NOTICE

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

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

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

Previously copyrighted materials (journal articles, published tests, etc.) are not filmed.

Reproduction in full or in part of this film is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30. Please read the authorization forms which accompany this thesis.

THIS DISSERTATION
HAS BEEN MICROFILMED
EXACTLY AS RECEIVED

AVIS

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

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

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

Les documents qui font déjà l'objet d'un droit d'auteur (articles de revue, examens publiés, etc.) ne sont pas microfilmés.

La reproduction, même partielle, de ce microfilm est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30. Veuillez prendre connaissance des formules d'autorisation qui accompagnent cette thèse.

LA THÈSE A ÉTÉ
MICROFILMÉE TELLE QUE
NOUS L'AVONS REÇUE

SYNTACTIC DEVELOPMENT IN 'IMMERSION' AND 'NON-IMMERSION' TEN YEAR-OLDS

by

Ann Stuart Laubstein

B.A. (Hons.), Carleton University, 1978

A Thesis

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Master of Arts

Department of Linguistics

University of Ottawa

April 1980

UNIVERSITÉ D'OTTAWA
UNIVERSITY OF OTTAWA

CONTENTS

LIST OF TABLES

LIST OF FIGURES

Acknowledgement

I would like to thank my thesis supervisor Dr. P. G. Patel for his valuable guidance throughout, and, just as important, for his kindness.

Special thanks are also extended to Dr. Frances Morrison and members of her staff at the Research Centre of the Ottawa Board of Education.

CHAPTER I

INTRODUCTION

1.1 Preliminary Remarks

Conflicting results as to the effects of bilingualism on language development, and more generally, on cognitive development, have been plausibly accounted for (Cummins, J.) in terms of the 'additive' and 'subtractive' learning situations.[1] A dominant culture learning the second language out of choice characterizes the 'additive' situation. In the 'subtractive' case the individual is forced by his own social situation to learn the 'élite' language -- often at the expense of his own language. The French Immersion programmes in the Ottawa Public and Separate School systems are 'additive' situations. Recent studies (Barik, H. and Swain, M.) (Lambert et al.) suggest that in this latter type of situation immersion programmes have 1) positive effects on cognitive growth in general and 2) no lasting deleterious effects on language development. If this is indeed the case it has important implications, both theoretical and practical.

1.2 The Research Problem.

Are the development of language and other cognitive faculties separate, independent processes or do they interact? Are the answers to these questions the same for both first and second language development? That is, do first and second language

---

1 Cummins limits positive effects to those individuals who have reached a certain 'threshold' -- this being a minimum level of competence.

learning involve different cognitive processes, and do they interact differently with other cognitive domains? In 'additive' situations do second language learning and other aspects of cognitive growth interact with each other to each other's benefit? Are the answers to these questions dependent upon factors like age, sex, IQ and level of competence in the language being learned? At a practical level, questions like these must be answered if our schools are to develop curricula geared to the maximum development of our students at the individual level.

These are only a few of the myriad questions dealing with the relationship of first and second language learning to cognitive growth in general. We are still at the stage where the goal is to formulate the appropriate empirical questions so that in the end attempts may be made to answer them. This study attempts to focus on a small part of the more general problem. Namely, on the effects of the immersion programme on first language syntactic development. To do this, it compares some aspects of syntactic development in Immersion children and their Non-immersion counterparts.

Questions concerning the effects of second language learning on first language acquisition would perhaps be more tractable if researchers agreed on what they considered the pertinent areas of language development to be. For instance, spelling, punctuation and reading tests would seem to be measuring peripheral areas of language development in the sense that one can be 'fully competent' linguistically and yet illiterate. It would seem that any effects of immersion on first language growth might be

most discernible if one were to look at the three components characteristic of all languages, in isolation. It seems possible that the phonological, lexical and syntactic components might be affected in different ways by second language learning, and that these various effects may depend on the first and second languages concerned. Learning French might augment an English lexicon but it probably would not affect a Chinese lexicon in this way. A child in the Immersion programme might know the meaning of the word 'ameliorate' because of the fairly common French word 'améliorer' and his English counterpart probably would not. There is also the possibility of lexical interference where words mean different things in the two languages but are orthographically the same -- for example 'sensible'. If learning French is going to affect the development of English its effects are going to be substantively different for the three major areas and therefore it seems reasonable to look at the three areas separately.

Adding new items to the lexicon is presumably something speakers never lose the capacity for, but what of phonological and syntactic rules? For how long can a speaker go on adding, deleting or modifying these types of rules? The answer to this is not known. For this reason it seems salutary to examine these two areas for any effects of the immersion programme on first language development. The present study has chosen syntax.

1.3 Background: Tremaine's Study (1975)

Tremaine's study of Immersion and Non-Immersion children is the only published Canadian study which has focused entirely

on syntax. She used syntactic comprehension tests to study two questions; one being the relation between Piagetian operational level and syntax, and the other being whether exposure to French improves the comprehension of English syntax. In both cases her results were in the affirmative. The latter was support for the Edwards and Casserly (1973) results that, not only do Immersion programmes benefit second language learning but also first.

Tremaine's study was interesting in that it could be replicated. There were some weaknesses in the design, however. As she herself pointed out, her random samples from the two groups could not be considered as unbiased because children who could not cope (p. 15) in the Immersion programme had been shifted into the English programme, and so in some sense the Immersion and English children were not equal to start with. This was confirmed by the fact that there were more children classed as 'operational' in the Immersion group. In her statistical analysis she controlled for this effect but as her sample size was only 5 boys and 5 girls from each group at the three grade levels this would make the number of cases to compare even smaller. Another weakness was that she did not measure IQ. Various studies have linked language development and IQ (Peal, E. and Lambert W.,) (Kessler, C.) and controlling for this variable might have been more valuable than controlling for 'operational'. She also had children from grades 1, 2, and 3 and it is conceivable that there might have been interactive effects of age and IQ. And, too, her subjects were all tested individually and so the size of the sample had to be kept down

but a design which would permit a larger more homogeneous sample might be preferable. The small size of her sample (she only tested part of her group on this structure) led her to posit a base order of Goal Objective[1] in the Ottawa dialect. This was to account for the fact that her subjects had understood the structure which has the indirect object before the direct. For instance 'Bobby showed the dog the cat'. Kessler (1971) testing the same structure had found that children had mistaken 'the dog' for the Objective Case. With a sample of 108 children this study tested the same structure and found that over 70% of both Immersion and Non-immersion children in Ottawa did not understand the structure. That is, using a larger sample (which was homogeneous in terms of grade and age) the results matched the Kessler results. Increasing the number of structures tested would also lead to greater prediction accuracy with regard to overall syntactic competence. (Tremaine's study consisted of 11 structures.) One wonders too if one can generalize about the effects of immersion on first language with such a young and varied group. Her subjects ranged from grades one to three. First, it seems possible that the effects of the programme might be different for the different age levels. Not only because of the individual differences in levels of maturation but also because the children would have spent different lengths of time in the programme and these two things might interact with each other. To recapitulate, it

---

1 Tremaine and Kessler both analyse their structures from a Case Grammar point of view.

was felt that the nature of Tremaine's sample was not conducive to generalizing about the effects of Immersion on native language syntactic growth. There were only 5 boys and 5 girls from each group at the three grade levels (and as mentioned these were perhaps not equal to start with). Controlling for grade level left you with a sample size 10 (and if sex differences are part of experimental error the sample size was 5). Controlling for grade level seems essential in view of the possible interactive effects of age, mental ability and amount of exposure time to the French Immersion programme. In designing the present study these shortcomings were taken into consideration.

1.4 The Present Study

1.4.1 The Dependent Variable

The aim of the present study was to test for the effect of the French Immersion Programme on the development of English syntactic rules. Children in the English and Immersion Programmes (the control and experimental groups, respectively) were to be compared on some variable which was considered to be an appropriate measure of syntactic competence.

'Competence' is rather a loaded term but it is useful for our purposes. In the broad Chomskyan sense 'competence' is that internalized set of rules in the idealized speaker-hearer which relates sounds to meanings. Syntactic rules are only a subset of the whole. We shall use 'syntactic competence' to refer to that set of rules which relate syntactic configurations to their meanings. Clearly language does not exist in a vacuum; it is used in context whether it be a monologue,

dialogue, or whatever. Nevertheless, it seems useful to distinguish between what we shall term 'linguistic' and 'communicative' competence. Loosely, the latter would be all those rules which appropriately encode or decode all linguistic and non-linguistic cues. Linguistic and communicative competence exist in varying levels in different individuals. Second language learners can be quite clear examples of this. Rather paltry levels of linguistic knowledge sometimes accompany quite high degrees of whatever else makes up 'communicative competence' -- resulting in more or less effective communication. At the other end of the scale one can imagine that if Henry James had talked the way he wrote he might have often failed to communicate with mortals less linguistically competent than he. In other words the correlation between linguistic and communicative competence need not be high in any one individual.

For this reason a measure of syntactic competence was wanted which would limit the intrusion of non-linguistic cues as much as possible. Since language is part and parcel of the extra-linguistic world this seems a highly unnatural way of looking at language. However, the question of whether there is a cognitive structure made up of purely linguistic rules as opposed to 'rules for use' is an empirical one. We are assuming that there is such a structure and that support for its existence is the appropriate decoding of items where only linguistic information is available. Following naturally from this is the assumption that if a structure cannot be decoded

using only linguistic information then there is no linguistic rule for that structure. In other words, rather than assume that performance factors have intervened to block the execution of some existing rule, we would conclude that no rule existed. With regard to this study, then, we are going to compare two groups of individuals in order to see whether the set of syntactic rules is the same for each group and we have suggested that, in order to ensure that our dependent variable is measuring linguistic and not communicative competence, contextual clues be omitted. More specifically, we want our instrument to measure one small part of linguistic competence, namely, syntactic competence. What would be a valid measure from a substantive point of view? That is, which syntactic structures would be the most likely to show up the differences in our comparison groups? Presumably, 'complex' structures are learned late relative to others and since our sample was a group of ten and eleven year-olds 'complex' structures seemed appropriate. 'Complex' in what sense?

Complexity was operationally defined. Those structures were deemed complex which had been found to be complex by different researchers on the basis of different measures. Some of the different measures used were, for instance, structures

(i) rarely or incorrectly used,

(ii) complex relative to other structures according to a response time metric,

and/or (iii) those structures not understood in comprehension tasks.

No attempt has been made to give reasons for the complexity of the items chosen. The 'whys' are interesting and important but

irrelevant in terms of the present study. The area of difficulty was (broadly) considered to be syntactic. In certain cases this is certainly a question of one's point of view. For instance in Sentence no. 5 (S5) 'Starbuck believed Apollo not to have been examined by the doctor' the lexical item 'believe' might be considered the culprit. We have chosen to consider this use of the non-finite clause, rather than the finite analogue viz. 'that Apollo had not been examined by the doctor' to be a syntactic matter. Without trying to justify them the following have been considered matters of syntax: (The test sentences exemplifying the areas have been noted (see App. D).)

(i) the omission of closed system items, (S6, S10, S14)

(ii) the reordering of items, (S9, S2)

(iii) the use of non-finite clauses instead of their finite analogues (with a subordinating conjunction), (S5, S7, S11, S12, S13, S20, S21)

(iv) auxiliary items (tense, voice, modal, aspect) (S1, S8, S16, S18, S22)

(v) the use of adverbs in a main clause to replace subordinate conjunctions (S19)

(vi) adverbial scope (S15, S4)

(vii) pronomial scope (S8)

(Various sentences include more than one of the areas cited. For example all of those in area (iii) would also come under area (i).) Of the 22 test sentences only one is not here accounted for, S3. It was used as a control for S10. Also, some reservations about S4 and S18 should be noted here.

S4 is a double negative 'Susan doesn't not like cake.' This may be conceptually rather than syntactically complex.

That is, if it is difficult to get at the meaning of some structure in one form, but the meaning is not a problem in some other form, then the complexity is not conceptual and might well be syntactic. For example, in 'the boy (who was) offered...' the verb is clearly passive and subordinate without the deletions but otherwise it must be disambiguated later in the sentence. Presumably the meaning is clear when there are no deletions. The syntactic form is posing the problem. In the case of the double negative however, negative (she negative (like cake)) does not equal positive (she like cake). Nevertheless it seems possible that the conceptual complexity would only exist for those who in fact had a syntactic rule which related the positve and negative forms. Those who decode it as a simple negative would be doing so because they had not found it any more complex than the simple negative which they thought it was. For this reason the structure was considered to be a test of syntactic competence.

S18 and S22 are structurally identical. 'Andy had watched T.V. for a couple of hours when he went to bed' (S18) and 'Christopher had studied for three hours when he went to bed' (S22). This study was piloted in January 1979 to ensure that the task itself would not mask any experimental effects. Ten Grade 5 children were tested on the same structures twice, over a one week period. Different lexical items were used for the two tests. S18 had a higher error rate than S22. It was felt that this might have been due to the two 'd's interfering with comprehension in S18 because the 'had' was pronounced 'əd' in both cases. In other words we had

phonological complexity added to syntactic. So perhaps S18 is not only a test of syntax.

1.4.2 The Task

As mentioned above the test was piloted in January. When people are being tested there is always the possibility that some aspect of the task itself will interfere, will add to experimental error. We wanted to be sure that when there was an incorrect response it was due to a lack of comprehension of the syntactic structure and not because 1) the task was too difficult or too long (ten year-olds do not like being bored) or 2) the vocabulary used was unfamiliar.

First. Why use a comprehension task and not a production task? Lambert et al. (1973) partly measured language development in the Immersion and Non-immersion groups with an oral expression task where a marker graded children on a scale of 1-5 -- a production task. Kessler (1971) and Tremaine (1975) used pictures to test comprehension of various syntactic structures. There is no evidence that levels of development in production and comprehension (encoding and decoding capacities) are equal. Nor is it clear which is the more appropriate measure in different areas of language. However, using a production task to determine level of development is fraught with a number of problems. Subjective measures (scales) from markers are unreliable and cannot be replicated. With objective measures it is difficult to specify the criteria; word counts? clause counts? auxiliary counts? And further, just because some structure has not been produced there is no

reason to assume it cannot or will not be. Comprehension tasks can be devised where the criterion is unambiguous; namely, has X been understood by Y or not? (Your scope, of course is limited to those structures you test.)

For these reasons we have used a comprehension task and as mentioned earlier we will assume that there is comprehension of a given structure if, and only if, (iff) there is a rule for that particular structure. (This involves the assumption that performance factors are randomly distributed in both the Immersion and Non-immersion groups and thus can be ignored. In other words, it is quite clear that here and there on the test, a child will choose the correct or incorrect sentence as a paraphrase choice whether he understands the examiner's sentence or not. That is, by chance he will be right or wrong. We assume that this will be true of both groups and hence, does not matter.)

To help prevent any effects from boredom the test was kept short (less than 20 minutes) and the sentences themselves were kept short so that short-term memory would not be over-burdened. The 'complex' structure was read out loud by Examiner but the two paraphrase choices were only read silently by the children, they were not heard. It was felt that the 'complex' sentence would be remembered better if it were the only one read aloud.

1.5 Independent Variables Controlled for

Various studies in the past decade have indicated that there are correlations between language development and

1) socio-economic class-home environment (McCarthy, D., 1962),

2) IQ (Peal, E., and Lambert, W., 1962),

3) age (Loban, W., 1976) and

4) sex.

With regard to the latter, 'it is well established that girls are more advanced than boys' (Peal, E., and Lambert W., p. 1, 1962). We were unable to control for home environment. However, verbal IQ, age, and sex were all controlled. These were controlled on an individual, rather than on a group basis. For example, if a child from the random sample of French Immersion children happened to be a boy, 10 years old, with a verbal IQ of 110, then a boy was found from the English programme who matched the French boy on these variables. This was to control for possible interactive effects of age, IQ, and sex. We believed that an increase in age by 6 months (we actually controlled at 3 month intervals) in a child with a high verbal IQ might affect performance on the test differently from the same increase in age in a child with a low IQ. Therefore, to control for these variables on a group basis might prove inadequate. The mean age and IQ might be the same in both the English and French groups, but the high IQ's might be young in one group and old in the other. As it turned out it was not far-fetched to consider that high IQ's might not be normally distributed throughout the age range. In our two random groups (French girls and French boys) it happens that high IQ's are more frequent at the lower age levels. Since we 'paired' in the above sense this is also

true of the English groups. What we have done, then, is to match each of the 30 girls and 24 boys from the Immersion group with their 'twin' in the English programme on the three variables, age, sex, and verbal IQ. The children were all from Grade 5 because we felt that any real effects of Immersion were more likely to be discernible at this age than at the young age used by Tremaine. The group was also homogeneous in that all the children had had 6 years in the programme.

With these variables controlled in this way, the hypothesis was that there would be no difference in the performances of the English and French groups on the dependent variable and that we could conclude that there was no reason to assume that the French Immersion programme had any effects on the syntactic development of English (neither good nor bad.) We expected our results to conflict with the results obtained by Lambert et al. (1973), Barik, H. and Swain, M. (1975) and Tremaine, R. (1975).

The experiment was designed specifically to test the above hypothesis. But added to this were two areas of interest which could be called exploratory. Results of various studies have been indicative (as mentioned earlier) of relationships between language development and our controlled independent variables, age, sex, and verbal and non-verbal IQ. We wanted to look at these relations in the whole group. This was not strictly legitimate because the members of our group are not independent of each other because of the matched pairs design. Similarly, were relationships between the variables the same

for each group? For instance would there be a significant correlation of test score and age in one group but not in another? Would the relationships be the same for girls as for boys? The second peripheral area of interest involved the specific structures tested. Would the two groups, Immersion and English differ significantly on particular items? Nothing definitive could be concluded if such differences existed because, in a narrow sense, there was only one example of each of the sentence types. Differences would be suggestive however, and indicate areas for further more detailed research.

1.6 Summary

Within the more general area of language's relationship to cognitive development we have identified the specific area of interest for this study viz. syntactic development's relation to Immersion. Reasons were given for choosing our dependent variable -- 22 syntactically complex sentences. And for believing that using the appropriate task the dependent variable was a reliable measure for comparing syntactic competence in the English and French groups. Finally reasons were given for using a sample of matched pairs. It was predicted that with the chosen dependent variable and using matched pairs we would find no reason to assume that French Immersion positively or negatively affects English syntactic development.

CHAPTER II

METHOD

2.1 Design

Syntactic comprehension tests were given to a random sample of children from the French Immersion programme matched with their 'twins' from the English programme in the Ottawa Public School System. Results of the tests were then analysed in order to compare the levels of English syntactic development in both groups.

2.2 Subjects

The children were 108 grade five pupils (54 from Immersion and Non-immersion classes) from five Ottawa Public schools. Two considerations determined the selection of schools; whether the school had both Immersion and Non-immersion classes (to allow testing of both groups together) and whether the school principal could schedule the testing in an already overloaded testing period -- June. Pupils were then chosen at random from the Immersion classes of the five suitable schools. The thirty girls and twenty-four boys thus selected[1] were then matched with controls from the schools' regular classes on the basis of age (within three months), verbal IQ (within four points on the Lorge-Thorndike test) and

---

1 Actually 33 of each were originally selected to allow for absences and the like, with the intention of getting 30 of each.

sex. Social and ethnic background were not considered although the language spoken at home was noted in all cases. There were, therefore, two major subgroups -- girls and boys from the English programme (54), and girls and boys from the French Immersion programme (54), and four minor subgroups -- English girls (30), English boys (24), French girls (30) and French boys (24). Descriptive statistics of each variable for each group appear in Appendix A.

2.3 Materials

The test consisted of twenty-two 'syntactically complex' sentences (keeping in mind the reservations in this regard discussed above). They were relatively short (that is, short-term memory was not overloaded) with simple vocabulary and some proper names chosen from children's current 'folk heroes'. (See Appendix D2.) For each sentence there were two paraphrase choices-- which of the two was the correct choice (a) or (b) was randomly decided. The lexical items were kept relatively more invariant in the incorrect paraphrase so that any strategy such as 'map morpheme to morpheme' would result in the wrong choice.

2.4 Procedure

Immersion and Non-immersion children were tested together in groups of not more than twenty in five different Public schools during the first week in June 1979. The time of day was determined by the school principals but as the experimental and

the control groups were tested at the same time neither group suffered in this respect.

All the children were told that they were representing their respective groups Immersion and Non-immersion, and that the test was being used to compare their knowledge of English. Examiner then explained the task and the children were given two practice examples to ensure that they understood. (See Appendix D1.) Examiner read the example. All sentences read by Examiner were only heard by the children and not seen. They were read only once and with slightly exaggerated, but the appropriate, intonation for the intended interpretation. Each child had a copy of two typed paraphrases (a) and (b), and was asked to circle the one he felt was closest in meaning to Examiner's. (Of all groups tested only one child made the wrong choice on practice example 1 because she had misunderstood the instructions. All children were correct on example 2.)

After Examiner read each test item the children were given as much time as they needed to make their selections. The complete test took less than twenty minutes and all groups commented on how easy they found it.

2.5.1 Statistical Analysis

Each of the 22 test items was considered as a separate random variable with a possible value of 1 for a correct response, or a value of 0 for the wrong answer. The total test score was the 23rd variable with possible scores running from 0 to 22. Scores on these variables were then used to test for the effects of sex, age, verbal and non-verbal IQ and Immersion.

The 108 subjects were classified according to the purpose of the particular test. The tests were aimed at comparing the relationship between the independent and dependent variables in each of the two groups -- Immersion and Non-immersion, and at comparing group performance on the dependent variable.

2.5.2 T-Tests

The random sample of 54 Immersion students were considered as paired with their Non-immersion counterparts and 23 paired T-Tests (with 54-1 degrees of freedom (d.f.)) were used to test for differences in means on each of the test sentences and on the total test score. Means for Immersion and English girls and for Immersion and English boys were compared in the same way (with 30-1 and 24-1 d.f. respectively). (See Table 2.1.)

Table 2.1

T-Tests for differences in means

Groups		Variables	
Classification Criteria	No. of cases		
Gr. 1 English subjects	54		
Gr. 2 Immersion subjects	54		
	108		
Gr. 1 English girls	30		
Gr. 2 Immersion girls	30	Type	No.
	60	Sentence$_i$ i=1...22	22
Gr. 1 English boys	24	Total test score	1
Gr. 2 Immersion boys	24		23
	48		
Gr. 1 Girls	60		
Gr. 2 Boys	48		
	108		
Gr. 1 English girls	30		
Gr. 2 English boys	24		
	54		
Gr. 1 Immersion girls	30		
Gr. 2 Immersion boys	24		
	54		

Children from both groups were classified according to IQ level (see Table 2.2) and Paired T-Tests were run to compare mean test scores in the French and English group at each of these IQ levels. (See Table 2.3.)

Table 2.2.

Variables recoded into high and low classes

	Verbal IQ	Non-verbal IQ	Age	Test Score
High	$\geq$ 116	$\geq$ 116	$>$ 11.0	$\geq$ 14
Average	100-115	100-115		
Low	$\leq$ 99	$\leq$ 99	$\leq$ 11.0	$\leq$ 13

Table 2.3

T-Tests for differences in means

Groups		Variables	
Classification Criteria	No. of cases		
Gr. 1 English Low IQ	12		
Gr. 2 Immersion Low IQ	12		
	24		
Gr. 1 English Average IQ	28	Type	No.
Gr. 2 Immersion Average IQ	28	Total test score	1
	56		
Gr. 1 English High IQ	14		
Gr. 2 Immersion High IQ	14		
	28		

Note: These tests were run for the Immersion and English groups as a whole and also for the separate sexes.

Unpaired tests were used to compare boys' means on each of the test items and test score to the girls' means on the same variables. This was done for the group as a whole and within each of the two subgroups (106 d.f. and 52 d.f.). (See Table 2.1.)

The children were classified into groups for each of the 22 test sentences depending on whether they were right (Gr. 1) or wrong (Gr. 2). T-Tests were run for differences in means for verbal and non-verbal IQ and for age. (See Table 2.4.)

Table 2.4

T-Tests for differences in means

Groups		Variables	
Classification Criteria	*No. of cases		
Gr. 1 Test score $\geq$ 13	67		
Gr. 2 Test score $<$ 13	41		
	108		
Gr. 1 Correct on sentence$_i$ (i=1,...,22)	*See App. B, Table 3.	Type	No.
Gr. 2 Incorrect on sentence$_i$ (i=1,...,22)		Verbal IQ	1
		Non-Verbal IQ	1
		Age	1
			3
Gr. 1 Correct on sentence$_i$ English	*See Figure 3.1		
Gr. 2 Incorrect on sentence$_i$ English			
Gr. 1 Correct on sentence$_i$ French	*See Figure 3.1		
Gr. 2. Incorrect on sentence$_i$ French			

Total scores were taken as a classification criterion -- Gr. 1 being all subjects with 13 or better on the test and Group 2 those with less than 13. Again the means tested were IQ's and age. (See Table 2.4)

All the children were grouped on the basis of test score again but this time within the Immersion and English groups ($\geq$14 versus $<$14) and the mean IQ's and ages of the two groups were compared. That is those children with high scores in the English group were compared on these variables to the high scoring children in the French group. (See App. B, Table 4.)

At each of the three verbal IQ levels there are children who do well ($\geq$14) and those who do not ($<$14). The mean ages and non-verbal IQ's of those who did well were compared to the ages and non-verbal IQ's of those who did not at each of these verbal IQ levels.

2.5.3 Chi Square Tests for Independence

Chi Square tests were run to test for dependencies between particular test items and verbal and non-verbal IQ, sex, and age and similarly for relationships between test scores and these same variables. The tests were done both between and within groups. (For statistical hypotheses and details of T-Tests and Chi Square tests see App. B, Tables 1-5.)

2.5.4 Pearson's 'r'

Pearson's 'r' and 'r^2' were computed in order to test the strength and the direction of the relationships between test scores on the one hand and age verbal and non-verbal IQ's on the other. These statistics were computed for the group as a whole and within the subgroups.

2.5.5 Regression analysis

Test scores were regressed on age, verbal and non-verbal IQ assuming the existence of a linear relationship. The statistical model assumed was:

$$Y_i = B_o + B_1X_{1i} + B_2X_{2i} + B_3X_{3i} + B_4X_{4i} + E_i \qquad i = 1,\ldots,108$$

where Y = test score

X_1 = age

X_2 = verbal IQ

X_3 = non-verbal IQ

and $X_4 = X_1X_2X_3$

CHAPTER III

RESULTS

3.1 Introduction

Statistical tests were run to test for the effects of the independent variables age, sex, verbal and non-verbal IQ and Immersion on the dependent variable, syntactic comprehension tests. For ease of reference the various statistical tests will be considered in the order in which they appear in Section 2,5, tabulated in Tables 2.1 - 2.4 inclusive.

3.2.1 T-Tests for Means on Test Items and Total Test Scores

The mean scores on the test for each group are fairly close as you can see from Table 3.1. The French girls have the highest mean -- 12.17 and the English boys have the lowest -- 11.33. The English boys also have the lowest maximum with 14. Both the Immersion boys and girls have top scores of 18 and the English girls reach a maximum of 17. (For frequency of errors on each separate test sentence for the different groups refer to Figures 3.1 - 3.4.)

TABLE 3.1

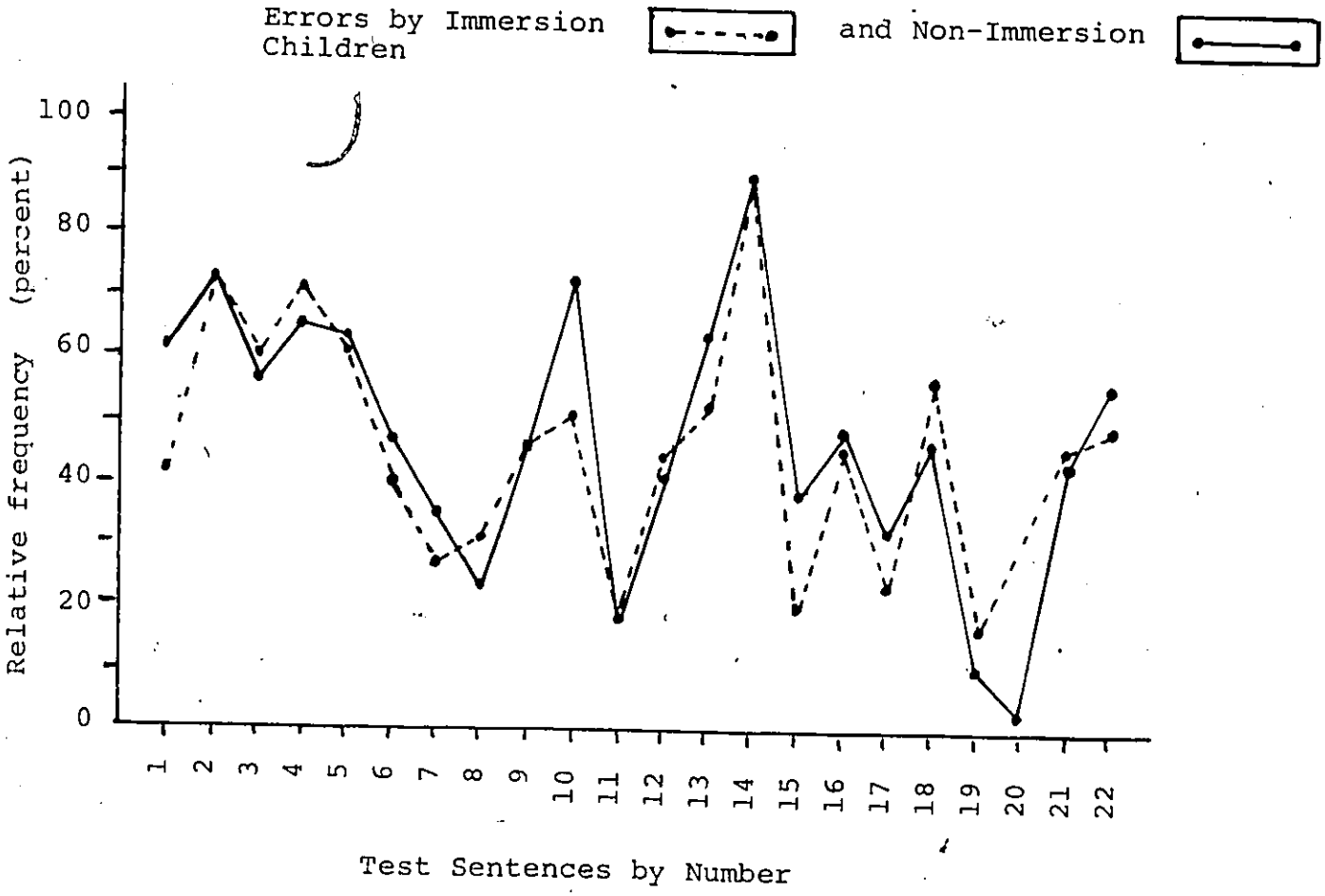
Descriptive statistics of total test for each group.
(Possible no. correct = 22)

Group	Mean	Standard deviation	Minimum	Maximum
English girls	11.43	2.46	7	17
French girls	12.17	2.68	7	18
English boys	11.33	1.10	7	14
French boys	12.00	2.73	7	18
English girls/boys	11.39	2.24	7	17
French girls/boys	12.09	2.68	7	18
All children	11.74	2.49	7	18

Comparing the Immersion and Non-immersion groups using a two-tailed paired T-Test (alpha = .05), of the 22 structures tested there is a significant difference in means in the case of three sentences. (The result is the same with a one-tailed test.) The Immersion children perform better on sentences 1, 10 and 15 (S1, S10, S15).

They have slightly higher means in 63% of the items but this proportion is not significant ($p. > .10$). Figure 3.1 is a graphic representation of the errors made by both groups. In S1, S10 and S15 the Immersion children make about 20% fewer errors than their English counterparts.

Figure 3.1



Note: (1) Each group includes 30 girls and 24 boys

The mean number correct out of 22 is 11.39 for the English children and 12.09 for the Immersion girls and boys. This difference is not significant (although with a paired One-tailed test it is ($p < .048$).

The paired test was used to compare the English to the Immersion girls' means. (Again using a two-tailed test with alpha = .05.) The Immersion girls do significantly better on S7 and S17 whereas the English girls have a significantly higher mean on S18. In this case the English girls have a slightly higher proportion of higher means--52% versus 48%.

This difference and the difference on the total test score are not significant. The total test mean for English girls was 11.43 and for Immersion girls, 12.16.

The boys' means were 11.33 and 12.00, for English and Immersion respectively. Both groups had exactly 50% with higher means so on the basis of this and the latter statistics there is no reason to reject the null hypothesis. The only significant difference among the boys is S15 where the Immersion boys are better at the .005 level of significance. Figures 3.2 and 3.3 show the boys' and girls' error rates respectively.

Figure 3.2

Boys' Errors: Immersion  , Non-Immersion 

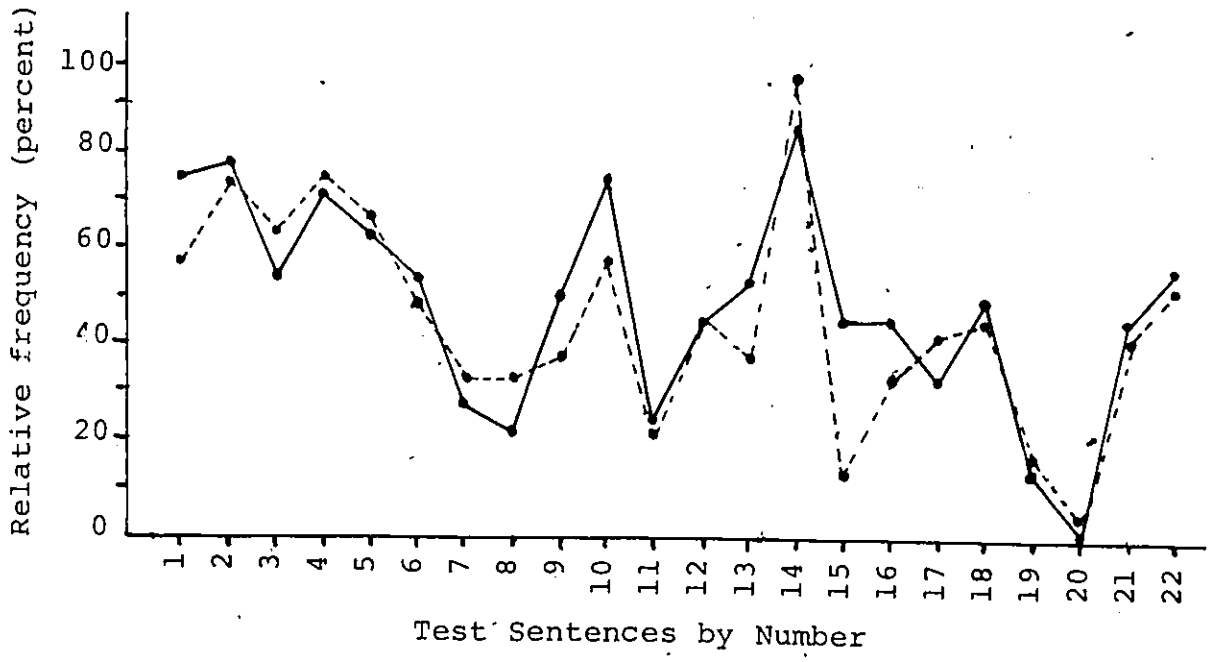
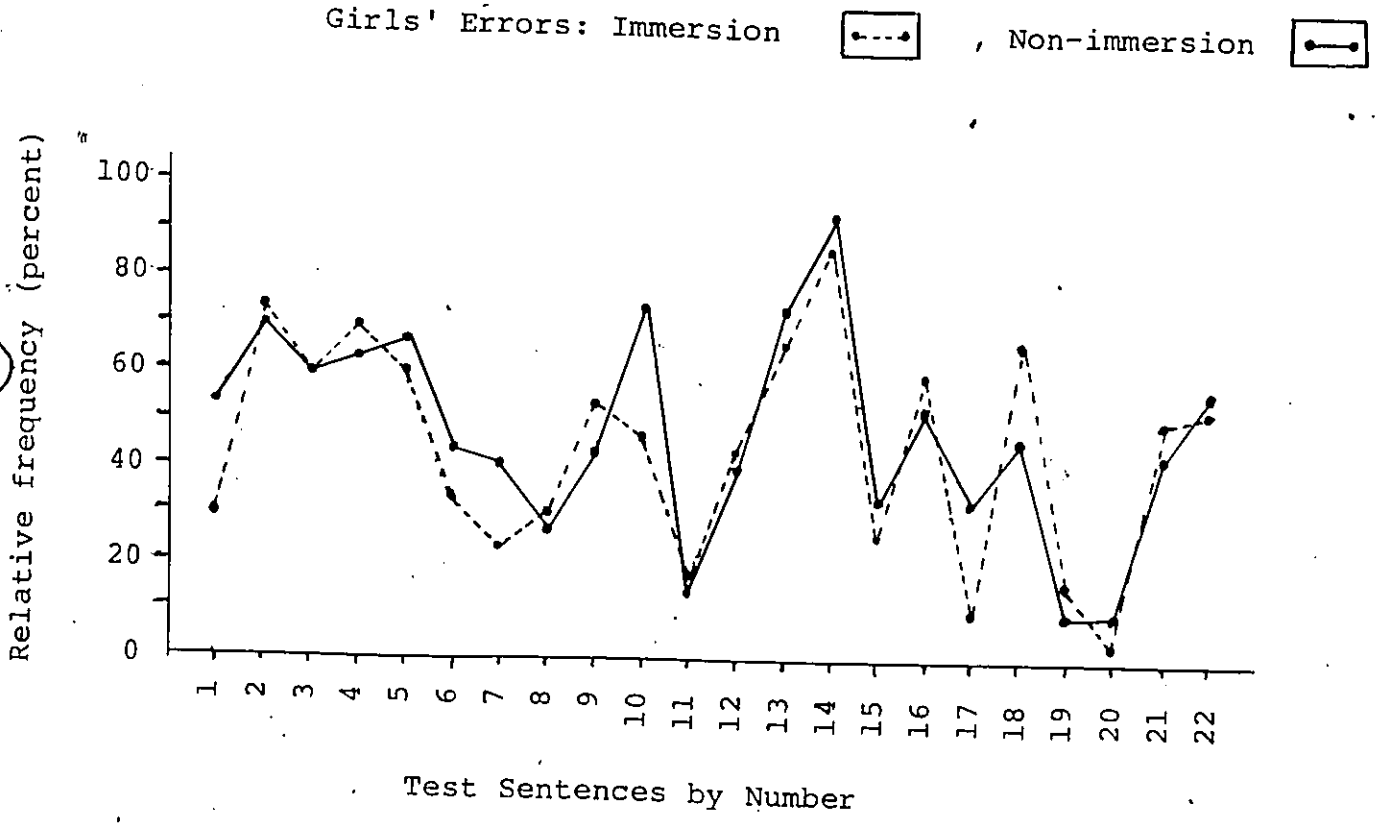


Figure 3.3



The Immersion and Non-immersion children were divided into three groups on the basis of verbal IQ -- low (≤ 99), average (100-115) and high (≥ 116). The paired T-Test was used to compare the Immersion and English total test score within each of these groups. It was considered possible that differences that were not apparent in the whole group might show up at the different IQ levels. (Verbal IQ was the measure.) A comparison of the mean number correct for boys and girls at each of the three IQ levels showed no significant differences. In other words the test means for children with low IQ's were the same in the English and Immersion groups, as was true of the average and high levels.

The boys and girls were also looked at separately. There

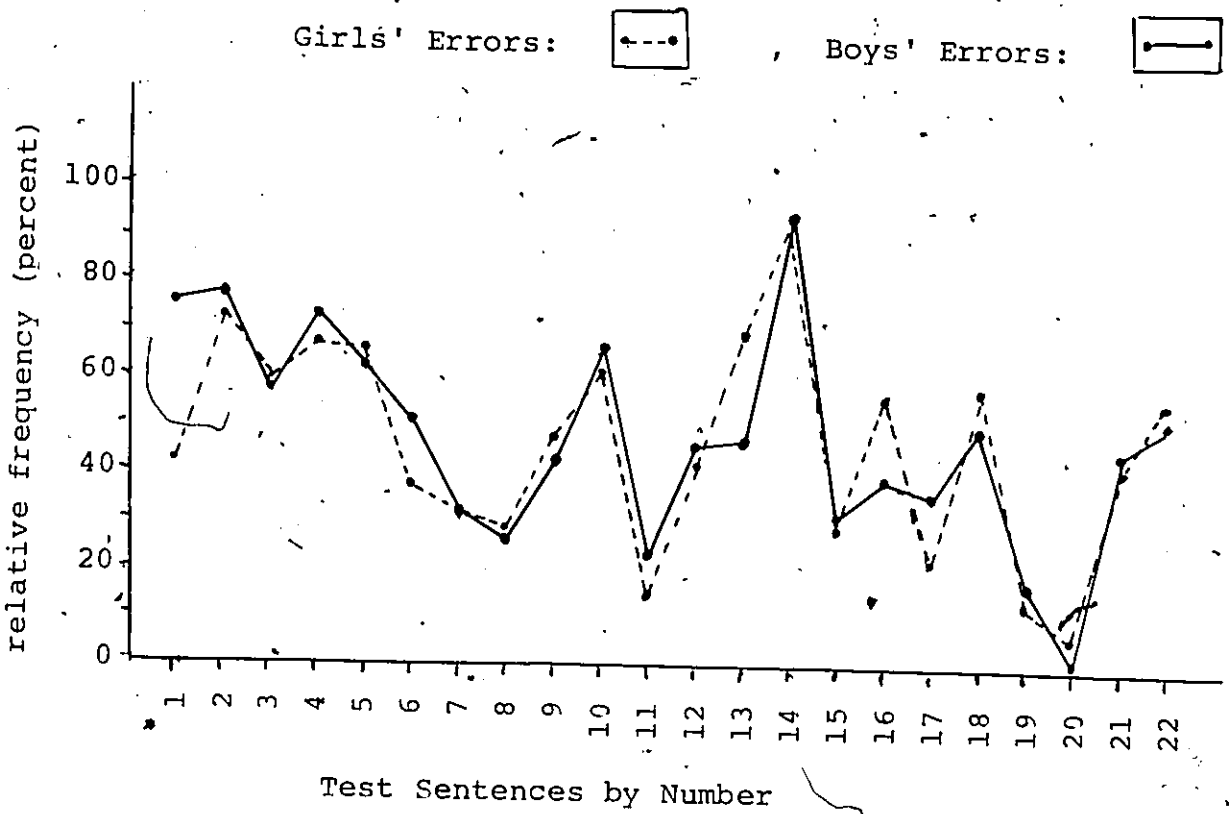
were no differences except in the case of the English and Immersion boys from the low IQ group. The mean for Immersion boys was 12.2 and for their English 'twins' was 9.8. This was significant ($p<.04$). It should be noted that there were only 5 boys from each group with verbal IQ's ≤ 99 and so there were only 4 d.f. for this test. There was also a significant difference in the non-verbal IQ's of these two groups. The Immersion boys with low verbal IQ's had a non-verbal average of 106 and their English counterparts had a mean of 96 ($p < .05$). As was mentioned there were no other differences among the boys and there were none at all between the girls at the three IQ levels.

Combining the girls from both groups and comparing their means to those of the boys they do significantly better on S1 whereas the boys are better on S13 ($p < .01$ and $p<.02$ respectively). Figure 3.4 shows how close the error rates of the two sexes are.[1] The mean for the total test differ by only .14 (girls in the lead) and this is not significant.

---

1 This supports the Tremaine results for no effect of sex.

Figure 3.4



Comparing English boys to English girls there are no significant differences in their performance on any of the sentences nor on the total test score. In the French group the test score difference is not significant but girls and boys differ on some specific items. Girls do significantly better on S1 and S17 and the boys have significantly higher means on S13 and S16.

3.2.2 T-Tests for Means on Verbal and Non-Verbal IQ's and Ages

Using the group as a whole and dividing it into groups according to whether a test item was answered correctly or incorrectly T-Tests for differences in means were run on each of the variables, verbal and non-verbal IQ and age. A one-

tailed test was used -- the hypothesis being that the means for Gr. 1 (correct response) would be higher than the means for Gr. 2. That is, higher IQ's and ages would go with right answers. This proved to be true in the case of some test sentences. Table 3.2 gives the results for the three variables which were significant at a level of at least .05 for the group as a whole. The relevant means are noted. The children who got S13 and S22 correct have higher verbal and non-verbal IQ's whose differences are significant. For S2 the verbal IQ's are higher but the non-verbal IQ's are lower. The group which got S4 and S6 right have higher non-verbal IQ's than those who were wrong. The mean age is higher for those who are right on S19 but significantly younger on S17. This is possibly explained by the fact that children with high IQ's were younger than the others.

TABLE 3.2

Significant differences in means of IQ's and age for sentence$_i$ within the entire group

Sentence	Verbal IQ		Non-verbal IQ		Age	
	$\bar{X}_1$	$\bar{X}_2$	$\bar{X}_1$	$\bar{X}_2$	$\bar{X}_1$	$\bar{X}_2$
S2	113	107	113	119		
S4			122	114		
S6			118	113		
S13	112	107	120	114		
S17					10.9	11.1
S19					11.0	10.9
S22	112	106	120	113		

The same tests were run within the Immersion and English subgroups. Those Non-immersion children who answered S17 and S15 correctly had significantly higher verbal IQ's than those who did not. Similarly in the case of S9 and non-verbal IQ's. There were only 6 children who got S19 wrong in the English group. Their age (10.6) was significantly lower than was that of the children who were right (11.0). Within the Immersion group there was a significant difference in the expected direction in both verbal and non-verbal IQ's on S6. Table 3.3 tabulates the above results.

TABLE 3.3

Significant differences in mean IQ's and ages for sentence$_i$ within the sub-groups

Group	Sentences	Verbal IQ		Non-verbal IQ		Age	
		X_1	X_2	X_1	X_2	X_1	X_2
English girls/boys	S9			120	111		
	S15	112	104				
	S17	111	103				
	S19					11.0	10.6
French girls/boys	S6	112	104	121	104		

Students were compared using the total test score as a classificatory criterion. Gr. 1 were those who had scores of 13 or more (n = 41) and Gr. 2 the remaining 67. Gr. 1 had significantly higher verbal and non-verbal IQ's but there was no age difference. (Tests were one-tailed.) (For details

on mean ages and IQ's at each test score level see App. C.)

	Gr. 1	Gr. 2	p
IQV	112	107	$<.02$
IQNV	123	113	$<.001$
Age	11.0	10.9	$>.05$

The four subgroups were also looked at separately and the means for Gr. 1 non-verbal IQ's were significantly higher in the Immersion girls and the Non-immersion boys groups. Non-immersion girls differed significantly on verbal IQ.

Means tests were run comparing (a) Non-immersion to Immersion girls, (b) Non-immersion to Immersion boys and (c) the whole Immersion group to all the Non-immersion students. The criterion used was greater than or equal to 14 as a test score versus less than 14. In other words in the case of (a) all girls with high test scores from the English group were compared to all girls with high scores from the Immersion group. And similarly all those with low scores were compared. The variables tested were age, verbal IQ and non-verbal IQ. There were no significant differences on any of the eighteen tests. That is, children who did well on the test from the Non-immersion group matched those from the Immersion group who did well, on these three variables. This was also true for the low scoring children.

Children were grouped into three IQ groups on the basis of verbal IQ -- low, average and high. Within each of these groups there were two classes, those who did well on the test (≥ 14) and those who did not (<14). T-Tests were run comparing the mean ages and non-verbal IQ's in each of these mutually exclusive

classes. The tests were not run separately for Immersion and Non-immersion groups because there were too few children who had scores of 14 or better in the low or average or high IQ groups (6, 8 and 11 respectively). There were differences in mean age in the low and average groups ($p<.05$). The mean age for the high scorers in the low IQ group was 11.3 whereas the age for the low scorers in the same IQ group was only 10.7. In the average IQ group the mean ages for high and low were 11.2 and 10.9 respectively. This group also had a significant result for non-verbal IQ -- 111 versus 100. In the high IQ group there were no significant differences.

It happens that the sample is lopsided when one looks at age and IQ. For some reason the older children tend to be in the lower IQ groups. This is true of boys and girls in both groups. Since the English children were matched on the basis of age, IQ, and sex with members of the Immersion group it is not surprising that they follow the Immersion children's pattern. However there seems to be no reason for this to have happened with a random sample of Immersion boys and girls. The mean ages for the different IQ levels are listed in Table 3.4.

TABLE 3.4

Mean ages for the different IQ levels

Group	High IQ	Average IQ	Low IQ
English girls	10.7	11.0	11.2
English boys	10.8	10.9	11.2
Immersion girls	10.8	11.0	11.1
Immersion boys	10.7	11.0	11.1
English girls/ boys	10.7	11.0	11.2
Immersion girls/ boys	10.8	11.0	11.1
All Groups	10.8	11.0	11.1

A two-tailed test comparing the ages in the high and low groups was significant ($p < .000$).

3.3 Chi Square tests for independence

Chi Square tests were run in the group as a whole and in the two subgroups Immersion and Non-immersion in order to determine whether there were any dependency relations between the total test scores and particular sentences on the one hand (23 variables) and the 4 variables age, sex, and verbal and non-verbal IQ on the other.

Testing within groups was to enable us to ascertain whether dependencies which existed in one group were, or were not, the same as those which existed in the other. This was in fact the case both when specific sentences were looked at and when test scores were considered. For instance, S6 and S17

were not independent of non-verbal IQ in the Immersion group whereas in the Non-immersion group this dependency relation existed for S4 but not for S6 or S17. Similarly for relationships with verbal IQ. This variable was significantly related to S2, S3 and S12 in the English group but only the S12 dependency existed in the Immersion group. There were no significant results for age in either group when looking at specific sentences. Each of the items was tested in the group as a whole and within the subgroups for any sex dependencies. The only significant results were for the group as a whole. On S1 girls were better ($p<.02$) and on S13 boys performed better ($p<.03$).

It might be noted again here that T-Tests comparing the means of boys and girls in the French group have significant results for S1, S17, S13, and S16 -- the girls having higher means on the first two and the boys on the latter. Since there was only one degree of freedom in these Chi Square tests there is more chance of a Type 11 error here than with the T-Tests with 52 d.f., that is the T-Test is possibly the better measure here.

Tests for dependencies between total test scores and age, sex, and IQ in the whole group were first done with 13 as the boundary between high and low scores on the test. The result was significant for non-verbal IQ ($p<.002$) but not for verbal IQ. A reclassification of scores using 14 as the pivotal point showed a dependency between verbad IQ and test score. However when looking at dependencies and controlling for other variables 14 as a cutoff left too many cells empty so 13 was again used. The relationships of verbal and non-verbal IQ controlling for

age were looked at as was verbal IQ and test score controlling for non-verbal IQ. There were no significant results except for non-verbal IQ within the high age group. As noted above, when non-verbal IQ was tested with test score without controlling for age the result was also significant.

Looking at the results in the subgroups separately, test score was related to non-verbal IQ in the Immersion group, but not to either verbal IQ or age. Controlling for age did not make any difference to the verbal IQ test score relationship -- they remained independent. The non-verbal IQ and test score dependency was true of the upper age group but, as in the case of the whole group, non-verbal IQ was independent of test score in the low age group. Using Fisher's Exact Test the probabilities were .049 and .057 respectively. It should be mentioned that this difference, albeit small, is not because there were fewer children in the low age group -- in fact there were more (30 versus 24).

In the English group, unlike the French group, non-verbal IQ was not related to test score but verbal IQ was (again unlike the Immersion group). Controlling for age did not effect the independence of non-verbal IQ. The verbal IQ dependency remained significant in the lower age group but not in the upper age group (27 in each group). When controlling for non-verbal IQ the verbal IQ relationship is significant in the high non-verbal group but not in the low non-verbal group. The results in this case are not close, but there are only two cases of children with high verbal IQ's who have low non-verbal IQ's.

3.4 Pearson's 'r'

As was mentioned in 3.2.2 the older children tend to have the lower IQ's in both the Immersion and Non-immersion groups. In testing for correlations between the variables test score, age, verbal IQ and non-verbal IQ the only significant negative relations were between age and both IQ's. 'r' was equal to -.47 for age and verbal IQ in the group as a whole and -.23 for age and non-verbal IQ. Looking at the groups individually the French boys have the highest negative 'r' for verbal IQ and age -- -.56 but each group has a significant negative value for this relationship. The only minor subgroup which has a significant result for the age non-verbal IQ correlation is the Immersion boys where 'r' = -.39.

The positive relationship between test scores and verbal IQ's is not particularly strong in either the French or English group. The coefficients are .23 and .35 respectively which means that the strengths of the relationships as measured by 'r^2' are only .05 and .12 respectively. The coefficients are almost equally weak for the test score, non-verbal relation. Tables 3.5 and 3.6 give all the significant 'r' values. Table 3.6 tabulates the results for the minor subgroups. It should be noted that the correlation between test score and non-verbal IQ in the English boys group is high -- 48% of the variation in test scores can be explained by non-verbal IQ in this subgroup (r = .69). As might be expected, the relationship between verbal and non-verbal IQ is positive and fairly high. The highest coefficient is .67 (English boys) and the lowest is

.58 (English girls).

TABLE 3.5

Correlations ('r') which were significant for the whole group and the two major subgroups ($\alpha = .05$)

	Group	Test Score	Age	Verbal IQ	Non-verbal IQ
Test Score	All			.28	.33
	English			.35	.34
	French			.23	.32
Age	All			-.47	-.23
	English			-.48	-.25
	French			-.45	-.23
Verbal IQ	All	.28	-.47		.61
	English	.35	-.48		.61
	French	.23	-.45		.61
Non-verbal IQ	All	.33	-.23	.61	
	English	.34	-.25	.61	
	French	.32	-.23	.61	

TABLE 3.6

Correlations which were significant within each of the minor subgroups (α = .05)

	Group	Test Score	Age	Verbal IQ	Non-verbal IQ
Test Score	Eng.girls			.32	
	Eng.boys			.43	.69
	Imm.girls			.38	..41
	Imm.boys				
Age	Eng.girls			-.43	
	Eng.boys			-.54	
	Imm.girls			-.34	
	Imm.boys			-.56	-.39
Verbal IQ	Eng.girls	.32	-.43		.58
	Eng.boys	.43	-.54		.67
	Imm.girls	.38	-.34		.60
	Imm.boys		-.56		.61
Non-verbal IQ	Eng.girls			.58	
	Eng.boys	.69		.67	
	Imm.girls	.41		.60	
	Imm.boys		-.39	.61	

It seemed possible that relationships might be different at different IQ levels. As indicated in Table 3.7 this turned out to be the case. The relationship between test score and non-verbal IQ for English boys of high IQ is .94 whereas there is no significant relationship for boys of average or low IQ on

these two variables. Similarly the verbal non-verbal IQ relation has a value of .89 for French girls of high IQ but there is no significant relationship between these variables for the French girls of low and average IQ. The French girls and English boys match here but the opposite is true of the English group of girls. For verbal and non-verbal IQ in the low IQ group there is a .77 coefficient while there is no correlation at the average or high levels. In fact there is an almost significant negative relationship at the high IQ level, $r = -.62$ and $p < .07$.

It is noteworthy that the negative age and verbal IQ correlation is already beginning to surface at the average IQ level. This might have implications for any IQ age interactive effects with regard to the test score. However, the relationship, although significant is weak with 'r' only -.24 for the group as a whole. By the high IQ level 'r' has increased to -.44 in the whole group, but a good part of this must be due to the English boys subgroup whose correlation coefficient is -.81.

TABLE 3.7

Significant correlations at the three IQ levels

Group	No.	Variables	'r'
LOW IQ. significant correlations (α = .05)			
All	25	Test score and age	.35
		Test Score and non-verbal IQ	.42
		Verbal and non-verbal IQ	.44
English girls/boys	13	Verbal and non-verbal IQ	.42
English girls	8	Age and non-verbal IQ	.72
		Verbal and non-verbal IQ	.77
AVERAGE IQ significant correlations (α = .05)			
All	54	Age and verbal IQ	-.24
		Verbal and non-verbal IQ	.32
French girls/boys	28	Test Score and non-verbal IQ	.33
		Age and verbal IQ	-.31
		Verbal and non-verbal IQ	.38
French girls	16	Test Score and verbal IQ	.45
		Test Score and non-verbal IQ	.54
French boys	12	Verbal and non-verbal IQ	.64
HIGH IQ significant correlations (α = .05)			
All	29	Age and verbal IQ	-.44
French girls/boys	14	Verbal and non-verbal IQ	.63
English boys	8	Age and verbal IQ	-.81
		Test score and non-verbal IQ	.94
French girls	7	Verbal and non-verbal IQ	.89[1]

1. For English girls 'r' = -.62 ($p<.07$)

3.5 Multiple Regression

Assuming a linear relationship existed between the total test score and the variables age, verbal and non-verbal IQ a regression analysis was done for the whole group, the major subgroups and the minor subgroups. As can be seen in Table 3.8 there is a linear regression for the group as a whole where multiple 'r' = .37 and $p < .01$. But looking at the subgroups there are significant results only in the case of the English boys and girls considered together and English boys alone. There are no significant indications of regression for the Immersion groups. Although r = .51 in the case of the French girls F is not significant.

TABLE 3.8

Results for the regression of test score on age, verbal and non-verbal IQ (See 2.5.5 for statistical model)

Group	Multiple 'r'	F
All boys/girls	.37	4.07**
English boys/girls	.46	3.15*
French boys/girls	.33	1.53 N.S.
English girls	.42	1.31 N.S.
French girls	.51	2.19 N.S.
English boys	.70	4.09*
French boys	.38	.81 N.S.

Note: N.S. not significant
* significant at .05 level
** significant at .01 level

Because of the significant correlation of non-verbal IQ and test

score when Pearson's 'r' was computed we attempted to fit a reduced model to the data:

$$Y_i = B_o + B_1X_{1i} + B_2X_{2i} + E_i$$

where i depends on the number in the group being considered

X_1 = non-verbal IQ

X_2 = (non-verbal IQ)(verbal IQ)(age)

Using this reduced model the results are similar to those for the full model which included two more independent variables, verbal IQ and age. Table 3.9 tabulates the results.

TABLE 3.9

Results for reduced model:
$(Y_i = B_o + B_1X_{1i} + B_2X_{2i} + E_i)$

Group	Multiple 'r'	F
All boys/girls	.36	7.87**
English boys/girls	.45	6.14**
French boys/girls	.32	2.96 N.S.
English girls	.40	2.51 N.S.
French girls	.48	3.93*
English boys	.69	8.81**
French boys	.25	.67 N.S.

Note: N.S. not significant
* significant at .05 level
** significant at .01 level

Using only these two variables the result for French girls is significant at the .05 level, and the English boys' result is here significant at the .01 level. There is still no reason

to assume regression in the case of English girls and French boys. English girls had a significant simple 'r' correlation with verbal IQ and so a second reduced model was fitted with X_1 = verbal IQ (instead of non-verbal IQ). The results remained non-significant.

CHAPTER IV

DISCUSSION

4.1 Weaknesses in the design

The main purpose of this study was to determine if French Immersion had the positive effects on English syntactic development claimed for it by Tremaine (1975). One obvious limitation of the study in this respect was that the 22 sentences represent only a subset of the gamut of English syntactic structures. Thus one's conclusions can only be tentative regarding syntactic development in general.

Using a larger sample (vis-à-vis Tremaine's) which was homogeneous with respect to number of years in the Immersion Programme, and controlling for age, sex and verbal IQ in the individual members of the group (that is, using 'matched pairs') the hypothesis was that there would be no confirmation of the Tremaine results. Strictly speaking, the nature of the design did not permit us to address the central question at all. Whatever the results of the test no causal relation between the Immersion Programme and syntactic development could be inferred. At most we could cite differences and show the existence of negative or positive correlations. And the problem with correlations is, of course, the old problem of the high correlation between drinking milk as an infant, and being a murderer as an adult.

This is the problem which faces all those researching the effects of Immersion. A causal relationship could only be

inferred if the French Immersion 'treatment' were the only difference between the Immersion and Non-immersion children. Among the countless and perhaps unspecifiable things that seem to be related to learning are, to cite two, mental ability and motivation. It is not clear that Immersion and English children in the Ottawa system are equal in these two areas (ignoring any other possible differences). With respect to mental ability children in the Public School System in Ottawa are tested in 4-year-old kindergarten and parents of children who do not do well on the test are strongly dissuaded from entering their children in the Immersion programme. This kind of screening goes on throughout the primary school years. Looking at motivation, Lambert (1972) refers to the highly motivated parents of the children in his study. From casual observation this seems equally true of parents of Immersion children in Ottawa, whatever the various reasons might be. How much of an effect does this have on the cognitive development of their children?

Mental ability and its interaction with age, and motivation differences might have been part of the reason for the Lambert and Tremaine results. This study tried to eliminate the former[1] as a possible source of variation by using 'matched pairs' rather than random samples from both programmes. Verbal IQ was used -- not non-verbal. Results of this study suggest that in the case of English boys the more appropriate measure would have been non-verbal IQ, however verbal seems the

---

1 As measured by IQ tests.

relevant measure for girls. (These judgements are based on the correlation coefficient for test score and IQ). Perhaps both should be used. How one eliminates factors like motivation and all other sources of experimental error remains an enigma. Unfortunately children are not albino white mice and it is rather difficult to control for individual differences.

The central question of this study then could not be answered in any definitive way. What of the peripheral questions? We wanted 1) to compare Immersion and Non-immersion performance on specific structure types as exemplified in each of the 22 sentences and 2) to determine whether relations between the dependent variable and the independent variables existed and if so, to examine their direction, strength and form. With regard to the former (1) any results could only point to areas of interest for further research because several examples of each type of structure would be needed for any conclusion to be reached. In the case of the latter (2) this researcher is at a loss as to how to interpret the results. This one might call a weakness in the researcher rather than in the design.

4.2 Inexplicable results

Let us first consider these 'uninterpretable' results. The dependencies and linear relations that were found to exist were a 'rag-bag'. What was true of one group was not true of another. And this pattern did not follow the Immersion Non-immersion dichotomy. As measured by the Chi Square statistic for instance, there were 4 sentences for which there were

significant dependencies for the English group, and this was true of 3 sentences for the French group. The dependency was the same only in the case of one sentence -- verbal IQ and S12 were not independent. 'The witch hoped to seem to be a nice person' (S12). This suggests that the French and English groups might be different. But then we look at Pearson's 'r' for test score and the different variables. Here we have significant correlations for non-verbal IQ in the French girls group ('r' = .41) and the English boys group ('r' = .69). (See Table 3.6.) This relationship goes as high as .94 for English boys of high verbal IQ but there is no significant non-verbal IQ test score relationship for French girls of high IQ (See Table 3.7). The form of the relationships were different too as indicated by the regression analysis. There was no multiple regression of the test score on the variables for the French boys group but the result was significant for English boys. 'Helter-skelter' might best describe the relationships. Do they suggest the possibility that the processing of syntactic data may indeed be quite separate from the processing of other data? Tremaine refers to a correlation with 'r' = .35 as a high correlation between her syntactic comprehension test and Piagetian 'operational' level, because it was significant at the .01 level. We had a similar result for the test and non-verbal IQ in the group as a whole -- 'r' = .33 and $p < .01$. But as mentioned above this correlation was only true when looking at the subgroups of French girls and English boys, not of French boys or English girls. Because the

English boys' result was so high -- .69 -- a good proportion of our .33 probably refers to them. What does Tremaine's 'r' = .35 mean? One interpretation based on r^2 is that 12% of the variation in the test score can be accounted for in terms of operational level (or in our case 11% in terms of non-verbal IQ).[1] What of the other 88%? Clearly there is a lot of something affecting comprehension which is not operational level. (Might this missing 'something' be Chomsky's separate language faculty?)

4.3 Ambiguous results

What did the results of this study indicate with regard to Immersion and native language syntactic development? They tended to confirm our hypothesis, which, as stated was the null hypothesis. However they were ambiguous.

First let us consider the results in terms of total test scores. Paired T-Tests indicated that Immersion girls did not differ from Non-immersion girls. Neither did the boys differ. Neither did the group as a whole differ using a two-tailed test with alpha = .05. In terms of this study's hypothesis a two-tailed test indicated the bias of the researcher so a one-tailed test was run. The .7 difference in means between the two groups became significant. Perhaps a brief look at

---

1 Not only were verbal and non-verbal IQ not sufficient conditions for doing well (≥ 14) on the test they were not even absolutely necessary conditions. Three examples:

	IQV	IQNV
English boy	98	102
French girl	89	95
French boy	89	95

significance is in order here. The larger your sample the more powerful your test -- a difference of .0001 could be significant at the .0001 level if your sample were large enough. Is the difference of .7 in fact interesting? A more extensive study would answer that. The possibility was considered that the Immersion Programme might have different effects depending on IQ level. For instance that Immersion children of high IQ might do better on the test than Non-immersion children of high IQ, whereas this might not be true of children of low IQ in both groups. The null hypothesis was supported in all cases. No difference in mean test scores at each of the three IQ levels when comparing the girls and boys together. (There was one exception when looking at boys and girls separately, see 3.2.1).

Similarly with respect to the children who did well on the test from both groups (≥ 14) and those who did badly (<14). No differences in age, verbal IQ or non-verbal IQ in the two groups. In other words almost all tests comparing the performance of the two groups on the test as a whole supported the hypothesis -- no difference.

This similarity between the group performances is evident too when one compares the error rates on the separate sentences. (See Figures 3.1-3.3) But why then were the results referred to as ambiguous? Because there are significant differences on 5 individual items, and, with the exception of one of them where English girls are better (S18), the differences are all in favour of the Immersion group. One would

need to test several examples of the 5 structures concerned to be confident of these results, however they are suggestive.

4.4 Areas for further research

(a) Are the above mentioned differences real? If these differences with regard to specific items are real are they language independent or are they related to English and French? That is, does second language learning affect the processing of certain kinds of syntactic structures?

(b) What was surprising about the results of this study was the fact that almost all the children commented on how easy the test was and yet they did rather badly. That is, they thought they understood when in fact they did not. Would this be true of older children? of adults? How would adults compare to the ten year-olds of this study?

(c) This study focused on comparing two groups and completely ignored one, clearly crucial,area. What is at the root of the complexity of these sentences? Why is 'Bobby showed the sheep the lion' so much more difficult than 'The hammer wasn't in the tool-box, having been borrowed'? And why is this latter so much easier than its cognate 'The Smiths didn't have any money in the bank, having been robbed'?

CHAPTER V

SUMMARY AND CONCLUSIONS

Various studies have concluded that in 'additive' situations second language learning has positive effects on native language development. We considered that these positive results might have been due to inadequate controls of independent variables known to be related to language development -- namely, sex, age, verbal and non-verbal IQ. In order to remedy these possible weaknesses, to eliminate these possible sources of the positive results, this study used a 'matched pairs' design.

'Linguistic' and 'communicative' competence were differentiated. This distinction was drawn as an apologia for using a highly unnatural setting -- an 'out-of-context' comprehension test -- as a basis for inferences regarding syntactic development.

A comprehension, rather than a production task was used because of the lack of well defined criteria for interpreting the results of production tasks. Each of the sentences used to test comprehension was one which had been found, in the literature, to be 'complex'.

There were 108 girls and boys from the Immersion and English programmes who were tested together in small groups on the 22 test items.

The hypothesis of this study, that there would be no difference in the performance of the Immersion and Non-immersion

groups was

1) supported with regard to their performance on the test as a whole,

but 2) not clearly supported when their performance on particular structures was considered. There was some indication that the groups differed here, but a more extensive test would be needed to decide on this.

The effect of sex was also only visible in a few particular items, not on overall performance. There were positive correlations of verbal and non-verbal IQ both with specific test sentences and on the total test score. The effect of age was not apparent in the group as a whole, but this was possibly because those with high IQ's were young. The correlations differed both within and across the Immersion and Non-immersion groups. These differences could not be explained.

In general then, although the study could only be suggestive as to the peripheral areas it explored the design seemed appropriate for the specific task it was set up for; to test the null hypothesis that the experimental and control groups did not differ in English syntactic development. The results of the various tests gave us no reason to reject the null hypothesis. There is no reason to assume that second language learning improves native language syntactic development.

APPENDIX A

Table 1

Descriptive statistics for each subgroup for age and IQ

Group	Statistic	Age	Verbal IQ	Non-verbal IQ
English boys	mean	10.9	112	116
	stand. dev.	.4	15	17
	min.	10.0	90	86
	max.	11.6	150	145
French boys	mean	10.9	112	118
	stan. dev.	.4	13	15
	min.	10.0	89	91
	max.	11.5	144	141
English girls	mean	11.	106	115
	stan. dev.	.3	12	15
	min.	10.2	86	79
	max.	11.5	139	142
French girls	mean	11.	106	117
	stan. dev.	.3	12	15
	min.	10.4	79	75
	max.	11.6	135	142

APPENDIX B

Table 1(a)

Hypotheses: H_o: $U_1 = U_2$ H_1: $U_1 \neq U_2$

Statistical Test: Two-tailed paired T-Test
alpha = .05

Variables: $Sentence_i$ (i = 1,...,22)

Total test score

Groups	Degrees of freedom
Group 1 English girls/boys Group 2 French girls/boys	54-1
Group 1 English girls Group 2 French girls	30-1
Group 1 English boys Group 2 French boys	24-1

APPENDIX B

Table 1(b)

Hypotheses: H_o: $U_1 = U_2$ H_1: $U_1 \neq U_2$

Statistical Test: Two-tailed unpaired T-Test, alpha = .05

Variables: $Sentence_i$ (i = 1,...,22)
Total Test Score

Groups	Degrees of freedom
Group 1 English and French girls Group 2 English and French boys	108-2
Group 1 English girls Group 2 English boys	54-2
Group 1 French girls Group 2 French boys	54-2

APPENDIX B

Table 2(a)

Hypotheses: H_o: $U_1 = U_2$ H_1: $U_1 > U_2$

Statistical Test: One-sided unpaired T-Test
alpha = .05

Variables: Age
Verbal IQ
Non-verbal IQ

Groups		Degrees of Freedom
Group 1	Correct on sentence$_i$ ($i = 1,\ldots,22$)	108-2
Group 2	Incorrect on sentence$_i$	
Group 1	Correct on sentence$_i$ (English)	54-2
Group 2	Incorrect on sentence$_i$ (English)	
Group 1	Correct on sentence$_i$ (French)	54-2
Group 2	Incorrect on sentence$_i$ (French)	

APPENDIX B

Table 2(b)

Hypotheses: H_o: $U_1 = U_2$, H_1: $U_1 \neq U_2$

Statistical Tests: Two-sided unpaired T-Test
alpha = .05

Variables: Age
Verbal IQ
Non-Verbal IQ

Groups	Degrees of freedom
Group 1 Test Score ≥ 14 (English) Group 2 Test Score ≥ 14 (French)	(10-1) (15-1)
Group 1 Test Score ≥ 14 (English girls) Group 2 Test Score ≥ 14 (French girls)	(6-1) (10-1)
Group 1 Test Score ≥ 14 (English boys) Group 2 Test Score ≥ 14 (French boys)	(4-1) (5-1)
Group 1 Test Score < 14 (English girls/boys) Group 2 Test Score < 14 (French girls/boys)	(44-1) (39-1)
Group 1 Test Score < 14 (English girls) Group 2 Test Score < 14 (French girls)	(24-1) (20-1)
Group 1 Test Score < 14 (English boys) Group 2 Test Score < 14 (French boys)	(20-1) (19-1)

Table 3

Group Sizes for T Tests (see Table 2.4) based on number of correct and incorrect responses on each test sentence

Groups	Sentences (1-22)																						
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	Total
Group 1 No. of correct responses	51	28	45	32	37	60	74	78	58	40	88	61	43	10	76	55	78	51	92	103	59	49	1268
Group 2 No. of incorrect responses	57	80	63	76	71	48	34	30	50	68	20	47	65	98	32	53	30	57	16	5	49	59	1108
Total	108	108	108	108	108	108	108	108	108	108	108	108	108	108	108	108	108	108	108	108	108	108	2376

APPENDIX B

Table 4

Hypotheses: H_o: $U_1 = U_2$ H_1: $U_1 \neq U_2$

Statistical Test: Two-tailed Paired T-Test
alpha = .05

Variable: Total Test Score

Groups	Degrees of Freedom
Group 1 English girls/boys low IQ Group 2 French girls/boys low IQ	11
Group 1 English girls/boys average IQ Group 2 French girls/boys average IQ	27
Group 1 English girls/boys high IQ Group 2 French girls/boys high IQ	13
Group 1 English girls low IQ Group 2 French girls low IQ	6
Group 1 English girls average IQ Group 2 French girls average IQ	15
Group 1 English girls high IQ Group 2 French girls high IQ	7
Group 1 English boys low IQ Group 2 French boys low IQ	5
Group 1 English boys average IQ Group 2 French boys average IQ	11
Group 1 English boys high IQ Group 2 French boys high IQ	6

APPENDIX B

Table 5

Hypotheses: H_o: X_i and Y_i are independent
H_1: X_i and Y_i are not independent

Statistical Test: Chi Square Test for independence of criteria

Criterion Variables	Groups	Degrees of Freedom
X_i (i 1,...,4) 1. Age 2. Sex 3. Verbal IQ 4. Non-verbal IQ Y_i (i=1,...,22) 1-22. Sentences 1-22	Immersion girls/boys Y_{1-22} by X_{1-4}	1
	English girls/boys Y_{1-22} by X_{1-4}	1
	All boys/girls Y_{1-22} by X_2	1
X_i (i 1,...,4) 1. Age 2. Sex 3. Verbal IQ 4. Non-verbal IQ Y_i (i 1,...3) 1. Test score ≥ 13 2. Test score ≥ 14 3. Test score ≤ 10	Immersion girls/boys Y_2 by X_{1-4} Y_1 by X_3 by X_4 Y_1 by X_3 by X_1	1 1 1
	English girls/boys Y_1 by X_{1-4} Y_1 by X_3 by X_4 Y_1 by X_3 by X_1	1 1 1
	All girls/boys Y_{1-2} by X_{1-4} Y_3 by $X_{1, 3, 4}$ Y_1 by X_3 by X_1 Y_1 by X_4 by X_1 Y_1 by X_3 by X_4	1 1 1 1 1

APPENDIX C

Table 1

*Mean Ages and IQ's for children at each i^{th} score level

Score (no. correct out of 22)	No. of cases	Verbal IQ	Non-Verbal IQ	Age
7	5	105	115	11.3
8	8	103	117	10.8
9	10	101	102	11.0
10	9	107	111	11.0
11	13	108	110	10.9
12	22	110	118	10.9
13	16	111	122	10.9
14	13	109	121	11.1
15	6	116	123	10.9
16	2	114	122	11.1
17	2	110	132	11.2
18	2	126	131	11.2

*See Table 3.1 for descriptive statistics of the test scores of each group

APPENDIX D (1)

(1) Practice sentences

1. Susie really likes her baby brother.

1.(a) Susie likes her baby brother a little bit.

(b) Susie likes her baby brother a lot.

2. John loves swimming in the ocean.

2.(a) John loves to swim in the ocean.

(b) John loves the ocean and swimming.

APPENDIX D(2)

(2) Test Sentences read by Examiner

(b) 1. Judy played with her cat for a few minutes when she went to bed.

(a) 2. Bobby showed the sheep the lion.

(a) 3. Ginger sent the food to Gilligan for Mrs. Howell.

(b) 4. Susan doesn't not like cake.

(b) 5. Starbuck believed Apollo not to have been examined by the doctor.

(b) 6. The ghost the witch frightened ran away.

(a) 7. Susan fed her fish, having fed her canary.

(b) 8. The kids were coaxed by Jack to try to ride the horse.

(b) 9. That Lou Grant was insulted by Mary Tyler Moore shocked everyone.

(a) 10. Captain Apollo brought the money for Starbuck to Boomer.

(a) 11. The hammer wasn't in the tool box, having been borrowed.

(b) 12. The witch hoped to seem to be a nice person.

(a) 13. Joe Hardy believed Nancy Drew not to have been seen by the bad guys.

(b) 14. Steve Shutt, paid $500 for his car, wasn't happy.

(a) 15. Yesterday Mr. Zadow told the kids to be quiet during the concert.

(b) 16. The Canadians couldn't have been beaten by the Maple Leafs last night because there wasn't any game.

(a) 17. As clever as Steven was he couldn't do the last math problem.

(a) 18. Andy had watched T.V. for a couple of hours when he went to bed.

(a) 19. Billy practised very hard and yet he didn't get into Hockey School.

(b) 20. Having taken off his hat Mr. James went into the house.

(b) 21. The Smiths didn't have any money in the bank, having been robbed.

(b) 22. Christopher had studied for three hours when he went to bed.

Paraphrase choices

1. (a) Judy played with her cat for a few minutes and then she went to bed.

   (b) Judy played with her cat in bed for a few minutes.

2. (a) Bobby showed the lion to the sheep.

   (b) Bobby showed the sheep to the lion.

3. (a) Ginger helped Mrs. Howell by sending the food to Gilligan.

   (b) Ginger sent the food for Mrs. Howell to Gilligan.

4. (a) Susan does not like cake.

   (b) Susan likes cake a little bit.

5. (a) Starbuck believed Apollo when he said he had not been examined by the doctor.

   (b) Starbuck thought the doctor had not examined Apollo.

6. (a) The ghost frightened the witch and ran away.

   (b) The witch frightened the ghost and the ghost ran away.

7. (a) Susan fed her canary and then she fed her fish.

   (b) Susan fed her fish and then her canary.

8. (a) Jack tried to coax the kids to ride the horse.

   (b) Jack coaxed the kids to try to ride the horse.

9. (a) That Lou Grant insulted Mary Tyler Moore shocked everyone.

   (b) It shocked everyone that Mary Tyler Moore insulted Lou Grant.

10. (a) The money that Captain Apollo brought to Boomer was for Starbuck.

    (b) Captain Apollo brought the money to Boomer for Starbuck.

11. (a) Someone had borrowed the hammer and so it wasn't in the tool box.

    (b) Because the tool box had been borrowed the hammer wasn't there.

12. (a) It seemed that the witch hoped to be a nice person.

    (b) The witch hoped that it seemed as if she was a nice person.

13. (a) Joe Hardy thought that the bad guys hadn't seen Nancy Drew.

    (b) Joe Hardy believed Nancy Drew not to have seen the bad guys.

14. (a) Steve Shutt paid $500 for his car and wasn't happy.

    (b) Somebody paid Steve Shutt $500 for his car and Steve wasn't happy.

15. (a) Mr. Zadow, yesterday, told the kids to be quiet during the concert.

    (b) The kids were told by Mr. Zadow to be quiet during yesterday's concert.

16. (a) The Maple Leafs weren't able to beat the Canadians last night because there wasn't any game.

    (b) It isn't possible that the Maple Leafs beat the Canadians last night because there wasn't any game.

17. (a) Steven was clever but not clever enough to do the last math problem.

(b) He was as clever as Steven but he couldn't do the last math problem.

18. (a) Andy watched T.V. for a couple of hours and then he went to bed.

(b) Andy watched T.V. in bed for a couple of hours.

19. (a) Although Bill practised very hard, he didn't get into Hockey School.

(b) Although Bill didn't get into Hockey School he practised very hard.

20. (a) Mr. James went into the house and then took off his hat.

(b) Mr. James took off his hat and then he went into the house.

21. (a) Someone robbed the bank and so the Smiths didn't have any money in it.

(b) Because someone robbed the Smiths they didn't have any money in the bank.

22. (a) Christopher studied in bed for three hours.

(b) Christopher studied for three hours and then he went to bed.

References

Barik, H. and Swain, M. "Three-Year evaluation of a large scale early grade French Immersion programme: the Ottawa Study," Language Learning, 1975, 25, 1-30.

Bernstein, B. "Linguistic Codes, Hesitation Phenomena and Intelligence," Language and Speech, 5, 1962.

Bever, T. G. "The Cognitive basis for Linguistic Structures" in J. R. Hayes (ed.), Cognition and the Development of Language, (New York: John Wiley and Sons, 1970), pp. 279-352

Bever, T. G., Garrett, M. F. and Hurtig, R., "The Interaction of Perceptual Processes and Ambiguous Sentences," Memory and Cognition, 1973, I, 277-286.

Bever, T. G., Katz, J. J. and Langendoen, D. T., An Integrated Theory of Linguistic Ability. (New York: Thomas Y. Crowell Company Inc., 1976).

Bowerman, M.,"Structural relationships in Children's Utterances: Syntactic or Semantic?" in T. E. Moore (ed.), Cognitive Development and the Acquisition of Language (New York: Academic Press, 1973), pp. 197-213

Braine, M. D. S., "The Acquisition of language in infant and child" in C. Reed (ed.), The learning of language (New York: Appleton-Century-Crofts, 1971), pp. 7-95.

Brown, H. D. "Children's Comprehension of relativized English sentences", Child Development, 1971, 42, 1923-1936.

Brown, R. W. Psycholinguistics (New York: The Free Press, 1970).

Brown, R. W. A first language: The early stages (Cambridge, Mass.: Harvard University Press, 1973).

Brown, R. W. and Fraser, C."The Acquisition of Syntax", in U. Bellugi and R. Brown (eds.), The acquisition of language. Monographs of the Society for Research in Child Development, 1964, 29 (Serial No. 92), 43-79.

Brown, R. W. and Hanlon, C. "Derivational complexity and order of acquisition in child speech" in J. R. Hayes (ed.), Cognition and the development of language (New York: John Wiley and Sons, 1970), pp. 11-53.

Chomsky, C. The acquisition of syntax in children from 5 to 10 (Cambridge, Mass.: M.I.T. Press, 1969).

Clark, H. and Clark, E., Psychology and language: An introduction to psycholinguistics (New York: Harcourt Brace Johanovich, Inc., 1977).

Cummins, J., The influence of bilingualism on Cognitive growth (Dublin: Educational Research Centre, St. Patrick's College, 1976).

Cziko, G. A., Lambert, W. E., Sidoti, N. and Tucker, R. Graduates of Early Immersion: Retrospective views of Grade 11 students and their parents. (Montreal: McGill, 1978).

Dyer, J. R. Understanding and evaluating educational research. (Reading, Mass.: Addison-Wesley Publishing Co., 1979)

Edwards, H. P. and Casserly, M. C. Evaluation of second language programmes in the English Schools. (Ottawa: Annual Report, 1972-1973, Roman Catholic School Board, 1973).

Ferguson, C. A., and Slobin, E. I. (eds.) Studies of Child Language development. (New York: Holt, Rinehart and Winston, 1973).

Flores d'Arcais, G. B. and Jevelt, W. J. M. (eds.), Advances in psycholinguistics (Amstercam: North-Holland Publishing, 1970).

Fodor, J. A.,Bever, T. G. and Garrett, M. F., The psychology of language: An introduction to psycholinguistics and generative grammar. (New York: McGraw-Hill, 1974).

Fodor, J. A. and Garrett, M. F. "Some syntactic determinants of sentential complexity," Perception and Psychophysics, 1967, 2, 289-296.

Fodor, J. A. Garrett, M. F., and Bever, G. "Some syntactic determinants of sentential complexity, II: Verb structure," Perception and Psychophysics, 1968, 3, 453-461.

Foss, D. J., and Hakes, D. T. Psycholinguistics: An introduction to the psychology of language. (Englewood Cliffs, N.J.: Prentice-Hall Inc., 1978).

Foss, D. J., and Lynch, R. H. Jr., "Decision processes during sentence comprehension: Effects of surface structure on decision times," Perception and Psychophysics, 1969, 5, 145-148.

Green, E. "On Children's comprehension of synonymous sentences," TESOL Quarterly, 3, 1969, 189-195.

Hakes, D. T. "Does Verb Structure affect Sentence Comprehension?" Perception and Psychophysics, 1971, 10, 229-232.

Hakes, D. T. "Effects of reducing complement constructions on sentence comprhension," Journal of Verbal Learning and Behaviour, 1972, II, 278-286.

Kessler, C., The acquisition of syntax in bilingual children. (Washington, D.C.: Georgetown University Press, 1971).

Lambert, W., Tucker, G. and d'Anglejau A.,"Cognitive and Attitudinal consequences of bilingual schooling: The St. Lambert project through grade Five," Journal of Educational Psychology, 1973, Vol. 65, No.2, 141-159.

Loban, W., Language development: Kindergarten through Grade Twelve. (Urbana, Ill.: NCTE, 1976).

Mackay, D. G. "On the retrieval and lexical struxture of verbs," Journal of Verbal Learning and Verbal Behaviour, 1976, 15, 169-182.

McCarthy, D. "Language development in children" in L. Carmichael (ed.), Manual of Child Psychology (New York: Wiley, 1954).

McNeill, D. The acquisition of language: The study of developmental psycholinguistics. (New York: Harper & Row, 1970).

Mehler, J. "Some effects of grammatical transformations on the recall of English sentences," Journal of Verbal Learning and Verbal Behaviour, 1963, 2, 250-262.

Menyuk, P., The acquisition and development of language (Englewood Cliffs, N.J., Prentice-Hall, 1971).

Menyuk, P. Language and Maturation (Cambridge, Mass.: M.I.T. Press, 1977).

Moore, E. E. (ed.), Cognitive development and the acquisition of language. (New York: Academic Press, 1973).

Nelson, K. Children's Language, (New York: Gardner Press, Inc., 1978).

Peal, E., and Lambert, W. "The relation of bilingualism to intelligence," Psychological Monographs, 1962, 76, no. 27, whole no. 546.

Tremaine, R. V. Syntax and Piagetian Operational Thought (Washington, D.C.: Georgetown University Press, 1975).