association of Pharmacy Regulatory Authorities website

North, B. (1993). *The Development of Descriptors on Scales of Language Proficiency*. Washington DC: The National Foreign Language Center.

Office of Francophone Affairs. (2002). *Staffing & Managing Designated Positions: A practical guide for managers* – Toronto: Government of Ontario

Omaggio, A.C. (1986). *Teaching language in context. Proficiency-oriented Instruction*. Boston: Mass. Heinle & Heinle.

Ontario College of Pharmacists website

Ordre des thérapeutes respiratoires de l'Ontario website

Pawlikowska-Smith, Grazyna (2002). *Canadian Language Benchmarks 2000: Theoretical Framework*. Centre for Canadian Language Benchmarks.

Public Service Commission of Canada. Determining the Linguistic Profile for Bilingual Positions – online

Public Service Commission of Canada. Second Language Evaluation: Oral Interaction Test – online

Public Service Commission of Canada. Second Language Evaluation: Reading Test – online

Public Service Commission of Canada. SLE OI Test Information – Brochure for Candidates – online

Public Service Commission of Canada. Second Language Evaluation: Writing Test - online

Reboul, M.-R. & A.S. Badière. (2001). Test de compétence linguistique à l'oral – Manuel du testeur. Toronto, Canada.

Shohamy E. & A.R. Walton eds. (1992). Language Assessment for Feedback: Testing and Other Strategies. Dubuque: Iowa Kendall/Hunt.

Stewart, G. (2001). Developing professionally-relevant and fair language testing. In Access Issues for Regulators.

Underhill, N. (1982). On Tesol 82 Pacific Perspectives on Language Learning and Teaching. Ed. by M.A. Clarke & J. Handscombe.

University of Ottawa – Second Language Institute – CanTEST website

Appendix 3: Canadian Language Benchmarks

INTRODUCTION

In 2005, the College Registrar and Chief Executive Officer established as a priority the enhancement of French language services following a series of consultations with French language stakeholders. In the fall of 2006, the Ontario College of Teachers launched a four-component project whereby:

1. French language proficiency assessment systems would be researched, compiled and analyzed;
2. linguistic profiles would be established for positions designated bilingual by the College;
3. Appropriate tools would be identified for authentic assessments of language competency;
4. agencies capable of administering French language assessments would be identified.

Research indicated that the most applicable system for the College would be an adaptation of a Canadian system created, not for language testing, but originally for adult immigrant language training. This system, which breaks language down into twelve (12) levels and the global descriptors it supplies, lends itself to adaptation for language testing as an organization attempts to find out, through appropriate assessment tools, what a candidate to or incumbent in a position can actually do in a second language.

The Canadian Language Benchmarks (CLB) are, in the words of their authors:

“A descriptive scale of communicative proficiency* ... expressed as 12 benchmarks or reference points;

a set of descriptive statements about successive levels of achievement on the continuum of language performance;

descriptions of communicative competencies and performance tasks in which the speaker, reader or writer demonstrates application of language knowledge and skill;

a framework of reference for learning, teaching, programming and assessing adult English as a Second Language in Canada;

a national standard for planning second language curricula for a variety of contexts;

a common “yardstick” for assessing the outcomes.”

The authors of the Benchmarks were careful to say that the Benchmarks are not “descriptions of discrete elements of knowledge and skills that underlie communicative proficiency; a curriculum guide... a test.”

Here is a link to the benchmarks:

http://www.language.ca/display_page.asp?page_id=570

Appendix 4: Groupings of Designated Bilingual Positions

DEPARTMENT/UNIT

Membership Services Department

Administration Unit

Executive Secretary	Secretarial
Membership Analyst	Evaluation

Evaluation Services Unit

Senior Evaluator	Evaluation
Evaluator	Evaluation
Evaluation Assistant	Evaluation
Secretary, Evaluation Services	Secretarial

Client Services Unit

Client Services Officer	Client Services
Correspondence Assistant	Client Services
Receptionist	Client Services
Client Services Assistants	Client Services
Client Services Researchers	Client Services
Secretary, Client Services	Secretarial

Membership Records Unit

Membership Records Officer	Evaluation
Membership Records Assistants/Researchers	Evaluation
Secretary, Records	Secretarial

Three leadership positions (Co-ordinator, Evaluation Services Manager, Client Services Manager, Records Manager)

Leadership/Educator

French Language Services Department

Administration Unit

Co-ordinator, French Language Services	Leadership/Educator
Executive Secretary	Secretarial

External Relations

External Relations Asst	Public Service
External Relations Officer	Public Service

Investigations and Hearings Department

Investigations Unit

Investigator

Investigation

Intake and Hearings Unit

Intake Assistant

Client Service

Resolution Officer

Investigation

One leadership position (Co-ordinator, Investigations Manager or Intake and Hearings Manager)

Leadership/Educator

Professional Affairs

Administration Unit

Executive Secretary

Secretarial

Accreditation Unit

Program Officer, Accreditation

Leadership/Educator

Accreditation Assistant

Evaluation

Secretary, Accreditation

Secretarial

Standards of Practice Unit

Program Officer, Standards of Practice

Educator

Secretary, Standards of Practice

Secretarial

One leadership position (Co-ordinator, Accreditation Manager or Standards Manager)

Leadership/Educator

Executive Department

Administration Unit

Human Resources Assistant

Client Services

Executive Secretary (one of four positions)

Secretarial

Communications Unit

Translator

Communication

Communications Officer

Communication

Production Officer

Communication

Policy and Research Unit

Library Assistant

Client Services

Information Technology

Help Desk Specialist

Client Services

One leadership position (one of Registrar, Deputy Registrar, Co-ordinator and unit managers)

Leadership/Educator

Appendix 5: Final Report of the Consultant: French Language Proficiency Assessment Project

This project was divided into three components.

Component 1 related to the identification of French language oral expression and comprehension, written comprehension and written expression assessment systems which might be relevant to the College's purposes.

Various models were reviewed, among them those used by the governments of Canada, Ontario and the municipal government of Ottawa. The Canadian Language Benchmarks, the ACTFL proficiency guidelines, the Common European Framework and Educational Testing Services assessment models were examined. Their assessment tools were studied with particular attention paid to their background, design and theoretical foundation, components, reliability and validity factors, method of scoring, test administration, administration time and cost.

The first recommendations were then proposed:

1. That the College adopt the Canadian Language Benchmarks as the basis for its French language assessment system.
2. That related assessment tools be created, using, as a first step, an oral interview conducted by an examiner and aiming to verify the oral expression and oral comprehension of candidates, as per the ACTFL Oral Proficiency Interview method for oral expression.
3. That oral comprehension assessments be created using a number of different exercises.
4. That lexis and grammar be created and assessed separately.
5. That reading comprehension exercises be created and assessed on the basis of a four (4) section assignment.
6. That written expression tools be created and assessed on the basis of a six-exercise tool.
7. That developments on internet-based methods of second language assessments be monitored so the College can avail itself of same as soon as it becomes a reliable testing method for oral skills.
8. That the Consultant share with the College representatives her findings on assessment tools while attending the ACTFL Annual Conference in November 2006 in Nashville, U.S.A and Expolangues in January 2007 in Paris, France.

In a November 10, 2006 meeting, the College's representatives, Mr. Frank McIntyre and Mrs. Lise Roy-Kolbusz met in Toronto with Mrs. Gorrie to discuss her recommendations. It was then decided that the College would adopt the **Canadian Language Benchmarks** as its scale for French language assessments.

It was also decided that any **assessment tool**, however designed, should **not last longer than two hours** and that any assessment tool design would only be created if no adequate assessment tool was found to meet the College's requirements.

Component 2 dealt with establishing French language competency requirements levels for approximately 35 different positions within the College. Linguistic profiles for oral expression, oral comprehension, written comprehension and written expression were drawn up following an in-depth study of the relevant job descriptions provided by the College.

Meetings with coordinators at the College were scheduled for February 2007 to review and amend the suggested profiles as per the coordinators' appraisal of their linguistic situation within their respective units. A revised version is now available.

Component 3 was to deal with the identification of appropriate tools for authentic assessments. However, it was determined, in the first component, that, barring any discovery in Nashville and Paris, occupationally-related language assessments, which would enable the College to meet its requirement of adequately and accurately providing French language services, do not exist at the present time.

Recommendation of the most appropriate system for College use in future staff recruitment relies on the design of made-to-measure tools, which will require an important investment from the College for test design. Component 3 also dealt with the concept of exemption and that of a portfolio assessment process.

Recommendations on these two issues were as follows:

Linguistic identity:

1. That the College determine how it will identify French as a first language speakers.
2. That the College formulate the question(s) it will use to determine its applicants' Francophone identity

Exemption:

3. That no linguistic group be exempt from initial second language testing.
4. That Francophones not be exempt from assessment in French written skills when their position's linguistic profile requires a minimum level of proficiency of 10.
5. That a portfolio (for written skills) be required of candidates who have been educated in French and whose linguistic profile requires a level between 5 and 9 in written skills.
6. That exemption be granted from further testing (in the case of transfer or promotion) only for candidates who have reached a level of 12 in a language skill during initial testing.

Portfolio:

7. That a certification of authenticity and ownership statement be included in the portfolio to guarantee authorship of the provided French written skills documentation.
8. That a self-assessment grid based on the Canadian Language Benchmarks be designed.

Identification of agencies capable of administering French language assessments in support of the College staff recruitment:

The Consultant noted that a previously conducted survey in October with testing agencies accredited with the Government of Ontario did not yield any information. The Consultant therefore suggested the College might meet with less resistance and better success than an external Consultant, who also happens to conduct language assessments. A list of these agencies was provided to the College.

It was also indicated that the assessment means were to be considered: face to face oral interaction assessments or assessments via telephone.

It was suggested that:

9. A request for proposals be drawn up with specific characteristics delineated in a provided appendix for the College to have language assessments conducted.
10. Assessments tools, which the testing agencies would use, be designed with the following components:
 - Assessment grid based on the Canadian Language Benchmarks
 - Oral interaction interview questions
 - Occupationally-related oral comprehension sentences
 - Occupationally-related role plays
 - Written comprehension tests with selection of text and questions or multiple choice distractors
 - Written expression tests based on the current assignments used by the College as companions to the hiring interviews
 - Grammar correction grid to assist with the above-mentioned written expression assignment
11. These designed assessment components cover testing for ten (10) position categories which would cover all College positions, for the time being. This approach enables a more time and cost-efficient approach rather than develop tests for all position titles. The identified categories are: leadership; supervisory; technical; client service; administrative; secretarial; communications; public relations; investigative and educator jobs.

12. That, to further reduce cost, an incremental approach be used so only a couple of tests be drawn up per category before moving incrementally as time goes by thereby creating a bank of tests later on.
13. That the College decide if it wants to call upon agencies offering their own French language assessment tools and ask them to correlate their scale to that of the Canadian Language Benchmarks or start designing its own assessment tools to be either given to agencies to administer or to be given by College staff and corrected externally with an agency using the Canadian Language Benchmarks as its scale.

The Consultant further offered thoughts on the validity, authenticity, cost effectiveness, efficiency in test administration, accessibility to test administrators and global test administration.

LATEST DECISIONS

Subsequent to meetings with the College's Human Resources Manager and with the French Language Services Co-ordinator in February, it was decided:

- That an assessment grid based on the Canadian Language Benchmarks, would indeed be created.
- That the ten proposed position categories would be reduced to seven (7) categories identified by the Human Resources Manager.
- That another project might be drawn up for test design, after reception of a test design estimate by the Consultant and approval by College authorities.
- That the timeframe for test design would not expand past September 2007.
- That a four skills test should not require more than two (2) hours excluding the written assignment presently given for hiring purposes.
- That managers would prepare unit-specific relevant written documentation to serve as a basis to develop written comprehension tests. Human Resources would share its written expression assignments during its hiring process to serve as a basis for occupationally-relevant written expression tests.

The Consultant suggested that any designed test be conducted with volunteers (present employees of the College) to verify its relevance to the College context before these assessment tools are officially used for hiring purposes.

CONSULTANT'S CONCLUDING REMARKS

This was a fascinating project which enabled the Consultant to verify what is presently available as far as language test systems and tools are concerned; suggest the best available system; suggest linguistic profiles for bilingual-essential positions; share thoughts on exemption and portfolios; make recommendations on testing agencies and the best way to conduct assessments with the appropriate tools to measure the occupationally-related linguistic proficiency of the College's candidates to a position.

The extraordinary cooperation received by the Human Resources Manager, French Language Services Coordinator and all Unit Coordinators has made this project a very fruitful experience. Seldom do we find people who take the time and demonstrate enthusiasm for what is usually considered a sensitive issue. The College approached it with a positive and rationale attitude which made the whole project a truly pleasurable endeavour.