

A Longitudinal Evaluation of an Early Immersion School Program

Fred Genesee

protestant school board of greater montreal

On a comparé le rendement des élèves participant à un programme précoce complet d'immersion en français à celui d'élèves homologues inscrits dans les programmes normaux d'anglais et de français. On a mesuré ce rendement à la fin de chaque année scolaire (1^e à 6^e année) en utilisant une série de tests d'anglais, de français et de mathématiques. Les résultats des élèves d'immersion ont été inférieurs à ceux des élèves de contrôle anglais sur les tests qui nécessitaient une connaissance de la langue anglaise — tels que la lecture — jusqu'à ce que soient introduits dans le programme les matières de langue anglaise, période à laquelle ils arrivent à égalité avec les élèves du groupe de contrôle. On ne note pas de différences entre les élèves d'immersion et ceux du groupe de contrôle anglais, même en première année, sur les tests de langue anglaise qui ne nécessitaient pas une connaissance de la langue anglaise, tels que ceux de décodage auditif et d'association de mots, ou sur ceux de rendement en mathématiques. En même temps, la compétence en français des élèves d'immersion était grandement supérieure à celle du groupe de contrôle anglais sur tous les instruments de mesure compris dans l'évaluation. Les résultats des élèves d'immersion étaient de beaucoup comparables à ceux du groupe de contrôle français sur les tests français nécessitant des compétences en décodage, tels que la lecture et la compréhension auditive. Toutefois, leurs résultats sur les tests nécessitant des compétences de production — tels que l'expression orale — étaient en général plus faibles que ceux du groupe de contrôle français. Néanmoins, les élèves d'immersion avaient acquis une bonne connaissance pratique du français.

The purpose of this report is to summarize the findings of a longitudinal evaluation of an early French immersion program, 1970 to 1976, patterned after the St. Lambert Elementary School French Immersion Program (cf. Lambert & Tucker, 1972). Much has been written and reported about immersion programs for majority-group children in Canada and the U.S., most of it in agreement with the St. Lambert program results.

The French immersion program of the Protestant School Board of Greater Montreal (PSBGM) was inaugurated in 1967 in response to a growing concern of parents about their children's ability to function in a bilingual community. During the first two years of the program enrolment was limited to two classes a year and to relatively select groups of students. Registration in the program during its third year was completely open, and children with a fairly wide range of ability levels were enrolled. There were four classes, in total, in four schools. The early French immersion program has continued to grow, such that during the 1977/78 academic year there were 35 kindergarten classes in 15 French immersion schools with 721 students, or 24% of the kindergarten population of the PSBGM.

THE PROGRAM

During the first three years, kindergarten to grade 2, all classroom instruction is presented in French by native French-speaking teachers. This means that all language arts instruction, including initial reading, as well as content material such as arithmetic and science, is taught in French during these years. Even though the teacher addresses all of his/her comments to the children in French from the outset, the students communicate amongst themselves and with the teacher in English at least until approximately December of grade 1. During kindergarten and the first four months of grade 1 emphasis is placed on developing listening comprehension skills. By December of grade 1 students in most classes are encouraged to make all their classroom comments in French; this depends, of course, upon the particular teacher. English is introduced into the curriculum for one hour a day in grade 3 in the form of an English language arts course. (English is introduced in most other early immersion programs in grade 2.) Use of English is increased to two hours a day in grade 4, when selected course material is taught via English, for example, mathematics. English is increased to a maximum in grades 5 and 6, when it is used approximately 60% of the day (the morning), and French is used during the remaining 40% of the day (the afternoon).

In contrast to the immersion program, in the regular English program French is taught as a second language for 20 to 30 minutes a day from K–6.

The early immersion program has been evaluated systematically since its inception. The research presented in this report was carried out between 1970 and 1976 and can be divided into two blocks: that in grades 1 to 3, 1970 to 1973 (cf. Polich, Note 1) and that in grades 4 to 6, 1973 to 1976 (cf. Genesee, Morin, & Allister, Note 2; Genesee & Chaplin, Note 3; Genesee & Stefanovic, Note 4). The students who participated in these evaluations represent the third generation of students to enter the program; that is, they entered the program the first year that enrolment was completely open. Changes in research procedures to be reported here can be attributed in large part to the change in evaluators and, to a lesser extent, to other circumstantial factors.

The research was designed to answer three major questions:

1. What effect does participation in a French immersion program of this type have on the development of the students' first language?
2. How effective is the French immersion program in promoting acquisition of second-language skills?
3. Are students who participate in the French immersion program able to master academic material, such as mathematics, even when it is taught via the second language?

PROGRAM EVALUATION

The Samples

Three groups of students participated in the evaluations:

- 1. *Immersion*: anglophone students in the French immersion program since kindergarten;
 - 2. *English Control*: anglophone students in the regular program with French as a second language (FSL);
 - 3. *French Control*: francophone students in regular French schools within the PSBGM or the Commission des Ecoles Catholiques de Montréal (CECM).
- A summary of the exact sample sizes is presented in table 1. The anglophone samples (immersion and English control) were selected from immersion and regular-program schools within the PSBGM. The sample size and composition of the immersion group varied from year to year; notwithstanding these fluctuations, the immersion students were selected from the same classes every year.

Table 1 / Summary of Sample Information

		Sample Immersion	English Control	French Control
A. From Polich (Note 1)				
Grade 1:	Size	54	26	22
	Raven's	22.45	19.86	17.81**
	Lorge-Thorndike	15.83	13.14	N.A.***
Grade 2:	Size	54	18	16
	Raven's	26.90	25.44	22.67**
	Lorge-Thorndike	16.85	17.50	N.A. ns
Grade 3:	Size	47	17	17
	Raven's	28.11	24.93	24.35 ns
	Lorge-Thorndike	108.27	106.87	N.A. ns
B. From Genesee et al. (Notes 2, 3, 4)				
Grade 4:	Size	54	70	40
	Lorge-Thorndike	108.67	105.68	106.36 ns
Grade 5:	Size	46	64	58
	Lorge-Thorndike	108.60	106.20	107.40 ns
Grade 6:	Size	63	66	60
	Lorge-Thorndike	107.10	106.18	N.A. ns

N.A. = Not administered.

**p < .01

***p < .001

ns = Not significant according to analysis of variance procedures.

All three groups were administered Raven's Progressive Matrices (1958) in grades 1, 2, and 3; the immersion and English control groups were also administered the Lorge-Thorndike Intelligence Test. At the grade 1 and 2 levels, the latter test comprises three subtests — Vocabulary, Not-belonging, and Go-together — the results of which are combined to produce a total score. At the grade 3 level, the test comprises two subtests — Verbal and Non-Verbal.

The two grade 4, 5, and 6 anglophone samples were equated and stratified according to IQ. There were three IQ groupings: above average (full-scale

IQ above 115), average (full-scale IQ between 95 and 115), and below average (full-scale IQ below 95). The basis of the IQ classification was the student's score on the Canadian Lorge-Thorndike Intelligence Test (1967), which is administered annually in grades 4 and 6 in the PSBGM. The IQ scores reported in table 1 represent the average of the three groups combined. Test scores will not be presented for each IQ grouping separately; reference will be made to this distinction in the text, where appropriate. (See Genesee, 1976, for a complete report of the French-language test results according to IQ.) The grades 4 and 5 French control groups were similarly equated with the two anglophone groups; however, in this case, IQ scores were based on the Lavoie-Laurendeau Test Collectif d'Intelligence Générale (1960). IQ scores were not available for the grade 6 French control group. We were assured, however, by the principal and teachers of these students that they were "average" students. Most students, anglophone and francophone, came from predominantly middle-class neighborhoods.

The Test Battery

English Language Tests

1. *Achievement tests.* Appropriate levels of the Metropolitan Achievement Test (MAT) were administered to students in grades 1 to 4 inclusive; similarly, the California Achievement Test (CAT) was administered to students in grades 5 and 6. The following subtests of the MAT were administered to the grade 1, 2, and 3 students: Word Knowledge; Word Discrimination; and Reading Comprehension. The Word Knowledge and Reading Comprehension subtests were administered to the grade 4 students. The Word Knowledge, Reading Comprehension, and Spelling subtests of the CAT were administered to the grade 5 and 6 students.
2. *Peabody Picture Vocabulary Test* (English version). This test (Dunn, 1959) was group administered to grade 1 and 2 students.
3. *Word Association Test.* This test consists of a series of 51 English words, to each of which the student is requested, upon presentation, to say the first word that comes to mind. The following response classifications were used to analyse the responses. *Syntagmatic* responses are those where there is a syntactic link between the stimulus word and the response given (e.g., bird — flies; red — flower). *Paradigmatic responses* are those which (1) are of the same grammatical class as the stimulus word (e.g. red — brown), and (2) cannot be linked by a determiner to form a commonly occurring verbal sequence. Other categories included *rhyming responses* (e.g. say — day), *transformations* (kill — killed), and *idiosyncratic responses* (street — four).
4. *Listening Comprehension Test.* This test consists of a tape-recorded story, read twice, once in its entirety and a second time in sections, with each section followed by a number of true/false questions. Students in grades 1, 2, and 3 were administered different versions of this test.

5. *Decoding Test*. This test, which is group administered, is designed to assess the ability to understand spoken descriptive messages. The children listened to a series of 24 tape-recorded descriptions and selected from their test booklets the design or photograph being described. In one condition, the students listened to child speakers and in a second condition to adult speakers (cf. Samuels, Reynolds, & Lambert, 1969).

6. *Speaking skills*. Each student in the grade 1, 2, and 3 samples was asked to tell the story depicted by a sequence of cartoon pictures. The students' responses were tape-recorded and subsequently rated by a linguist along a number of dimensions: *overall fluency*, *grammatical correctness*, *enunciation*, and *rhythm and intonation* (and *liaisons* for grade 3 students only). Ratings ranged from 5, indicating "excellent," to 1, indicating "very poor." In addition, the stories were analysed according to (1) the number of words used to tell the story; (2) the percentage of words that were nouns, verbs, and adjectives; and (3) the percentage of words that were different nouns, different verbs, and different adjectives.

7. *Writing Skills*. This test was administered to students in grades 4, 5, and 6 during one 45-minute period. Topic sentences were provided to the students. Each composition was subsequently marked independently by two PSBGM teachers from the appropriate grade level along nine scales: spelling, sentence accuracy, sentence complexity and variety, organization, originality, overall quality, length, vocabulary, and punctuation. On the first six scales, the ratings range from 5, indicating excellent performance for a student of that grade level, to 1, indicating totally inadequate performance for a student of that grade level. A rating of 3 indicates a level of performance to be expected of a student of that grade level. There were only two response categories for the length and punctuation scales - "appropriate" and "inappropriate." There were three categories for the vocabulary scale - "superior," "average," and "below average." The average percentage frequency of occurrence of each rating was calculated for each of these dimensions.

The compositions were rated in a mixed random order and the raters did not know the school group of any author. The ratings of both teachers were averaged to form a single score for each student on each dimension.

French Language Tests

1. *Peabody Picture Vocabulary Test, Form B* (French Version). This test was administered to students in grades 1 and 2.

2. *Word Discrimination Test*. This test was administered to students in grades 1 and 2. The student is required to select, from among four written alternatives, the one which matches a word presented orally by the tester.

3. *Word Association Test*. This test was administered to students in grades 1 and 2. Test administration and scoring procedures were the same as those for the English Word Association Test.

4. *French Listening Comprehension Tests*. These were administered to

students at all grade levels, 1 to 6. Although the content of each test was different, test formats were the same, namely, a short story read in French followed by multiple-choice questions. The stories were read twice, with the questions following the relevant passages during the second reading. The grades 1, 2, and 3 tests used a true/false answer format and there were four answer choices in the tests for grades 4, 5, and 6.

5. *French Decoding Test*. This is a French version of the English Decoding Test. The same administration and scoring procedures were used. The test was administered to students in grades 2 and 3.

6. *Tests de Rendement en Français*. These tests were administered to students in grades 2 to 6 inclusive. The tests include a number of multiple-choice items which, depending upon the particular grade level, assess the following types of language skills: phoneme/grapheme correspondences, orthography, vocabulary, grammar, logical sequencing of ideas, and comprehension of text. Stanine norms based on the francophone student population of the CECM were analysed statistically.

7. *French Speaking Skills*. In grades 1, 2, and 3 speaking skills in French were assessed by a French linguist in the same way and with the same linguistic rating scales and word counts as in the case of English. In grades 4, 5, and 6, the test took the form of an interview between each student and a female native-French speaker. Each story and interview was tape recorded for later evaluation. The grade 4, 5 and 6 speech samples were rated by two native-French speakers, independently, using a 5-point rating scale, along five dimensions: *aural comprehension*, *pronunciation*, *grammar*, *vocabulary*, and *communicativeness*. The dimension of communicativeness was intended to assess the student's willingness to speak French and his/her ease while doing so. A rating of 5 indicates excellent or native-like performance and a rating of 1 indicates very poor performance. The ratings of both raters for each student were averaged to give one score on each dimension. Ratings were made blind with respect to the school program and IQ level of the students and were carried out in a mixed random order.

8. *French Reading Comprehension Tests*. Modified versions of the Test de Lecture "California", 4e and 5e année (Editest, Note 5) were administered to grade 4 and 5 students respectively. These are multiple-choice tests for native-French-speaking students. The grade 4 test includes five subtests: (1) discriminations visuelles, (2) antonymes, (3) synonymes, (4) directives à suivre, and (5) interprétations des textes, which yielded two scores: *vocabulaire*, which includes the scores from subtests 1, 2, and 3, and *compréhension*, which includes the scores from subtests 4 and 5. The discriminations visuelles section was deleted from the grade 5 test and therefore from the vocabulaire score. The grade 6 French reading test consisted of a modified version of the Test de Lecture Silencieuse, Niveau 6-7, Forme B (Joly, Note 6). This test is designed for French Canadian students. It uses a cloze procedure with choices.

Mathematics

- 1. *English.* Grade 1 students were administered the Arithmetic Concepts and Skills subtests of the MAT (1959) Primary I Battery. Grade 2 and 3 students were administered the Arithmetic Problem Solving and Concepts subtests of the MAT (1959) Primary II Battery and Elementary Battery respectively. Grade 4 students were administered the Concepts subtest of the MAT (1970) Elementary Battery, Form F. Grade 5 students were not administered any English mathematics test. Grade 6 students were administered the Computation, Concepts, and Problems subtests of the CAT (1970), Level 3, Form A.
- 2. *French.* Students in grades 2 to 6 inclusive were administered appropriate levels of the Test de Rendement en Mathématiques. These tests are standardized and used in French schools of the CECM. The test uses a multiple-choice format and measures a variety of mathematical skills.

Test Administration

The English language tests and the English mathematics test were administered to the immersion and English control groups. The French language tests and the French mathematics test were administered to the immersion, the English control, and the French control groups. All students in each sample were administered the tests as described with one exception: in the case of the French-speaking skills test at the Grade 4 to 6 levels, subsamples of the students were included of the following size:

	Grade 4	Grade 5	Grade 6
Immersion	39	28	51
English Control	35	35	58
French Control	31	34	41

These subsamples were selected so as to equate and stratify the language subgroups at each grade level according to IQ and sex.

The tests were administered during May and June of each academic year, 1970–76, except for the grade 5 and 6 Tests de Rendement, which were administered in November 1975 and 1976 to coincide with the standardized testing program of the CECM.

RESULTS: STATISTICAL ANALYSES

The results for each test at each grade level were analysed separately by analysis of variance for unequal sample sizes (either a least squares solution or an unweighted means solution). The main variable for the grade 1 to 3 results is school group (immersion, English control, French control) and for the grade 4 to 6 results the main variables were school group, IQ classification (above average, average, below average), and sex of student. Since there were no important sex differences in all of the analyses, no further mention will be made of this variable. Results according to IQ classification will be included only where they are particularly relevant to an understanding of second-language learning (cf. Genesee, 1976).

Results of the grade 1 to 3 English language and French language test results are summarized in tables 2 and 3. Analyses of the corresponding grade 4 to 6 tests are summarized in tables 4 and 5. *F* ratios and *df* values have not been reported here in order to simplify data presentation; significance levels have been indicated by means of the one, two, and three asterisk convention.

Grades 1 to 3

It should be noted that there was a statistically significant difference among the grade 1 groups on the Lorge-Thorndike Intelligence Test in favor of the immersion students. As well, there were significant differences among the grade 1 and 2 groups on Raven's Progressive Matrices. At the grade 1 level, the results favor the immersion group over the English control group, who similarly scored more favorably than the French control group. Thus, results on the other grade 1 tests may have been influenced by this differential; this possibility must be kept in mind when interpreting these results. At the grade 2 level, it seems that the significance is attributable primarily to the French control group's relatively low score; there is minimal difference between the immersion and English control groups. Therefore, the results of the grade 2 English language testing are probably not confounded by IQ differences, whereas the results of the French language testing may be.

The progressive change in relative IQ standing among the groups from grade 1, where there were significant between-group differences on both IQ measures, to grade 3, where there were no significant differences on either measure, probably reflects annual changes in sample size and composition.

English Language Tests

The immersion students' performance on the MAT — Word Knowledge, Word Discrimination, and Reading Comprehension — was poorer than that of the English control students in grades 1 and 2, prior, that is, to the introduction of English language arts. By the end of grade 3, the year when English language arts was introduced, the immersion students were scoring at par with the English control group.

The "language lag" which the immersion students demonstrated during grades 1 and 2, and which has been reported by others (Barik & Swain, 1976; Edwards, 1976; Lambert & Tucker, 1972), seems to be restricted to tests which require reading skills. The immersion students' performance on the remaining English language tests, which do not require literacy, is for the most part equivalent to that of the English control groups. Thus, the immersion students' oral vocabulary skills (PPVT) were at par with those of the English control group at the end of grade 1, although they were somewhat behind by the end of grade 2. The oral vocabulary skills of the grade 1 and 2 immersion students seem to develop in much the same way as those of the English control students, as assessed by the Word

Table 2 / Summary of English Language Test Results: Grades 1 to 3 (from Polich, Note 1)

	GRADE 1			GRADE 2			GRADE 3		
	Max	Immersion	English Control	Max	Immersion	English Control	Max	Immersion	English Control
1) M.A.T.:									
a) Word Knowledge	ss	46.90	57.29 ***	ss	49.41	57.31 **	ss	54.22	56.75 ns
b) Word Discrimination	ss	46.53	56.64 ***	ss	49.81	56.13 *	ss	52.84	54.94 ns
c) Reading	ss	43.69	54.64 ***	ss	46.00	52.93 *	ss	49.27	50.25 ns
2) Peabody Picture Vocabulary Test	70	47.31	46.82 ns	100	64.41	67.35 *		N.A.	N.A.
3) Word Association Test :									
a) Syntagmatic	%	.55	.55 ns	%	.66	.63 ns			
b) Paradigmatic	%	.22	.19 ns	%	.17	.15 ns			
c) Semantic cluster	%	.10	.09 ns	%	.08	.10 ns			
d) Rhyming	%	.00	.04 **	%	.00	.03 **			
e) Transformation	%	.00	.02 ***	%	.01	.03 *			
f) Idiosyncratic	%	.09	.08 ns	%	.08	.05 n.s.			
g) No response	%	.04	.04 ns	%	.00	.01 *			
4) Listening Comprehension	20	10.02	11.00 ns	20	10.91	10.29 ns	20	13.47	17.73 ns
5) Decoding :									
a) Children's Descriptions	12	N.A.	N.A.	12	8.16	7.38 ns	12	9.39	9.07 ns
b) Adults' Descriptions	12	N.A.	N.A.	12	7.07	7.06 ns	12	7.68	7.80 ns
6) Speaking Skills (Story Invention) :									
a) Linguistic Ratings									
- Overall fluency	5	2.76	2.00 ns	5	3.51	3.31 ns	5	3.18	2.60 *
- Grammatical correctness	5	2.91	2.86 **	5	3.18	3.38 ns	5	3.27	2.80 *
- Enunciation	5	2.50	2.43 ns	5	3.56	3.75 ns	5	3.62	3.30 ns
- Rhythm & Intonation	5	2.67	2.57 ns	5	3.64	3.44 ns	5	3.31	3.20 ns
b) Word Counts									
- Total no. of words		61.18	43.23 *		73.82	80.50 ns		150.55	80.14 **
- Nouns	%	.18	.22 *	%	.19	.20 ns	%	.19	.24 **
- Different nouns	%	.65	.63 ns	%	.45	.46 ns	%	.43	.56 **
- Verbs	%	.18	.17 ns	%	.16	.17 ns	%	.18	.19 ns
- Different verbs	%	.75	.82 ns	%	.69	.67 ns	%	.69	.78 ns
- Adjectives	%	.03	.03 ns	%	.07	.06 ns	%	.08	.07 ns
- Different adjectives	%	.40	.08 ns	%	.62	.62 ns	%	.69	.80 *

N.A. = Not Administered; ss = Standard Score * p<.05; ** p<.01; *** p<.001; ns = not significant

Table 3 / Summary of French Language Test Results: Grades 1 to 3 (from Polich, Note 1)

	GRADE 1			GRADE 2			GRADE 3		
	Max	Immersion	French Control	Max	Immersion	French Control	Max	Immersion	French Control
1) Peabody Picture Vocabulary Test	70	35.67	45.35 ***	100	50.78	62.33 ***		N.A.	N.A.
2) Word Association Test:									
a) Syntagmatic	%	.44	.55 ns	%	.57	.70 ***		N.A.	N.A.
b) Paradigmatic	%	.23	.12 ***	%	.17	.12 ns			
c) Semantic cluster	%	.14	.11 ns	%	.10	.08 ns			
d) Rhyming	%	.01	.02 ns	%	.00	.01 ns			
e) Transformation	%	.01	.07 ***	%	.01	.04 ns			
f) Idiosyncratic	%	.10	.07 ns	%	.11	.05 ns			
g) No response	%	.07	.06 ns	%	.04	.00 ns			
3) Word Discrimination	32	27.54	24.74 *	32	31.63	30.39 ns		N.A.	N.A.
4) Test de Rendement en Français		N.A.	N.A.	30	16.74	21.11 **	40	19.92	28.75 ***
5) Listening Comprehension	20	8.25	8.90 ns	20	10.22	13.11 ***	20	11.23	14.65 ***
6) Decoding:									
a) Children's Descriptions		N.A.	N.A.	12	5.35	6.44 **	12	6.23	9.06 ***
b) Adults' Descriptions		N.A.	N.A.	12	5.00	6.61 **	12	6.49	8.31 ***
7) Speaking Skills (Story Invention):									
a) Linguist's Ratings:									
- Overall fluency	5	3.11	3.77 ns	5	2.61	3.06 ns	5	3.41	4.72 ***
- Grammatical correctness	5	2.19	4.18 ***	5	2.20	4.61 ***	5	3.05	4.61 ***
- Enunciation	5	3.19	4.68 ***	5	2.44	4.94 ***	5	3.02	4.72 ***
- Rhythm & Intonation	5	2.53	4.68 ***	5	2.49	5.00 ***	5	2.86	4.06 ***
- Liaisons	5	N.A.	N.A.	5	2.39	4.94 ***	5	2.48	4.17 ***
b) Word Counts :									
- Total no. of words		62.13	87.00 ***		64.52	82.28 *		116.00	118.00 ns
- Nouns	%	.27	.27 ns	%	.21	.20 ns	%	.21	.26 ***
- Different nouns	%	.42	.40 ns	%	.47	.48 ns	%	.54	.57 ns
- Verbs	%	.16	.16 ns	%	.16	.17 ns	%	.18	.20 ns
- Different verbs	%	.69	.74 ns	%	.67	.61 ns	%	.62	.57 ns
- Adjectives	%	.05	.08 ns	%	.06	.05 ns	%	.09	.11 *
- Different adjectives	%	.25	.36 ns	%	.62	.60 ns	%	.50	.51 ns

N.A. = Not Administered; ss = Standard Score; ns = not significant * p<.05; ** p<.01; *** p<.001

Table 4 / Summary of English Language Test Results: Grades 4 to 6 (from Genesee et al., Notes 2, 3, 4)

	GRADE 4			GRADE 5			GRADE 6		
	Max	Immersion	English Control	Max	Immersion	English Control	Max	Immersion	English Control
1) <u>M.A.T./C.A.T.:</u>									
a) Reading Comprehension	ss	77.97	77.30 ns	42	29.03	31.24 ns	40	28.64	26.73 ns
b) Word knowledge	ss	77.02	76.90 ns	40	32.96	26.70 *	40	32.81	31.09 ns
c) Spelling		N.A.	N.A.	32	23.03	22.70 ns	32	22.93	24.50 ns
2) <u>Writing:</u>									
a) Spelling	5	3.00	3.46 *	5	3.40	3.40 ns	5	3.22	3.27 ns
b) Sentence Accuracy	5	2.61	2.42 ns	5	2.89	2.94 ns	5	2.88	3.36 ns
c) Sentence Complexity	5	2.53	2.36 ns	5	2.62	2.67 ns	5	2.74	2.93 ns
d) Organization	5	2.66	2.43 ns	5	2.81	2.77 ns	5	2.71	3.06 ns
e) Originality	5	2.20	2.01 *	5	2.10	2.08 ns	5	1.92	2.17 ns
f) Overall	5	2.57	2.37 ns	5	2.82	2.83 ns	5	2.64	2.80 ns

N.A. = Not Administered; ss = Standard Score;
 * p<.05; ** p<.01; *** p<.001; ns = not significant

Table 5 / Summary of French Language Test Results: Grades 4 to 6 (from Genesee et al., Notes 2, 3, 4)

		GRADE 4				GRADE 5				GRADE 6			
		Max		English		French		Max		Immersion		English	
1) Test de Lecture California:				Control		Control						Control	
a) Vocabulaire	52	38.17	30.83	42.55	***			28	20.21	12.09	24.90	N.A.	N.A.
b) Compréhension	21	14.78	9.54	15.54	***			30	21.33	13.92	22.47	N.A.	N.A.
Test de Lecture silencieuse		N.A.		N.A		N.A.		N.A.		N.A.		11.73	
2) Test de Rendement en Français		9		4.36		1.33		9		3.47		1.79	
3) Listening Comprehension		37		30.30		16.61		17		14.62		6.16	
4) Speaking Skills:													
a) Comprehension	5	4.97	3.32	5.00	**			5	4.92	3.67	5.00	4.50	5.00
b) Pronunciation	5	4.03	3.17	4.92	**			5	4.00	2.01	5.00	4.01	5.00
c) Grammar	5	4.49	2.64	4.91	**			5	4.00	2.02	5.00	4.13	4.88
d) Vocabulary	5	4.81	2.36	5.00	**			5	3.59	1.74	5.00	3.92	4.99
e) Communicativeness	5	4.61	3.12	4.94	**			5	3.65	2.53	5.00	3.74	4.96

N.A. = Not Administered

** p<.01, *** p<.001

1. Mean scores that have been joined by a line differ significantly from one another according to the Newman Keuls Multiple Comparisons procedure.

Association Test (Item 3). Indeed, the immersion students were somewhat advanced according to this test insofar as they used fewer rhyming and transformation responses than the English control students; these response types are generally regarded as less mature (Entwisle, 1966; Ervin, 1961). In addition, the immersion students' listening comprehension skills (Items 4 and 5) were as good as those of the English control students.

Finally, the immersion students' speaking skills (Item 6) were rated by the linguist as at par with or somewhat superior to those of the English control groups (on grammatical correctness at grade 1 and on grammatical correctness and overall fluency at grade 3). In terms of speaking style, as assessed by the word counts, the immersion students were more talkative, that is they used more words in total, although they tended to use fewer nouns (grades 1 and 3) and fewer different nouns and adjectives (grade 3) than the English control students. For the most part, it would appear that the immersion students were using their language in much the same ways as the English control students.

French Language Tests

Results from the French PPVT (Item 1) indicate that the immersion students had not developed as extensive an oral vocabulary as the French control students. At the same time, the results from the Word Association Test (Item 2) suggest that the immersion students had developed the same kinds of vocabulary skills as the French control students. Results from the Word Discrimination Test (Item 3) and from the Tests de Rendement en Français (Item 4) indicate that the reading skills of the immersion students were as advanced as those of the French control students when reading single words is considered; but they were less advanced when more sophisticated reading skills are considered, such as logical sequencing and comprehension of text, which are included in the Test de Rendement. It is interesting to note that this differential was found to reverse in the higher grades — that is, the immersion students were at par with the French control students on reading comprehension skills but somewhat behind on reading vocabulary skills.

The results from the Listening Comprehension Test (Item 5) and the Decoding Test (Item 6) indicate that, in general, the immersion students were less able to comprehend spoken French than were the French control students. It is not clear why there were no differences between the grade 4, 5, and 6 immersion students and the corresponding French control students when differences had been found in the case of grade 1, 2, and 3 immersion and French control students. Contrary to our grade 1, 2, and 3 results, Samuels, Reynolds, and Lambert (1969) did not find significant differences between grade 1, 2, and 3 immersion and French control students from the South Shore area of Montreal on the same French Decoding Test.

Let us now consider the assessments of the students' speaking skills. The

immersion students at all three grade levels were rated lower by the linguist than the French control students on most dimensions of the assessment (Item 7a). It appears that the linguist was using very rigorous criteria in his assessments when one considers that not even the French control students were rated 5 or excellent. It is not possible from these data to determine the immersion students' growth in speaking skills from grade 1 to grade 3, since the ratings for each grade level were made at different times and perhaps by different linguists; the latter information is not contained in the report.

Inspection of the word counts that were made on the speech samples (Item 7b) reveals more similarities between the groups than differences, the major difference being that the grades 1 and 2 French control students were more talkative than the immersion students. Otherwise, the immersion students used the same proportion of nouns, verbs, and adjectives, and different nouns, different verbs, and different adjectives. Additional analyses of the students' stories indicated that immersion students at both grade levels (grade 1: .11; grade 2: .06) tended to use "là" less frequently than the control students (grade 1: .81; grade 2: .72); this probably reflects the prevalence of "là" in colloquial Canadian French and its relative absence in more standard French, which would characterize the language of the classroom. The immersion students (grade 1: 5.78; grade 2: 1.44) made significantly more gender errors than the control students (grade 1: .62; grade 2: .11) as one would expect. The grade 2 immersion students' speech was characterized by a greater use of "et" and less use of pronominal appositions (4.71 and .11, respectively) than was that of the French control students (1.50 and 1.67, respectively), probably reflecting the immersion students' general tendency to simplify their speech.

Mathematics Tests

There were no significant differences between the immersion and English control students at any grade level on the MAT math tests: grade 1 — immersion 54.29, English control 51.00; grade 2 — immersion 57.06, English control 53.88; grade 3 — immersion 55.58, English control 55.06. Furthermore, there were no significant differences between the grade 2 and 3 immersion and French control students on the French math test: grade 2 — immersion 22.43, French control 22.06; grade 3 — immersion 20.02, French control 22.97.

Summary of the Grade 1 to 3 Test Results

The picture that emerges at the end of grade 3 immersion is of a group of youngsters who have developed the same competencies in all aspects of the English language assessed as similar youngsters in a regular English program despite an initial lag in English literacy skills and despite having had only one hour of English language arts instruction in school for less than a year. At the same time, the immersion youngsters demonstrated

considerable competence in the French language (reading, speaking, and listening comprehension) albeit generally not at the level of native French-speaking children of the same age. Furthermore, course instruction in French did not retard the participating students' achievement in mathematics.

Grades 4 to 6

Since the grade 4 to 6 samples (with the exception of the grade 6 French control group) were systematically equated on IQ at the outset, group differences on the language and mathematics testing cannot be attributed to IQ differences. Although the test results have not been presented by IQ subgroup, reference will be made in the text to the performance of the below-average students. This will be done in order to assess the effectiveness of the immersion program for these particular students.

English Language Tests

Analyses of the English language test results have been summarized in table 4. The only statistically significant difference between the immersion students and the English control students on the achievement tests (MAT, CAT) occurred in grade 5 on the Word Knowledge subtest — the immersion students (mean = 32.96) scored significantly higher than the English control students (mean = 26.70). The immersion students at all three grade levels performed as well as the English control students on the Reading Comprehension subtest. The grade 5 and 6 immersion students also scored as well as the English control students on the Spelling subtests. It has been found by other evaluators (Lambert & Tucker, 1972; Swain & Barik, 1976) that immersion students sometimes have a persistent problem in learning English spelling patterns long after English language arts has been introduced and after lags in other English language skills have been redressed. Evidence of this same difficulty was found in the current grade 4 writing evaluation, where the immersion students (mean = 3.00) received a lower rating on spelling than did the English control students (mean = 3.46). The results from the grade 5 and 6 language achievement testing and writing evaluation indicate that the spelling difficulty is resolved by grade 5.

The results of the writing evaluation indicate that the writing skills of the immersion students are generally comparable to those of the English control students. There were some differences between the groups on the dimensions of punctuation and length, the results of which are not included here. These differences were quite variable and therefore do not indicate any systematic problems. Similar results have been reported by Swain (1975) working with French immersion students in Ottawa.

When the English language test results were analysed according to IQ subgroups, there was no evidence that the below-average immersion students were differentially handicapped by participation in the program.

That is to say, the below-average immersion students scored at the same level and were rated as highly as the below-average English control students on the language achievement tests and writing evaluation respectively. Both the below-average immersion students and the below-average English control students scored significantly lower than the average and above-average groups. In other words, the below-average students in one program were comparable to the below-average students in the other program.

French Language Tests

Analyses of the French language test results have been summarized in table 5. The English control students at all grade levels scored significantly lower than the immersion students and the French control students on every French language measure. This is not surprising in view of the considerable difference in classroom time spent in French.

The grade 4 and 5 immersion students scored as well as the French control students on the Compréhension section of the Test de Lecture "California" and on the group and individual listening comprehension tests. On the other hand, the immersion students scored consistently lower than the French control students on the Vocabulaire section of the Test de Lecture "California" and on the Test de Rendement en Français. In terms of test norms, the grade 4 immersion students scored in the "average" stanine range on the Test de Rendement, while the grade 5 and 6 students scored somewhat "below average" on this test. The disparity between the grade 4 results and the grade 5 and 6 results is probably attributable to the earlier administration date of the grade 5 and 6 tests.

It is interesting to note that the Compréhension section of the Test de Lecture and the listening comprehension test assess integrative language skills, that is to say, language skills which involve the integration of contextual information. On the other hand, the Vocabulaire section of the Test de Lecture and the Test de Rendement assess discrete-point skills, that is, the students' knowledge of specific points of vocabulary, orthography, or grammar. Thus, these findings suggest that the immersion students are relatively more proficient in language tasks which are characterized by redundancy that can be used to compensate for gaps in their knowledge or understanding of specific linguistic rules. When the redundancy in a reading passage is reduced by deletion of lexical items, as was done in the grade 6 Test de Lecture, the immersion students can no longer retain parity with native French-speaking youngsters. In order to perform successfully on a "cloze" test, the student must use knowledge of the language as well as comprehension of the text.

The speaking skills of the immersion students, and particularly those in grade 4, were rated quite high. There were no significant differences between the grade 4 immersion and the French control students on any of the dimensions of the speaking skills evaluation except pronunciation.

The grade 5 and 6 immersion students were rated significantly lower than the French control students on most of the dimensions.

When the French language test results were analysed by IQ subgroupings, a rather interesting pattern of findings occurred. On the Test de Lecture and Test de Rendement, at all grade levels, below-average students scored significantly lower than average students, who in turn scored lower than the above-average students. These findings are not surprising, since standardized group IQ tests are designed to differentiate among students with different scholastic achievement levels. What was surprising was the failure to find a similar stratification according to IQ level on the two tests related to interpersonal communication, namely, the listening comprehension tests (group and individual) and the speaking skills evaluation. This pattern was evident at all grade levels and for students in both the immersion and the regular FSL programs. It appears, then, that students in the Early Immersion Program are differentially successful in acquiring academic or literacy skills in the second language in much the same way that students in regular first-language programs are differentially successful in traditional academic areas. Apparently, however, there is more "equipotentiality" among students in the learning of language skills which are non-academic, that is, related to interpersonal communication. These results have also been found for students at the grade 7 level in regular FSL and immersion programs (Genesee, 1976).

Mathematics Tests

There were no significant differences between the immersion and English control groups on the English math tests: Grade 4 — Immersion 84.27, English Control 82.47; Grade 6 — Computation: 35.60, 34.94; Concepts: 21.67, 20.88; Problems: 12.17, 12.48 (immersion results listed first in each case).

At the same time, the immersion students scored consistently higher than the English control students on the French math test: Grade 4 — 5.20, 2.19; Grade 5 — 5.36, 3.04; and Grade 6 — 4.53, 3.39. There was no significant difference between the immersion and French control students at grades 4 and 5 (5.20, 4.95; and 5.36, 6.40, respectively) whereas the French control students scored higher at grade 6 (4.53, 5.83). At all three grade levels, the immersion students scored in the average stanine range, as did the French control students, whereas the English control students scored in the below-average category.

Summary of the Grade 4 to 6 Test Results

The pattern of results that had emerged by the end of grade 3 was largely sustained throughout grades 4, 5, and 6. There was no indication from the English language testing, in either standardized achievement tests or a writing evaluation, that the immersion students were falling behind their peers from the regular program in native-language development.

Moreover, the immersion students were able to maintain achievement levels in non-language subjects such as mathematics. At the same time, they had acquired skills, in all areas of the French language under investigation, that were superior to skills of students following the regular French course. The performance of the immersion students was most comparable to that of native-French-speaking students of the same age on tests which required integrative language skills, such as are involved in comprehension of written or spoken material. The immersion students demonstrated considerable competence on tests which required discrete-point skills, such as the Test de Rendement and the speaking skills evaluation, but they did not attain native-language performance levels on these tests. It was noted that while acquisition of academic language skills was associated with the IQ level of the student, the acquisition of interpersonal communication skills was not.

DISCUSSION

The findings from this longitudinal evaluation of an early immersion program are largely consistent with those reported by other investigators — the characteristic lag in English language skills development prior to the formal introduction of English language instruction; the quick recovery of English language skills one year after English language arts is taught in the classroom; and the somewhat protracted difficulty immersion students have in learning English spelling patterns. On the basis of our results we are able to add that below-average students in immersion programs demonstrate the same pattern of development, although at a lower level, in their first language as do academically stronger immersion students; furthermore, relative to comparable students in the regular English program, below-average students in immersion are able to achieve comparable levels of competence in their first language.

There was no indication that academic achievement was impaired as a result of participation in the immersion program. In the current evaluations, evidence for this came from the results of mathematics testing and to a certain extent from the writing evaluations which assess school-based learning. Additional evidence comes from another evaluation which included a survey of the Quebec Secondary School Leaving Examination results for secondary school students who followed either a Grade 7 Immersion Program (Genesee, Polich, & Stanley, 1977) or a regular English Program. (See Genesee & Chaplin, Note 7.) In this survey, the performance of immersion students who wrote examinations in French for courses such as chemistry and mathematics, which were offered to them in French as post-grade-7 immersion follow-up options, was compared (1) to the performance of other English students who wrote the equivalent examinations in English and (2) to regional norms based on native French-speaking secondary school students. In Quebec parallel courses and examinations are offered in French and English at the secondary school level. In com-

parisons of the results of the immersion students with those of the non-immersion students it was found, even after the influence of IQ had been neutralized statistically by analysis of covariance procedures, that students from the secondary school immersion program scored at least as well as students from the regular program on exams in English language arts; in fact, the immersion students scored statistically higher than their peers in the regular program; and the immersion students scored as well as their peers in the regular program on other exams which assessed mastery of academic material — physics, chemistry, history. When the results of the immersion students were compared with regional norms, it was found that PSBGM students in special post-grade-7 immersion French courses scored higher on the average than their francophone peers in the rest of Quebec writing the same examinations. These examinations included *histoire, géographie, mathématiques, dactylo*, and *arts plastiques*. Interpretation of these latter results must, however, be qualified in accordance with the variability of socio-economic, academic, and geographical factors in the groups under discussion.

Before we leave the issue of academic achievement, it should be noted that in the present study the performance of below-average immersion students was comparable to that of below-average English control students in mathematics, albeit lower than that of average and above-average students; this finding indicates again that these students are not differentially handicapped as a result of the program.

The immersion students demonstrated considerable competence in the French language, particularly in tasks which involve integrative-type language skills. In fact, the immersion students scored at the same level on these types of tests as did comparable groups of French-speaking students. The immersion students' performance on tests of discrete-point language skills was less comparable to that of the French control groups, indicating a somewhat greater difficulty in generating language than in decoding it. It was also found that the students' IQ level was differentially associated with the level of language acquisition such that while acquisition of literacy or academic language skills in French varied with IQ level, acquisition of interpersonal communication skills did not seem to. These findings are consistent with those of other researchers who have found that IQ is not necessarily the only or most important determinant of second-language acquisition (Dockrell & Brosseau, 1967; Gardner & Lambert, 1972).

We are continuing to monitor the development of these students as they progress through secondary school, where most of them will take special maintenance courses in French.

NOTES

This research was made possible by grants from the Ministry of Education, Quebec, to the Protestant School Board of Greater Montreal. I would like to express my thanks

to Dr. E. G. Cochrane, Director of Instructional Services, and to Mr. F. Méloul, Coordinator of Modern Languages, PSBGM, for their comments on a draft of this report.

REFERENCE NOTES

1. Polich, E. *Report on the evaluation of the lower elementary French program through grade 3*. Montreal: The Protestant School Board of Greater Montreal, Instructional Services Department, 1974.
2. Genesee, F.; Morin, S.; & Allister, T. *Evaluation of the 1973-74 grade 4 French immersion class*. Montreal: The Protestant School Board of Greater Montreal, Instructional Services Department, 1974.
3. Genesee, F., & Chaplin, S. *Evaluation of the 1974-75 grade 5 French immersion class*. Montreal: The Protestant School Board of Greater Montreal, Instructional Services Department, 1975.
4. Genesee, F., & Stefanovic, B. *Evaluation of the 1975-76 grade 6 French immersion class*. Montreal: The Protestant School Board of Greater Montreal, Instructional Services Department, 1976.
5. Editest, *Test de Lecture "California", 4e et 5e année*. 94, rue Général Capiaumont, Bruxelles, 1967.
6. Joly, J. M. *Test de lecture silencieuse, Type CR, Niveaux 6-7*. Centre de psychologie et de pédagogie, 260 ouest, rue Faillon, Montréal, 1964.
7. Genesee, F., & Chaplin, S. *Evaluation of the 1974-75 grade 11 French immersion class*. A report submitted to the Protestant School Board of Greater Montreal, 1976.

REFERENCES

- Barik, H., & Swain, M. Primary-grade French immersion in a bilingual English-Canadian setting: The Toronto Study through grade 2. *Canadian Journal of Education*, 1976, 1(1), 39-58.
- Dockrell, W. B., & Brosseau, J. F. The correlates of second language learning by young children. *Alberta Journal of Educational Research*, 1967, 13, 295-298.
- Dunn, L. M. *Peabody Picture Vocabulary Test*. Minnesota: American Guidance Service, 1969.
- Edwards, H. P. Evaluation of the French immersion program offered by the Ottawa Roman Catholic Separate School Board. *Canadian Modern Language Review*, 1976, 33, 137-150.
- Entwisle, D. Form class and children's word associations. *Journal of Verbal Learning and Verbal Behavior*, 1966, 5, 558-565.
- Ervin, S. Changes with age in the verbal determinants of word association. *American Journal of Psychology*, 1961, 74, 361-372.
- Gardner, R., & Lambert, W. E., *Attitudes and motivation in second language learning*. Rowley, Mass.: Newbury House, 1972.
- Genesee, F. The role of intelligence in second language learning. *Language Learning*, 1976, 26, 267-280.
- Genesee, F.; Polich, E., & Stanley, M. An experimental French immersion program at the secondary school level — 1969 to 1974. *Canadian Modern Language Review*, 1977, 33, 318-332.
- Lambert, W. E., & Tucker, G. R. *The bilingual education of children: The St. Lambert experiment*. Rowley, Mass.: Newbury House, 1972.
- Samuels, M.; Reynolds, A.; & Lambert, W. E. Communicational efficiency of children schooled in a foreign language. *Journal of Educational Psychology*, 1969, 60, 389-393.
- Swain, M. Writing skills of grade 3 French immersion pupils. *Working Papers on Bilingualism*, 1975, 7, 1-38.
- Swain, M., & Barik, H. A large scale program in French immersion: The Ottawa Study through grade three. *ITL, A Review of Applied Linguistics*, 1976, 33, 1-25.

Fred Genesee is Research Officer, Instructional Services Department, The Protestant School Board of Greater Montreal, 6000 Fielding Avenue, Montreal, P.Q. H3X 1T4.