INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

Bell & Howell Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

UMI®
THE EFFECT OF FORM-FOCUSED INSTRUCTION ON
CONTROL OVER GRAMMATICAL GENDER BY FRENCH
IMMERSION STUDENTS IN GRADE 11

by

MICHAEL WARDEN

A thesis submitted in conformity with the requirements
for the degree of Doctor of Education,
Department of Curriculum, Teaching and Learning,
Ontario Institute for Studies in Education
of the University of Toronto

© Copyright by Michael Warden 1997
The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

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

L’auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

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

0-612-46512-8
The Effect of Form-Focused Instruction on Control over Grammatical Gender by French Immersion Students in Grade 11.

Doctor of Education (1997)
Michael Warden.
Department of Curriculum, Teaching and Learning, Ontario Institute for Studies in Education, University of Toronto

ABSTRACT

This study investigates the effect of form-focused instruction on control over grammatical gender by French immersion students in Grade 11. Previous research has shown that while native speakers of French acquire control over grammatical gender at a very early age, and seemingly with little or no instruction, anglophone learners of French have a great deal of difficulty with this concept, since it is not a feature of their first language. There is a growing body of research providing evidence that form-focused instructional strategies, when used in conjunction with a primarily experiential approach, can enhance the accuracy of immersion students' second language output. To date, most of this research has been carried out at the elementary level. The present study arose partly out of the need to address the issue of form-focused instruction at the secondary level.

Sixty-two students took part. There was an experimental class and two comparison classes. The study consisted of pretests, an eight-week treatment period in the experimental class, immediate posttests and delayed posttests. The five tests used were: listening test, written endings test, agreement test,
writing assignment, and oral production task. During the treatment period, the experimental group was exposed to form-focused tasks designed to make them aware of word-ending regularities as clues to the gender of French nouns, and to provide practice in making gender agreements. Other aspects of the study included teacher observations of the experimental class, and questionnaires designed to elicit the students' reactions to the treatment materials.

Statistical analyses showed that the experimental group made significantly greater progress over time than the comparison classes on three measures (listening test, written endings test, agreement test). On the other two measures (writing assignment and oral production task) the experimental class outperformed the comparison classes over time; however, the differences were not found to be statistically significant.

While further research is needed to follow up on these results, there is at least tentative support for the use of form-focused instruction as a means of improving French immersion students' control over grammatical gender.
ACKNOWLEDGEMENTS

There are many individuals who have assisted in bringing this thesis to completion. First of all, I would like to thank my supervisor, Dr. Birgit Harley, who provided constant inspiration and encouragement from beginning to end. Her comments on the initial and subsequent drafts of each chapter were consistently detailed and insightful, and she was always readily accessible and most generous with her time. The other members of my thesis committee, Dr. Sharon Lapkin and Dr. Merrill Swain, contributed many valuable suggestions which were incorporated into the final version. I am also grateful to them for their encouragement throughout my years at the Modern Language Centre at OISE. Dr. Roy Lyster, whose own exemplary thesis inspired me constantly, and who provided valuable input in his external appraisal, also deserves special thanks.

Many teachers, students and friends at the MLC provided moral support and practical suggestions, and were willing to share their own experiences of conducting research with me. They are too numerous to mention individually, and there would be the risk of omitting someone if names were used. I am most grateful to all of them.

The students in both the experimental and comparison classes were unfailingly cheerful and co-operative despite being subjected to the same series of tests on three occasions. The experimental students must also be commended for their enthusiastic participation in the treatment activities. I am also very grateful to the teacher of the second comparison class, who agreed without hesitation to participate in this project, and was so welcoming and supportive every time I visited her school.
Finally, I wish to express my gratitude to my wife, Carolina, and to my daughters Sara, Julia and Eva, who have lived with this thesis for several years, and who have provided both moral and practical support in countless ways.
# CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>ii</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>iv</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>x</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>xii</td>
</tr>
<tr>
<td><strong>CHAPTER 1: INTRODUCTION</strong></td>
<td></td>
</tr>
<tr>
<td>1.1. Purpose</td>
<td>1</td>
</tr>
<tr>
<td>1.2. The Nature of the Study</td>
<td>4</td>
</tr>
<tr>
<td>1.3. Plan of the Thesis</td>
<td>5</td>
</tr>
<tr>
<td><strong>CHAPTER 2: BACKGROUND</strong></td>
<td></td>
</tr>
<tr>
<td>2.1. Definition of Grammatical Gender</td>
<td>7</td>
</tr>
<tr>
<td>2.2. Native Speakers' Acquisition of Grammatical Gender</td>
<td>11</td>
</tr>
<tr>
<td>2.3. L2 Learners' Acquisition of Grammatical Gender</td>
<td>16</td>
</tr>
<tr>
<td>2.4. Reactions of Francophones to L2 Learners' Gender Errors</td>
<td>24</td>
</tr>
<tr>
<td>2.5. Implications for L2 Instruction</td>
<td>25</td>
</tr>
<tr>
<td>2.6. Analytic and Experiential Approaches to Second Language Teaching</td>
<td>30</td>
</tr>
<tr>
<td>2.7. The Role of Form-Focused Instruction in French Immersion</td>
<td>34</td>
</tr>
<tr>
<td>2.8. Form-Focused Experimental Studies</td>
<td>37</td>
</tr>
<tr>
<td><strong>CHAPTER 3: EXPLORATORY RESEARCH</strong></td>
<td></td>
</tr>
<tr>
<td>3.1. Context and Participants</td>
<td>47</td>
</tr>
<tr>
<td>3.2. Research Questions</td>
<td>48</td>
</tr>
<tr>
<td>3.3. Frequency of Errors</td>
<td>49</td>
</tr>
<tr>
<td>3.4. Prepared Writing vs. Writing Under Pressure</td>
<td>51</td>
</tr>
<tr>
<td>3.5. Types of Gender Errors</td>
<td>52</td>
</tr>
<tr>
<td>3.6. Patterns of Gender Errors</td>
<td>52</td>
</tr>
<tr>
<td>3.7. The Role of Natural Gender</td>
<td>53</td>
</tr>
<tr>
<td>3.8. Consistency of Errors</td>
<td>54</td>
</tr>
</tbody>
</table>
5.2. Qualitative Analysis  
5.3. Results of Questionnaire Data

CHAPTER 6: CLASSROOM OBSERVATIONS

6.1. Activity #1. Dictionary Activity Focusing on Written Word Endings

6.2. Activity #2. Word-Ending Clues to Noun Gender: Rules and Exceptions

6.3. Activity #3. Worksheets on Adjective Endings and Determiner-Adjective-Noun Agreements

6.4. Activity #4. Student-Invented Games Focusing on Gender

CHAPTER 7: DISCUSSION AND CONCLUSIONS

7.1. Summary of the Study

7.2. Summary of the Findings

7.3. Significance of the Findings

7.4. Limitations of the Study

7.5. Implications for Pedagogy

7.6. Implications for Future Research

BIBLIOGRAPHY

APPENDICES

Appendix A: Letter Requesting Parents' or Guardians' Consent

Appendix B: Student Answer Sheet for the Listening Test

Appendix C: Student Answer Sheet for the Written Endings Test

Appendix D: Student Answer Sheet for the Agreement Test

Appendix E: Oral Production Task, Picture Identification (Sample)

Appendix F: Oral Production Task, Picture Description
Appendix G: Questionnaires Completed by Students in the Experimental Class 176

Appendix H: Student Worksheet Used in Treatment Activity #2 181

Appendix I: Student Worksheet Used in Treatment Activity #3 182
LIST OF TABLES

4.1. Endings and nouns used in the Listening Test 68
4.2. Endings and pseudonouns used in the Written Endings Test 71
5.1. Class means on Listening Test 93
5.2. MANOVA of Listening Test administered to three classes on three occasions 95
5.3. Mean scores on Listening Test by ending (experimental class) 97
5.4. Class means on Written Endings Test 99
5.5. MANOVA of Written Endings Test administered to three classes on three occasions 101
5.6. Mean scores on Written Endings Test by ending (experimental class) 103
5.7. Class means on Agreement Test 105
5.8. MANOVA of Agreement Test administered to three classes on three occasions 106
5.9. Class means for articles on the Agreement Test 108
5.10. Class means for adjectives on the Agreement Test 108
5.11. MANOVA of Agreement Test administered to three classes on three occasions (articles only) 110
5.12. MANOVA of Agreement Test administered to three classes on three occasions (adjectives only) 110
5.13. Class means on Writing Assignment 112
5.14. MANOVA of Writing Assignment administered to three classes on two occasions 114
5.15. Class means on Oral Production Task 116
5.16. MANOVA of Oral Production Task administered to three classes on two occasions 117

5.17. Students' ratings of the treatment activities 132
LIST OF FIGURES

5.1. Line graph of Listening Test results 94
5.2. Bar graph of Listening Test results 94
5.3. Line graph of Written Endings Test 100
5.4. Bar graph of Written Endings Test 100
5.5. Line graph of Agreement Test 105
5.6. Bar graph of Agreement Test 106
5.7. Bar graph of Agreement Test (articles only) 109
5.8. Bar graph of Agreement Test (adjectives only) 109
5.9. Line graph of Writing Assignment 112
5.10. Bar graph of Writing Assignment 113
5.11. Line graph of Oral Production Task 116
5.12. Bar graph of Oral Production Task 117
5.13. Class E students' ratings of the four treatment activities, grouped according to the three criteria of 'interest', 'usefulness' and 'difficulty' 133
5.14. Class E students' ratings of each of the four treatment activities according to 'interest', 'usefulness' and 'difficulty' 133
Chapter 1

Introduction

1.1. Purpose

The present study aims to investigate the effect of form-focused, analytic instruction on control over grammatical gender by French immersion (FI) students in Grade 11. Studies have shown that while native speakers of French acquire control of grammatical gender at a very early age, and seemingly with little or no instruction (Tucker, Lambert and Rigault 1977; Karmiloff-Smith 1979), anglophone learners of French as a second language have a great deal of difficulty with this feature (Tucker, Lambert and Rigault 1977; Harley 1979; Stevens 1984; Taylor-Browne 1984). Furthermore, several studies that will be summarized in Chapter 2, Section 2.7, have demonstrated that despite years of classroom exposure to the target language, beginning as early as kindergarten or grade 1, FI students generally continue to display weaknesses in both grammatical accuracy and sociolinguistic competence. There is, however, a growing body of research (see Section 2.8 for details) providing strong evidence that form-focused instructional strategies, when used in conjunction with a primarily experiential approach, can enhance the accuracy of immersion
students' second language output. To date, most of the research dealing with form-focused instruction in French immersion has been conducted at the elementary level, up to grade 8. The present study arises partly out of the need to address the issue of form-focused instruction at the secondary level in immersion, and partly out of the author's research interests as a practising classroom teacher.

The question arises of whether it is possible to change the linguistic habits of learners who have been making the same errors with respect to a particular linguistic feature for a long time, or if the incorrect usage in fact becomes fossilized. Gass and Selinker (1994) stated that the concept of fossilization "generally refers to the cessation of learning" (p. 11). A fossilized feature, deviant from the target language norm, would continue to appear regardless of further exposure to the L2. Since it is virtually impossible to determine whether learning has ceased temporarily or permanently, Gass and Selinker suggested that stabilization might be a more appropriate term. This can be understood as an interlanguage plateau, possibly of long duration but not necessarily permanent. Further exposure to L2 input and instruction may, at some point, enable the learner to move beyond the plateau. Washburn (1994), after summarizing a number of definitions of fossilization, also noted that the cessation of learning is not necessarily permanent: "All the definitions do agree that when fossilization occurs, the second language acquisition process, or some aspect of this process, has ceased at least temporarily [my emphasis]" (p. 70).
Grade 11 was chosen for the present study in order to explore the hypothesis that it is never too late to change linguistic behaviour, even if a feature of the interlanguage appears to have been at a plateau for a long time. The aim is to determine whether students nearing the end of their immersion program can still benefit from form-focused intervention. Both Schmidt (1990) and Harley (1994a) have suggested that form-focused teaching strategies that require learners intentionally to pay attention to specific language features may be more effective with older learners than with young children, who appear less able to pay selective attention to features of a learning task.

Grammatical gender was chosen as a focus for two reasons. First, it is a topic which has fascinated and puzzled the author as both teacher and linguist for many years. Second, it is a feature which meets the conditions outlined by Harley (1993, 1994a) for aspects of second language learning that may require an analytic focus and awareness-oriented instruction. These criteria are: 1) lack of congruence with the first language, since English does not have grammatical gender; 2) lack of perceptual salience, since gender in French is most often indicated by articles (le, la, un, une), which are almost always unstressed syllables; and 3) low communicative load, since correct gender indication is seldom essential for comprehension of the message. Immersion students can understand each other and 'get the point across' without paying

\footnote{Corbett (1991:12) and others have argued convincingly that the English gender system is based entirely on semantic criteria. Oppositions such as he/she/it, his/hers/its, and himself/herself/itself reflect natural gender, not grammatical relationships. The reader is referred to Section 2.1 for a more detailed discussion of grammatical gender.}
attention to articles or adjective endings. Since they spend most of their time communicating with other immersion students, they do not often risk eliciting a negative reaction to their errors from native speakers. Only the teacher is likely to notice their mistakes on a regular basis, and he/she may be unwilling to interrupt the flow of communication for a feature which carries a low communicative load.

1.2. The nature of the study

Both quantitative and qualitative research methods were used in the present study. The quantitative component consisted of pretests, a treatment period, immediate posttests and delayed posttests in one experimental class that was taught by the researcher. The pretests, immediate posttests and delayed posttests were also administered in two comparison classes, one taught by the researcher, and the other taught by a teacher in a different school. Reactions of the students in the experimental class to the treatment materials were measured using questionnaires that asked them to rate different tasks using a five-point scale. The smaller, qualitative component of the study consisted of the researcher’s observations of the experimental class, items on the questionnaires that invited the students to add comments, and an examination of samples of oral and written responses from two of the tests.

Prior to the main study undertaken for this thesis, an exploratory study
was conducted to investigate some basic questions dealing with high school immersion students and the French gender system. It explored issues such as the frequency, type and consistency of gender errors found in the written production of students in a Grade 11 French immersion class. The purpose of the study was to provide information that would be used in the development of an instructional unit designed to assist high school FI students to improve their performance with respect to French gender.

1.3. Plan of the thesis

Chapter 2, the literature review, will summarize research that has been conducted in a number of areas that are pertinent to the thesis topic, such as the definition of grammatical gender, the acquisition of grammatical gender by both native speakers and second language learners, analytical and experimental approaches to second language teaching, and form-focused studies that have already been undertaken. Chapter 3 will describe the exploratory study that was conducted to determine the frequency, type and consistency of gender errors made by Grade 11 immersion students. Chapter 4 will explain in detail the methodology used in the main study, including the design, the participants, the testing instruments and scoring methods, the procedures and the treatment activities. Results of all pretests and posttests, and of questionnaire data, will be reported in Chapter 5. This will be followed by a brief chapter describing the researcher's observations of the students in
the experimental class during the treatment period. Finally, Chapter 7 will discuss the interpretation of the findings, implications for second language learning theory and practice, and suggestions for future research directions.
Chapter 2

Background

This chapter will review the literature that is relevant to the thesis research. It begins with a definition of grammatical gender, and touches on some historical issues, such as the disappearance of gender in English. Next a large number of studies dealing with the acquisition of grammatical gender in French and other languages by native speakers and L2 learners are summarized. The implications of these studies for second language instruction are then assessed. Analytic and experiential approaches to second language teaching are explained and compared, and evidence in favour of including form-focused instruction in French immersion programs is presented. Finally, several form-focused experimental studies that have been conducted for French and other languages are summarized.

2.1. Definition of grammatical gender

Corbett (1991) regarded gender as "the most puzzling of the grammatical categories" (p. 1). He noted that "[i]n some languages gender is central and
pervasive, while in others it is totally absent” (p.1). Fodor (1959), in a decades-
oid yet still highly relevant article, stated that “the category of grammatical
gender is one of the still unsolved puzzles of linguistic science” (p.1). He
discussed two basic and opposing hypotheses with regard to the origin of
gender in language. The debate dates back at least to the ancient Greek
grammarians, who were divided into two camps. One group saw gender as a
reflection of sex. They assumed that primitive humans had classified all beings
and objects into three categories -- masculine, feminine and neuter -- and that
this world-view was reflected in their emerging linguistic system. The opposing
group regarded gender as a grammatical category, largely devoid of semantic
content.

This fundamental division between proponents of ‘natural’ gender and
‘grammatical’ gender has persisted into modern times. In the nineteenth
century, for example, Grimm (1831) and Bopp (1883) looked upon grammatical
gender as a reflection of sex. Brugmann (1889), and later Meillet (1921) and
Martinet (1957), opposed Grimm’s and Bopp’s theories. They rejected natural
gender as a starting point and sought the origin of grammatical gender in the
inner laws of language.

Modern linguists, including Fodor (1959), have overwhelmingly favoured
the grammatical position. There is general consensus in the twentieth-century
literature that the link between natural gender and grammatical gender, in most
language families, is a tenuous one. Fodor argued that the sex theory fails to
answer one basic question: "What made it necessary to transfer perceptions about the world into grammatical form?" (p. 11). Many aspects of our worldview are not reflected in the grammatical structure of our language, so why should sexual classification be such a significant exception? Furthermore, linking grammatical gender to natural gender fails to account for the fact that the names of many sexless objects are assigned to masculine or feminine, even in languages which have a neuter (e.g. German der Tisch [m.] 'table' and die Tasse [f.] 'cup'). The disappearance of the neuter in the evolution from Latin to modern Romance languages, resulting in the reclassification of neuter nouns as either masculine or feminine, has been studied by a number of Romance philologists (e.g. Elcock 1960; Hall 1965; Posner 1966; Iordan and Manoliu 1972). The general consensus among these scholars is that the demise of the neuter gender resulted from phonological changes and the disappearance of the case system (Latin noun declensions). The fact that modern Romance languages function without the neuter, meaning that all inanimate nouns must be classified as either masculine or feminine, is cited as further evidence that grammatical gender does not reflect natural gender.

The reasons for the disappearance of grammatical gender in the evolution from Old to Modern English are similar to those which caused the demise of the neuter in the Romance languages. Old English, much like modern German, had three genders and a highly developed case system. Endings indicated the grammatical function of nouns (nominative, accusative, dative, etc.), and the relationships between nouns and other parts of speech,
such as adjectives. However, phonetic changes caused the erosion of final syllables, and this contributed to the disappearance of noun declensions. The evolving language came to rely more and more on word order, rather than agreement, to indicate the syntactic relations between sentence elements, with the result that gender no longer played an important grammatical role (see, for example, Bourcier 1981; Boiton 1982; Berndt 1983). Devoid of their grammatical purpose, gender distinctions ceased to exist.

Another argument against linking grammatical gender to sex is the fact that in many non-Indo-European languages there are more than three gender categories. Gleason (1961:228) gives the example of Bariba, a West African language with seven 'genders', which are in fact classes of nouns. Adjective endings vary for each of the seven noun classes.

A good starting point for an understanding of grammatical gender is Hockett's definition: "Genders are classes of nouns reflected in the behavior of associated words" (1958:231). Similarly, Gleason (1961) referred to "a set of syntactic subclasses of nouns primarily controlling concord" (p. 227). Concord, or agreement, means that some parts of speech, such as articles and adjectives, must agree in gender with the nouns that they accompany (e.g., le petit livre, la petite chaise). This is not to say that there is no correlation between gender and sex in languages such as French. Pairs of nouns such as le frère / la sœur, le taureau / la vache and le Canadien / la Canadienne, as well as the traditional labels 'masculine' and 'feminine', attest to the fact that
gender in French is a semantic, as well as a grammatical category. However, both synchronic and diachronic linguistic evidence point to the primacy of the grammatical function of gender. Hence it is argued that the teaching of gender to second language learners of French needs to focus essentially on its formal aspects, such as word endings that provide reliable indications of the gender of nouns, and article and adjective concord with nouns.

2.2. Native speakers' acquisition of grammatical gender

Several researchers have investigated the control of grammatical gender by native speakers of French. Tucker, Lambert and Rigault (1977), having noted that "[n]ative speakers, even very young children, have no apparent difficulty choosing the gender of nouns" (p. 13), conducted a series of experiments with 1035 francophones ranging in age from 7 to 18 years. Their hypothesis was that French speakers identify the gender of nouns on the basis of phonological information at the end of words. The subjects were required to identify the gender of both frequently and infrequently occurring nouns, and to predict the gender of pseudowords. In some experiments the nouns were presented orally, while in others they were presented both orally and in writing. The authors concluded that "French native speakers consistently assign gender to rarely occurring real nouns, to invented nouns, and to nonsense nouns in accordance with the distributional regularities of the corpus" (p. 57). The "distributional regularities of the corpus" refers to a listing, compiled at the
Université de Besançon, of 31,619 nouns in the 1962 edition of the *Petit Larousse* dictionary, congregated according to ending, and separated according to gender. For example, 99.8% of the nouns in the *Petit Larousse* ending in *-tion* are feminine, and the authors found that their subjects were almost unanimous in assigning unfamiliar nouns and pseudonouns with this ending to the feminine category. All 669 nouns ending in *-isme* are masculine, and subjects consistently assigned unfamiliar words with this ending to the masculine category. Subjects were split in their predictions of the gender of unfamiliar nouns ending in *-aire*, reflecting the fact that about 54% of nouns with this ending are masculine, and 46% feminine. Tucker, Lambert and Rigault observed that about 95% of the time, French-speaking children and adolescents will correctly predict the gender of an unknown noun. Furthermore, they claimed that "[t]his process, whatever it may entail, has certainly been mastered before the child begins formal schooling" (p.14), since even their youngest subjects obtained high scores on the tests.

Karmiloff-Smith (1979) agreed that native speakers' control of the French gender system is well established at an early age. As part of a large-scale investigation of determiners, she included a series of experiments involving oral use of gender. The 341 francophone children who participated in her studies were between the ages of 3 years 2 months and 12 years 5 months. She observed: "As early as 3 to 4 years ... the child constructed a very powerful, implicit system of phonological rules, based on the consistency, but not necessarily on the frequency, of phonological changes in word endings."
Classification by gender was clearly based on suffixes" (p. 167). When children were presented orally with syntactic clues (indefinite articles) or semantic clues (people depicted in drawings) that conflicted with phonological word ending clues, they tended to rely on the phonological information to predict gender. Only in the case of 'arbitrary' nouns (i.e. those with endings not clearly predictive of masculine or feminine) did the children pay more attention to the other clues. The nine-year-olds paid somewhat more attention to natural gender and syntactic clues than the younger children, but word endings remained the dominant influence for all age groups. Karmiloff-Smith concluded that "the phonological procedure remains predominant in French because it is through phonology that gender presents the most consistent patterns" (p. 169).

While Surridge (1989) appeared to accept that Tucker, Lambert and Rigault "ont démontré que le genre en français est attribué par les locuteurs natifs suivant une distribution qui dépend du phone final" (p. 664), she cautioned that the morphological structure of "suffixés" (nouns with suffixes) and "composés" (compound nouns) entails gender attribution rules that sometimes contradict the phonological clues. For example, 85% of nouns ending in [j] such as feuille and oreille are feminine (see Tucker, Lambert and Rigault, 1977, p. 79); however, portefeuille is masculine because compound nouns consisting of a verbal and a nominal element are almost without exception masculine in French. Surridge pointed out that when phonological and morphological clues are in conflict, "On doit donner la priorité aux règles morphologiques" (p.672).
Carroll (1989) claimed that morphological analysis is crucial to the native speaker's competence with respect to grammatical gender: "One need not encounter literally every noun in the language to know which gender it gets if one can do a morphological analysis of a word" (p. 567), and "Francophones can do morphological analysis and can assign gender to words on the basis of the gender of derivational suffixes" (p. 585). Carroll argued that the psycholinguistic processes used by native speakers to attribute gender to nouns cannot be explained simply as phonological analysis, as claimed by Tucker, Lambert and Rigault (1977), but must involve a complex interaction of phonological, morphological and syntactic features. However, her reference to "derivational suffixes" indicates that she agreed that noun endings are a key component of gender attribution.

Several studies of other languages have arrived at similar conclusions to the research on French. Perez-Pereira (1991), in a study of the early acquisition of grammatical gender by native speakers of Spanish, concluded that it is primarily intralinguistic (formal), rather than extralinguistic (semantic) criteria that are used by children to assign gender to unfamiliar nouns. Similarly, Delisle (1985) found that native speakers of German assign gender primarily on morphophonological criteria. Popova (1973) and Zakharova (1973), in experimental studies with Russian children up to the age of seven, demonstrated that phonological strategies were crucial in developing a system
of gender markers. Irregular nouns (i.e. those whose endings do not clearly predict masculine or feminine gender) presented difficulties even for the oldest children in the studies. Levy (1981a, 1981b), in two experiments involving monolingual Hebrew-speaking nursery school children, concluded that phonological criteria outweighed semantic criteria when the children assigned gender to an unfamiliar noun.

Hulstijn and Zekhnini (1994), in a study of the acquisition of nonarbitrary (predictable on the basis of rules or cues) and arbitrary (not predictable on the basis of rules or cues) gender specifications in Dutch, concluded that “[w]hen learning new nouns in their native language, native speakers of Dutch make fewer gender errors learning nouns with nonarbitrary gender than when learning nouns with arbitrary gender” (p. 10). In other words, for nonarbitrary nouns, native speakers make use of morphological and phonological statistical cues to predict the gender. The word ‘statistical’ is key to an understanding of Hulstijn and Zekhnini’s arguments. In opposition to Carroll (1989), who maintained that the process of gender learning by native and nonnative (anglophone) learners of French was fundamentally different, Hulstijn and Zekhnini argued that natives and nonnatives process new nouns in Dutch in basically the same way as far as gender is concerned. After comparing the results of native and nonnative speakers, Hulstijn and Zekhnini (1994) concluded that “[n]ative speakers of Dutch are superior to nonnative speakers of Dutch only when learning the gender of nouns with nonarbitrary gender. They are not superior, however, when learning nouns with arbitrary gender” (p. 10).
The authors attributed native speakers' superiority with nonarbitrary gender to the fact that they have had much more input than nonnatives, hence their vocabularies are much larger and they have accumulated stronger gender cues. The accumulation of vocabulary items provides them with the statistical information needed to predict the gender of new nonarbitrary items (e.g., if one already knows 100 words with the same ending, and 95 of them are masculine, it is logical to assign masculine gender to a new noun having the same ending).

All of the studies that have been reviewed in this section provide evidence that native speakers make use of word ending regularities (phonological, orthographic and morphological) to assign gender to unknown nouns, and that they learn to do so with little or no formal instruction.

2.3. L2 learners' acquisition of grammatical gender

Tucker, Lambert and Rigault (1977) contrasted the ease with which native speakers acquire grammatical gender in French to the difficulties encountered by nonnatives:

From a practical point of view, the necessity to master gender may be the single most frustrating and difficult part of the study of French as a second language, both for those whose native language lacks gender distinctions, and for those whose language makes use of some other basis for such distinctions. The apparently arbitrary classification system often proves to be a source of complete mystery, even to those who have studied the language for a number of years. Native French speakers, on the other hand, even the very young ones, appear to have no difficulty
mastering this feature of their language (p. 11).

Tucker, Lambert and Rigault (1977) conducted research with 139 English-speaking students at McGill University who were studying French as a second language. The students were divided into three subgroups: Introductory (those who had studied no French), Intermediate (those who had studied French for approximately seven years, but whose achievement on standard tests was poor), and Advanced (French honours majors or the equivalent). The authors found that the advanced students were "essentially as skilled as native speakers in making gender choices" (p. 60). It should be noted that the real and invented nouns used in the study all had endings that were "statistically reliable markers" (p. 24) of gender. Intermediate level students, even though they had studied French for about seven years, did not perform nearly as well as native speakers or advanced students. Curiously, introductory level students performed better than intermediate students, though not as well as advanced level students. Further investigation pointed to essential differences in the way the introductory and intermediate level students were taught. The intermediate students were very seldom corrected for making gender errors, as their teacher was more concerned with comprehension and fluency. The introductory level students had been exposed to a very limited sample of vocabulary items chosen for the most part from Le Français fondamental, but their teacher had told them about word ending regularities related to gender, insofar as she was aware of them. This led Tucker, Lambert and Rigault to suggest that:
...it should be possible to simplify the problem of gender acquisition for the non-native simply by structuring the manner of presenting vocabulary items. For example, words with distinctive ending characteristics could be congregated to highlight the regularities so that the student could infer the underlying regularities for himself in a fashion analogous to the native speaker of the language. To facilitate the concept attainment process, the existing regularities could also be called to the student's attention by the teacher" (p. 60).

In effect, the authors were calling for a form-focused approach to teaching grammatical gender to second language students of French.

A study of oral interview data by Harley (1979) demonstrated that children in early French immersion programs had great difficulty with gender, whereas monolingual francophones experienced few problems. Interestingly, the performance of bilingual Franco-Ontarian children appeared to fall in between the other two groups, with those from English-dominant homes more likely to make errors similar to those of French immersion students than those from French-dominant homes. Harley also reported that the immersion students tended to overgeneralize the third person masculine pronoun. Another characteristic was the neutralization of the phonological distinction between *la* and *le*, indicating either a strategy to avoid making gender choices, or simply a transfer of the English phonetic rule of reducing unstressed vowels.

Stevens (1984) examined the strategies used by anglophone children aged six to thirteen to acquire gender in French, and compared these with the strategies observed by Karmiloff-Smith (1979) for native speakers. The
anglophone students had spent from one to six years in a French immersion program. Stevens (1984) observed that the pattern of development of the anglophone children in many ways paralleled that of the francophones in Karmiloff-Smith's study: "Both groups used the phonological endings of nouns as a basis for determining gender and both had recourse to semantic and syntactic clues" (p. 149). However, the anglophones were not as accurate as the francophones in assigning gender, and their errors were overwhelmingly with feminine nouns. Stevens noted that the anglophone children were particularly weak with respect to the agreement of adjectives: "It was found that adjective concord was almost never made by any of the children, the masculine form of the adjective being used in most instances" (p. 106).

Tarone, Frauenfelder and Selinker (1976) conducted an analysis of tape-recorded interviews with Grade 2 French immersion students. These included conversations, picture descriptions and story-telling. They found overgeneralization of the masculine, specifically with respect to third person pronoun usage. Spilka (1976) analyzed free speech samples of immersion students in grades one through six and reported that little progress appeared to be made with gender control over the years.

Taylor-Browne (1984) conducted eleven different experimental tasks with subgroups from a total of 158 students from Grade 3 early total immersion, Grades 7, 8 and 9 continuing partial immersion, and Grade 9 late immersion. Many of the tasks were modelled after those used by Karmiloff-Smith (1979)
and Tucker, Lambert and Rigault (1977). Results showed that “unlike francophones, the majority of the anglophone immersion subjects were unable to use any type of phonological, syntactic or semantic clue to determine the gender of novel feminine nouns. Instead, masculine forms were used in almost all cases” (p. iii). Taylor-Browne concluded that the extremely poor scores provided evidence that “a large number of anglophone subjects are simply not attending to gender distinctions regardless of whether the mode of presentation is oral or graphic” (p. 127). She also suggested that “alternative instructional procedures may be required if anglophone children in immersion programmes are to acquire correct control of this particular grammatical category” (p. iii).

Carroll (1989) argued that English-speaking children do not develop the Francophone’s system for gender attribution because English does not have morphological gender, and therefore “[t]here is no trigger for the gender feature on the noun class” (p. 576). “I also argue that the inherent feature triggered in native acquisition cannot be available for children as old as 4 or 5 years” (p. 581). Carroll proposed that since anglophones do not have instantaneous access to gender information when they produce sentences, they need to develop mnemonic strategies and rules of thumb which “could provide the advanced learner not only with a reasonably accurate system but also with a mechanism for guessing the gender of new items” (p. 580).

Marinova-Todd (1994) compared the acquisition of gender in French by L1 speakers of English and German. Her first hypothesis was that learners
whose L1 is unmarked for gender (i.e. English) would display a critical period in acquiring the gender category in French. Learners whose L1 is marked for gender (i.e. German) would not have a critical period since they are familiar with the concept in their native language. The critical period is generally regarded as ending with puberty (e.g. Lenneberg 1967). Marinova-Todd’s experimental study with 26 female university students (18 anglophones and 8 native speakers of German) did not support the hypothesis. Participants who started learning French after puberty scored higher on two tasks requiring knowledge of French gender than did participants who started learning French before puberty. One explanation suggested by Marinova-Todd for this result was that most of the participants who started learning French before puberty were in French immersion programs “where the teaching is not concentrated on learning formal grammatical rules” (p. 31) and where “the less structured immersion environment also results in students having their grammatical mistakes corrected only very rarely” (p. 32). The participants who started learning French after puberty “were typically first exposed to the language in junior or senior high-school where teaching of formal grammatical rules is emphasized. Students were urged to learn and apply the grammar of the L2 and were often corrected for their mistakes” (p. 32). Marinova-Todd also observed that for the participants who started learning French after puberty, the German native speakers did better than the English native speakers, presumably because they were already familiar with the concept of gender in their L1. It should be noted, however, that the number of participants in the
post-puberty category was very small (3 English speakers and 6 German
speakers); thus the results are tentative and experiments with a larger sample
need to be conducted.

Marinova-Todd's second hypothesis was that English-speakers would
use a phonological strategy for assigning grammatical gender in French,
whereas German-speakers would use a combination of phonological, syntactic
and semantic strategies. She found that when participants were presented with
tasks where noun suffixes were inconsistent either with natural gender or with
the gender of the article (e.g. un goltine), German speakers used predominantly
syntactic and semantic strategies to assign gender. English speakers were
divided. Those who were “more proficient in French” (p. 33) used mainly the
phonological strategy, while “the less proficient ones” (p. 35) used
predominantly semantic and syntactic strategies. (Note that proficiency here
refers only to the participants’ scores on the gender assignment tasks. The
“less proficient” subjects were, for the most part, those who started learning
French before puberty in FI programs.) Marinova-Todd suggests that the
German speakers, because of their familiarity with the concept of gender in their
native language, were more sophisticated and more flexible when assigning
gender and thus were able to make use of a combination of strategies. They
recognized that the phonological rules, though useful, do not always apply, and
that other cues have to be taken into consideration when gender is assigned to
an unfamiliar noun. The English-speakers were more likely to rely on a single
strategy, and those who used mainly the phonological strategy were more
successful than those who used the semantic and syntactic strategies. Another finding in Marinova-Todd's study was a systematic trend on the part of the subjects to assign masculine gender more frequently than feminine. She attributed this to the use of the masculine as an unmarked or default morphological form which is more likely to be chosen than the marked feminine form when learners are uncertain of a noun's gender.

There is some research indicating that languages other than French present difficulties for the English-speaker with respect to gender. Delisle (1985), in a study of the strategies used by anglophones when assigning gender to German nouns, found that they tended to base their decisions on semantic, rather than formal considerations. This resulted in a high rate of errors.

In sum, it appears that children who learn French in a primarily experiential classroom setting make little progress in terms of improving their ability to correctly assign gender to French nouns (Spilka 1976; Tarone, Frauenfelder and Selinker 1976; Harley 1979; Taylor-Browne 1984; Marinova-Todd 1994). While Stevens (1984) observed that the pattern of development of anglophone learners paralleled that of francophones, and that both groups made use of phonological, syntactic and semantic clues to determine gender, she also noted that anglophone children were considerably less accurate in assigning gender than French native speakers, and that they were particularly weak with respect to adjective concord. A study by Harley (in press), which will
be outlined in Section 2.8, reported some success with form-focused tasks
designed to familiarize Grade 2 immersion students with the formal clues to
gender. The research on older learners (Tucker, Lambert and Rigault 1977;
Marinova-Todd 1994) appears to indicate that students who have been taught
French in an essentially analytic way (see Section 2.6 for a discussion of
analytic and experiential language teaching) and who have had their attention
drawn to word ending regularities, are more successful at assigning gender
than those whose classroom background in French has been mainly
experiential. Furthermore, anglophone learners, because of the absence of
grammatical gender in their L1, appear to be at a particular disadvantage when
dealing with this concept in a second language (Marinova-Todd 1994; Delisle
1985). The evidence seems to point clearly to the need for form-focused
instructional intervention to assist anglophone learners of French with the
category of gender. The present study proposes to investigate the efficacy of
such intervention at the high school level with F1 students.

2.4. Reactions of francophones to learners' gender errors

While gender errors seldom impede communication of the message, they
may be a source of irritation to native speakers. Magnan (1983) conducted a
study of age differences and sensitivity to gender in French. She found that, on
the whole, adolescents were more sensitive to and irritated by L2 learners'
gender and agreement errors than were adults, who tended to focus on other
features such as verbs and prepositions. However, all age groups expressed some irritation with non-native speakers' gender errors. Lepicq (1980) found that adult francophones rated gender mistakes as the least serious type of error made by eleven-year-old students in French immersion programs. Since the present study focuses on high school students, Magnan's finding with respect to adolescent francophones is more relevant. Students need to be aware that while inattention to gender will not prevent them from communicating with francophones, it may cause them to make a less favourable impression, especially with their peer group.

2.5. Implications for L2 instruction

The studies by Tucker, Lambert and Rigault (1977) and Karmiloff-Smith (1979) support the view that the assignment of gender in French is not haphazard or unpredictable, and that for most nouns (possibly up to 95%) francophones are able to use phonological, morphological and orthographic word-ending clues to correctly predict the gender of unfamiliar nouns. Furthermore, many of the exceptions to the word-ending rules are either nouns of very high frequency, and thus easily remembered, or nouns whose grammatical gender conforms to natural gender. It is only for unfamiliar nouns that are exceptions to the word-ending clues and that do not contain semantic clues that francophones are likely to make errors.
While native speakers acquire the word-ending clues to French gender seemingly without intentional learning, anglophones in a classroom setting experience much difficulty even when they have received a great deal of comprehensible input from a young age, as is the case of early French immersion students (see Section 2.3 for a review of the relevant studies). It would therefore seem worthwhile to explore the possibility that explicit, form-focused instruction in the phonological and written-ending clues could assist non-native speakers to improve their performance. Tucker, Lambert and Rigault (1977) conducted a pilot study with 154 students in the Intermediate French program at McGill University. These were students who had previously studied French for approximately seven years, but whose achievement on standard tests was poor. The students were divided into groups and given vocabulary training with nouns which were either congregated in lists according to ending and gender, or were randomly mixed. In addition, two groups were given explicit information about regularities, such as words ending with -ation being feminine. When pretest and posttest results were compared, only one group showed any improvement. It was one of the two groups that had been provided with organized lists and explicit information about predictable endings. While tentative, this result suggested that an instructional strategy that is both form-focused (i.e. organizes vocabulary items according to ending rather than meaning or function) and explicit may assist non-natives to make gender assignments more accurately. Unfortunately, Tucker, Lambert and Rigault (1977) provide limited information about some aspects of the study, such as the
length of the treatment period.

Surridge and Lessard (1984) tested the grammatical gender knowledge of adult learners of French. The participants were 113 students in their final year of French in six Canadian universities. They were asked to identify the gender of 60 nouns presented in written form. Participants scored well in determining the gender of frequently occurring nouns (94.6%). The figure decreased to 83.5% for nouns which are less frequent in occurrence, but whose gender is relatively easy to predict according to their endings. A further decrease to 66.4% was noted for nouns which not only occur infrequently, but have the gender which is either the opposite of the anticipated according to word-ending clues, or which would be difficult to predict since the ending is not strongly associated with either gender. The 17% difference between the results for predictable and non-predictable nouns would appear to indicate some awareness on the part of the students of the importance of word-ending clues for predicting the gender of unfamiliar nouns. However, the study did not explore the possibility that instruction focusing specifically on word endings might enable the students to score higher on subsequent testing.

Three studies by Hardison (1992) demonstrated that elementary (i.e. beginners and near beginners) and intermediate university-level L2 learners of French were able to take advantage of correspondences between gender and noun endings in making gender decisions. The author noted that “these students generally have had significant, explicit instruction in gender-noun
ending correspondences" (p. 294), although the exact nature of the instruction is not explained. The students were presented with vocabulary items having typically masculine (e.g. manteau) and feminine (e.g. cuisine) endings and asked to identify their gender. In one test the students only heard the words; in the others they received both oral and written clues. Exceptions such as peau (f.) and magazine (m.) were sometimes included. It was hypothesized that students would assign gender incorrectly to the exceptions as a result of applying their knowledge of oral and written word-ending clues. Results ranged from 69% to 84% depending upon three factors: 1) student level (intermediates outperformed beginners); 2) method of presentation (scores were generally higher when the items were presented not just orally, but in written form as well; and 3) the number of exceptions included in the test (the more the exceptions, the higher the rate of error). These results suggested that learners can improve their control over gender with practice, and that orthographic as well as auditory clues are important at least for older learners. Hardison also concluded "that learners appear to formulate rules of gender-noun ending correspondence based on the gender of the most salient member of each phonetic ending category in their input, even though the most salient member may constitute an exception to the gender pattern in the language as a whole" (p. 296). In other words, the students tended to focus on the noun with which they were most familiar as their exemplar for a particular ending. For example, more than 80% of nouns ending in -ème are masculine, but a student for whom crème (f.) is the most salient noun with this ending will tend to associate it with feminine
gender.

Luce (1979) completed the very useful task of categorizing written word-ending clues for the benefit of L2 learners. He included probabilities, such as "The -ie group, comprising some 1500 feminine nouns with better than 100:1 reliability" (p. 570). Some semantic guidelines were mentioned as well, such as months, days, seasons, colours and most trees being masculine, while fruits and cars tend to be feminine. Luce claimed that his guidelines (including lists of common exceptions, such as un incendie) "will identify the gender of close to 95% of the nouns listed in Le Petit Robert" (p. 568). Note that this percentage coincides with Tucker, Lambert and Rigault's (1977) observation that native French speakers will correctly predict the gender of an unknown noun 95% of the time (see Section 2.2). While Luce cautioned that "no noun can be automatically assigned to a certain gender because of its ending alone, for its meaning and form -- whether a simple or compound word -- have priority" (p. 570), it is clear that he regarded word endings as the most productive clues to gender, enabling learners to predict the gender of far more nouns than semantic or morphological information.

The studies by Tucker, Lambert and Rigault (1977), Surridge and Lessard (1984), and Hardison (1992) all demonstrated that awareness of word ending regularities can assist adult anglophone learners of French to correctly assign gender to French nouns. Two of these studies (Tucker, Lambert and Rigault [1977] and Hardison [1992]) also provided evidence that explicit
instruction in gender-noun ending correspondences can help students to improve their performance on gender identification tasks. The work completed by Luce (1979) provides learners with an organized and relatively simple classification system of the noun endings and their predictiveness of gender, with some semantic guidelines as well. The present research applies this information to a high school FI setting. The Grade 11 participants in the experimental class are exposed to explicit instruction focusing primarily on the formal clues to noun gender in French.

2.6. *Analytic and experiential approaches to second language teaching.*

Two basic orientations in language teaching have been labelled *experiential* and *analytic* (Allen, Swain, Harley and Cummins 1990; Stern 1990). The experiential approach is characterized as content-oriented, with students working on substantive topics or themes. They use the target language as they work on a variety of purposeful activities that are not geared exclusively to L2 learning. For example, in a French immersion program, instruction in many or all subjects such as geography, science and physical education takes place entirely in French. The students use French source materials, write their assignments and tests in French, and (theoretically!) use French as the only language of communication both with teachers and with one another. The experiential philosophy is that language learning takes place
largely incidentally, without formal instruction. Fluency and the communication of message are given priority over accuracy and error avoidance. Exposure to authentic language use, and the opportunity to communicate in a wide variety of situations, are intended to replicate as closely as possible, in the classroom, the acquisition process by which children learn their first language. The experiential approach has been criticized because the conditions of natural language use can never be totally reproduced in a classroom setting. Hence students may rarely if ever be exposed to certain linguistic features, language functions, types of discourse, and so on (Swain 1985). Even features to which students are exposed may not become salient for them. Since error correction is not stressed in the experiential approach, students repeatedly make the same mistakes and incorrect forms may appear to be fossilized, and these incorrect forms may serve as input to other learners (Lightbown and Spada 1993). Stern (1990:105) also pointed out that the activities in experiential classes are sometimes linguistically too demanding for the proficiency level of the learners. This can result in frustration, and students may revert to their first language in order to express their ideas.

The analytic orientation is a cognitive, systematic approach to language teaching. Language features are isolated and made salient for the learner. Rules, regularities and exceptions are noted, and students may practise specific topics, such as irregular present tense verbs, or possessive adjectives, through structured drills and exercises. Accuracy and error correction are important features of this orientation. Stern (1990) noted that critics of the analytic
strategy have often defined it far too narrowly, associating it exclusively with grammar rules and drills, whereas it "also refers to any other aspect of the language that can be identified and isolated, phonological, lexical, semantic, discoursal, and sociolinguistic" (p. 98). Analytic teaching has been criticized for decontextualizing and fragmenting language. Classroom practice of specific features does not automatically carry over to real situations. Fluency may suffer if too much emphasis is placed on accuracy, and students may not develop the confidence to take risks with the language and use it outside the classroom.

Allen et al. (1990) have pointed out that "in practice many teachers draw selectively upon the two approaches [analytic and experiential] in order to develop composite or mixed methodologies that are tailor-made for particular instructional settings" (p. 58). This pragmatic approach on the part of practitioners has in fact received theoretical support in the form of numerous proposals for comprehensive, integrated, multidimensional approaches to second language teaching, incorporating both analytic and experiential aspects. For example, Allen (1983) proposed a three-level, variable focus approach to L2 curriculum combining structural-analytic, functional-analytic, and non-analytic (experiential) components. The structural-analytic component focuses on grammar and other formal features of the language. The functional-analytic component focuses on communicative functions (such as requesting, instructing and advising), discourse, and sociolinguistics (the appropriateness of language to particular social contexts). The experiential component focuses on the spontaneous, unanalyzed use of language in a natural communicative
setting (insofar as this is possible in the classroom).

Stern (1983, 1992) proposed a multidimensional curriculum integrating four syllabuses: language, communicative, cultural, and general language education. The language syllabus is an analytic component comprising both structural and functional aspects. The cultural syllabus also contains analytic components, since it proposes that students observe and analyze the target culture as well as become familiar with it experientially. The general language education syllabus is also largely analytic in nature, and aims to increase students’ awareness of language in general and the language learning process. Allen's and Stern's models are examples of a philosophy that recognizes second language learning as a process that is too complex to be encompassed by any single orientation.

It is evident from the preceding discussion that the term analytic can encompass many aspects of the language learning process, from the formal structural analysis of linguistic features through discourse and socioinguistics to the investigation of the target culture and the exploration of communication in general. Throughout the present discussion the term form-focused is used to refer specifically to those aspects of the analytic orientation that are included in Allen's (1983) structural-analytic approach and the structural component of Stern's (1983, 1992) language syllabus. Lightbown and Spada (1993) defined form-focused instruction as "[i]ntuction which draws attention to the forms and structures of the language within the context of communicative interaction. This
may be done by giving metalinguistic information, simply highlighting the form in question, or by providing corrective feedback" (p. 122). This is a good working definition for the purposes of the present study, since all three strategies mentioned by Lightbown and Spada are incorporated into the treatment activities.

2.7. The role of form-focused instruction in French immersion

Traditionally, French immersion has been regarded fundamentally as an experiential approach to L2 learning. It emphasizes the exclusive and authentic use of the target language; students engage in a variety of communicative activities; tasks dealing with a wide range of topics focus on message and content. Assessments of the overall effectiveness of immersion programs have tended to be positive (see, for example, Swain and Lapkin 1982; Genesee 1984; Krashen 1984). Researchers have reported that neither the English-language skills nor the cognitive development of most students participating in French immersion programs are adversely affected, thus allaying the initial apprehensions of many parents and educators. Furthermore, test results have shown immersion students outperforming core French students in all aspects of second language learning. In terms of receptive French skills (listening and reading), immersion students have often scored as well or nearly as well as francophones on global tests of comprehension by grade (e.g. Swain, Lapkin
and Andrew 1981; Morrison, Pawley and Bonyun 1982; Swain and Lapkin 1982; Lapkin and Swain 1984a,b; Swain and Lapkin 1986).

Within this generally positive context, however, many researchers have noted that the proficiency and accuracy of many immersion students' productive (speaking and writing) skills remain far from native-like, despite many years of intensive exposure to French (see, for example, Swain 1985; Calvé 1986; Swain and Lapkin 1986; Cummins 1987; Lyster 1987; Harley 1993). Several explanations for this situation have been offered. Immersion students know that they can 'get by' in the classroom and be understood by their peers and teachers even though their speech and writing may contain many anglicisms and grammatical errors (Lyster 1987). Swain (1993) suggested that learners are not "pushed" enough to make use of their linguistic resources to produce accurate output. Harley (1993) stated that the "relatively restricted range of language functions that are naturally performed in a classroom context may also limit the range of vocabulary and types of discourse to which students are exposed. Even features that occur relatively frequently in natural teacher talk may still not be perceptually salient to L2 learners" (p. 247). Harley (1994a), in a paper not limited to French immersion, reviewed instructional studies that provide "growing support for the view that some level of awareness is necessary for second language learning" (p. 19).

Two factors appear to be at work: motivation and awareness. Learners will not produce formally accurate output unless they are motivated (externally
and/or internally) to do so. Learners cannot produce correct output if they are unaware of certain features of the target language. As Harley (1993) pointed out, for some linguistic features (être and avoir, gender distinctions, and lexical development are specifically mentioned), frequency in the input is not sufficient for learning to occur. She therefore suggested that analytic strategies are necessary to complement the experiential emphasis in immersion classrooms. Lyster (1990) also supported the inclusion of an analytic focus in immersion programs to "complement the already well established experiential component and counteract weaknesses in grammatical and sociolinguistic competence which persist among FL students" (p. 159). The immersion observation study (Swain and Carroll 1987), which was part of the large-scale Development of Bilingual Proficiency report (Harley, Allen, Cummins and Swain 1987) also strongly suggested the need for more focused language teaching in immersion classrooms, since students were observed to make a high proportion of errors with frequently used grammatical features such as gender, articles and verbs. Since students' errors were generally not corrected in a consistent or systematic way, there appeared to be little indication that students were being motivated to produce more accurate output, or indeed that they were aware of the mistakes they were making.

---

The Development of Bilingual Proficiency report presented the results of a five-year research project conducted in the Modern Language Centre at the Ontario Institute for Studies in Education. The purpose of the project was to examine a number of issues (the nature of language proficiency, the effect of classroom treatment, social factors and proficiency, age and proficiency) concerning the second language development of school-age children.
2.8. **Form-focused experimental studies**

There is growing empirical evidence that form-focused instruction can enhance second language learners' linguistic performance. Studies have included some that deal specifically with French as a second/foreign language. Tomasello and Herron (1988), for example, compared two methods for teaching grammatical exceptions in the language classroom. Thirty-nine students in an introductory French course at the university level in the United States participated in the study. Eight structures which were exceptions to a grammatical pattern were chosen. These included the use of the usually masculine possessive *mon* before a feminine noun beginning with a vowel, the irregular comparative *meilleur* (in place of "*plus bon"), the irregular second person plural *vous dites*, and so on. Two different teaching formats were chosen. In one case the students were simply told that the structures were exceptions. In the other case (dubbed the Garden Path method) they were presented with canonical exemplars encouraging them to induce the rule. The students were then told to generate the exceptional form, and their resulting overgeneralization errors were corrected. The group who were taught using the Garden Path method scored significantly higher than the other group on tests that were administered at three intervals after the instructional period ended. One explanation offered by Tomasello and Herron for the success of the error-inducing method was that it "helped students to focus attention both on the rule
and on the features of the particular structure that marked it as an exception" (p. 237).

Harley, Howard and Roberge (1996), in collaboration with two classroom teachers, explored the role of direct vocabulary teaching among Grade 9 extended French students and Grade 11 early immersion students in the Metropolitan Toronto area. The activities that were designed had two main focuses: thematically related words (e.g. surnaturel, fantôme, sorcellerie, spectre), and derivationally related words (e.g. clair, clairière, clairement, s'éclaircir, éclairage). One activity for thematically related words was semantic mapping. The students were given outlines for semantic maps and a central concept for each. Their task was to find related vocabulary items to fill out the branches in the network. For the word families, the students were provided with a stimulus word and told to find derivationally related words and to categorize them according to part of speech. The instructional period lasted from two to three weeks. When pretest and posttest scores were compared, the Grade 9 students made statistically significant gains on four tests measuring various aspects of vocabulary knowledge, and the Grade 11's made statistically significant gains on three of four measures. The authors cautioned that their study was exploratory in nature, and that they could not claim to have firm evidence that direct teaching of vocabulary with exposure leads to superior results than exposure alone. However, the students did make substantial gains, and both students and teachers considered the activities to have been, in general, useful and enjoyable. It should be noted that the activities required the
students to focus both on meaning and on form.

Harley (1989) conducted an experimental study with Grade 6 immersion students to determine the effect of curriculum materials dealing with the functional differences between two past tenses, the **passé composé** and the **imparfait**. An experimental and a comparison group were compared on three measures -- a cloze test, an oral interview and a composition. The experimental group outperformed the comparison group on the immediate posttest for the first two measures, but no statistically significant differences between the two groups were revealed for the composition task. On the delayed posttest, three months later, no significant differences were found between groups on any of the measures. However, this does not mean that the form-focused instruction was ineffective, since the scores for both groups improved between immediate and delayed posttesting, thus preventing statistically significant differences. Apparently, the comparison teachers had also devoted a considerable amount of time to teaching the **passé composé** and **imparfait**, though with different materials. Harley also suggested that the amount of time devoted to the topic in the experimental classes (twelve hours spread over eight weeks) may have been insufficient.

Day and Shapson (1991) conducted an experimental study in French immersion at the Grade 7 level involving the use of the conditional to express hypothetical situations (**Si j'avais un million de dollars , j'achèterais ...**) and as a politeness marker (**Je voudrais ...**). The activities, spread over a six week
period, all centred around the planning of a space colony, and included oral reports, written reports, the preparation of newspaper articles, linguistic games, and co-operative group work. There were six experimental and six comparison classes. On the immediate posttest, the experimental group scored significantly higher than the comparison group on a cloze test and written composition, but not on an oral interview. This significant difference on two measures was maintained eleven weeks later on the delayed posttest. Progress was also made by the experimental classes on the oral interview, but test scores were not statistically significant largely because one comparison class did very well on this measure. The teacher of this class had spent considerable time teaching verb forms and functions, including the conditional.

Lyster (1993, 1994) investigated the effect of functional-analytic teaching on the sociolinguistic competence of French immersion students in Grade 8. The study included the ability to use vous appropriately and accurately in formal situations and formal letters, the ability to use polite closings in formal letters, and the ability to recognize contextually appropriate French. Curriculum materials were implemented in three experimental classes over an average of 4.7 weeks. Two comparison classes also took part in the testing. The experimental group demonstrated substantial improvement on all three abilities, while the comparison group did not. Lyster (1993) concluded that the results "support the inclusion of analytic language teaching in the L2 curriculum" (p. 217).
Harley (in press) investigated the role of form-focused tasks in promoting the second language acquisition of children in Grade 2 immersion. Specifically, the students were exposed to classroom activities designed to draw their attention to formal clues of gender -- determiners and noun endings. Harley proposed several reasons for students' difficulty with gender, all of which relate to its lack of salience for the English-speaking learner. Since grammatical gender is not an issue in English, immersion students are generally not attuned to its importance in the L2. They may not notice the distinction between *le* and *la*, and between *un* and *une*, which are the most common indicators of noun gender in French. This is because vowel quality in English in unstressed syllables (articles are seldom stressed) can vary considerably and is usually reduced to the neutral mid-vowel known as *schwa*. If students do not notice the difference between the articles, they cannot produce the difference in their own speech, or they may deliberately develop an ambiguous pronunciation, approximating both the masculine and the feminine, but not clearly either. This strategy enables them to avoid paying attention to noun gender. Since gender carries a low communicative load, students have little motivation to pay attention to it. Incorrect input from peers, as well as lack of correction by teachers who may be focusing instead on subject matter content, may also contribute to gender's lack of salience for anglophone learners. Six experimental classes participated in Harley's study. Classroom activities included games, songs and rhymes, and the use of picture cards. The treatment period lasted for five weeks, with some time devoted to gender every
During the first two weeks, the emphasis was on the definite and indefinite articles, with the objective of making students aware that gender is an integral part of all French nouns. In the last three weeks, emphasis was placed on the word-ending clues to gender, such as /o/ and /ö/ for masculine, and /ɛ/ and /ɛz/ for feminine. Four tests, two listening and two oral production, were developed to assess the effects of the treatment. These tests were administered three times (pretest, immediate posttest, delayed posttest). Six comparison classes also took part in the delayed posttest. Test results indicated that students who received gender instruction outperformed those who did not. However, the results did not indicate that the instruction enabled students to generalize knowledge about noun endings to unfamiliar nouns. In other words, the treatment appears to have induced item learning rather than system learning. Interviews with the students revealed that some of them were consciously aware of the relevance of noun endings even if they had not answered correctly during the testing. This suggests that some explicit knowledge had been gained but that it had not been incorporated into their interlanguage system.

Form-focused studies have been carried out for languages other than French, and particularly in English as a Second Language. For example, Doughty (1991) conducted a study of the acquisition of relativization (relative clauses) in English by 20 university students with a variety of first language backgrounds. The students were divided into three groups, two experimental and one control. One experimental group participated in exposure to relative clauses plus meaning-oriented instructional treatment. The other received
exposure plus rule-oriented instructional treatment. The control group was exposed to the same relativization material but without any instructional treatment. Both experimental groups improved significantly more on the posttest over the pretest evaluation than did the control group, leading Doughty to conclude that "SL [second language] instruction was shown to have a positive effect on the rate of acquisition of relativization in English" (p. 463). It should be noted that both the meaning-oriented treatment and the rule-oriented treatment required the students to focus on form. The rule-oriented group were provided with explicit rule statements and on-screen sentence manipulation (lessons were conducted in a computer lab). The meaning-oriented group were provided with rephrasings of the sentences containing relative clauses, on-screen sentence manipulation, and highlighting and capitalization of the key components of relativization (head nouns and relative pronouns), but no explicit rule statements. Since both experimental groups improved almost equally, and did so significantly more than the control students, Doughty speculated that perceptual saliency, rather than metalinguistic rule statements, may have been crucial to the success of the instructional treatment.

White (1991) and White, Spada, Lightbown and Ranta (1991) investigated the effects of form-focused instruction and corrective feedback on the development of specific linguistic structures in the English of francophone ESL students in Quebec. White's study dealt with adverb placement as a grammatical feature that differs in English and French. Her hypothesis was that francophones need explicit instruction in the position of adverbs in English.
sentences or they will persist in using the rules for French. Experimental classes that received the instruction dramatically outperformed comparison classes on both immediate and delayed posttests. However, in follow-up tests a year later, the gains made by the experimental learners had disappeared, and their performance on adverb placement was the same as that of the comparison group. White's study introduces a cautionary note into the discussion of form-focused instruction. While studies generally show positive short-term results (even delayed posttests usually take place quite soon after the treatment period), long-term results may be less encouraging. Unfortunately, it is often difficult or impossible to include long-term follow-up testing in research projects. There is also, of course, the issue of what happens to the learners during the intervening period, even if long-term testing is possible.

White, Spada, Lightbown and Ranta's (1993) study, dealing with question formation in English, produced more positive long-term results. The group that received direct instruction made significantly greater gains than the comparison group on immediate posttests that required them to perform written tasks. The difference between the groups was maintained on delayed posttests (six weeks after the instructional period) and follow-up tests (six months after instruction). Improvements were also found in the oral performance of the students who received instruction.

The lack of long-term effects shown in some studies (e.g. Harley [1989] and White [1991]) led Lightbown and Spada (1993) to give only qualified
approval to the concept of form-focused instruction:

The overall results of the experimental studies in the intensive ESL and French immersion programs provide partial support for the hypothesis that enhanced input [Lightbown and Spada's emphasis] or form-focused instruction and corrective feedback within communicative second language programs can improve the learners' use of particular grammatical features. The results also show, however, that the effects of instruction are not always long lasting (p. 102).

Lightbown and Spada (1993) suggested several reasons for the fact that some form-focused studies obtain long-term effects, while others do not. These include differences in the experimental teaching materials and methodology, the frequency of use of the linguistic structures in classroom input after the treatment period, the ways that different linguistic features may respond to instruction, and the age of the learners. On the whole, however, Lightbown and Spada (1993) gave a positive appraisal of form-focused instruction:

Classroom data from a number of studies offer support for the view that form-focused instruction and corrective feedback provided within the context of a communicative program are more effective in promoting second language learning than programs which are limited to an exclusive emphasis on accuracy on the one hand or an exclusive emphasis on fluency on the other. Thus, we would argue that second language teachers can (and should) provide guided, form-based instruction and correction in specific circumstances (p. 105).

Among the 'specific circumstances' referred to in the preceding quotation, Lightbown and Spada (1993) included persistent errors which learners appear not to notice, and errors that are made by the majority of learners who share the
same first language. Both of these criteria apply to the learning of grammatical gender in French by English-speaking students.
Chapter 3

Exploratory Research

3.1. Context and participants

The preliminary exploratory investigation was designed to provide diagnostic information about the use of gender by French immersion students at the high school level. The participants were six students (four females and two males) in Grade 11. They had been enrolled in a 50/50 (half day French, half day English) immersion program in Grades 1 to 8, and had continued with immersion in high school. In Grades 9 and 10, four of their eight subjects (français, histoire/géographie, mathématiques, clavigraphie) were taught in French. By Grade 11, however, they were taking only one course (français) in French. A small corpus of written material was collected from the students. The students were chosen on the basis of their French marks in Grades 9 and 10. Two 'A' students, whose marks had been consistently over 80% for the two years, were chosen. Similarly, two 'B's' (70-79%) and two 'C's' (60-69%) were selected. The 4:2 female to male ratio approximately reflected the gender split in the class. Each student submitted two samples of writing. The first, written within an hour, was a test essay question requiring the students to compare a
film to the novel from which it was loosely adapted. Both the novel and the film had been studied and discussed in class. The second writing sample was a book report that the students were given three weeks to prepare. The writing samples were not initially collected for the purpose of analyzing them for gender use, but rather as part of a study comparing written output under two different sets of conditions. The original focus of the study was time, to see if the students' written work contained fewer errors when they had a great deal of time to revise and correct it than when they were writing under pressure. During the analysis of the data, it became apparent that the writing samples might yield some interesting information with respect to gender.

3.2. Research questions

During the analysis of the writing samples, the following questions were considered:

1. How frequently do students make gender errors; i.e., what is the percentage of correct gender marking through the agreement of other sentence elements such as articles and adjectives? Does this percentage vary widely from one student to the next, or is it fairly consistent?

2. Does time play a role? Do learners make as many errors when they have plenty of time to check over their work as when they are writing
under pressure?

3. What types of gender errors do students make? For instance, are adjectives more likely to be used incorrectly than articles?

4. Do patterns of usage emerge from the data? For example, is the masculine more likely to be overused than the feminine, or vice versa?

5. Does natural gender play a role? Are students less likely to make a mistake such as "une père" than "une crayon"?

6. Are learners consistent in their choices? If a noun is used several times in a sample, is it always assigned the same gender (either correctly or incorrectly) or does usage fluctuate?

7. Are mistakes more likely with unfamiliar than with familiar words?

8. Are mistakes less likely with words whose endings (such as -tion and -isme ) are highly predictive of gender, according to statistics published by Tucker, Lambert and Rigault (1977) and Luce (1979)? In other words, do the students demonstrate that they have acquired some of the 'rules' that enable francophones to predict the gender of most nouns?

3.3. Frequency of errors

Twenty-six pages of writing were analyzed. Twelve were 'prepared' writing -- a book report for which the students were given three weeks to read a
novel of their choice and write a review. Fourteen pages were produced in a time-limited testing situation lasting about one hour. A total of 261 gender errors were noted, for an average of ten per page (approximately 250 words were written on each page). Of more interest than the total number of errors is the percentage of correct and incorrect choices. In order to get a rough estimate of this, six pages of writing (three prepared and three test) by three students (one 'A', one 'B' and one 'C') were analyzed. The total number of gender-marked nouns was counted (une école and la grande voiture are marked; l'école and leur jeune fille are not), the number of errors was noted, and the percentage was calculated. There was some question of how to treat items for which double or multiple marking of gender was required, such as un livre intéressant or la petite fille est allée. It was decided to score the marking of each noun simply as right or wrong. Thus  *une livre intéressante* and *un livre intéressante* were both scored as one mistake. One might argue that the second example could be interpreted as correct identification of the gender of livre but carelessness or lack of knowledge with respect to adjective agreement. However, if one's fundamental definition of grammatical gender is based on the notion of agreement or concord (Hockett 1958; Fodor 1959; Gleason 1961; Corbett 1991), then 'knowing' that livre is masculine is not good enough. This knowledge must be applied every time that the agreement rules of the language require it. *Une livre intéressante*, while clearly a gender error, does at least follow the rule of agreement between articles and adjectives.

In the six pages, there were 183 instances of required gender marking.
An error was found in 65 of these (35.5%). It should be noted that many of these were repeated errors; for example, one student used "la film" three times in one page. For the 'A', 'B' and 'C' students, the error percentages were 16.3%, 40% and 48.8% respectively, indicating a wide range of control over or attention to gender agreements in written work.

3.4. Prepared writing vs. writing under pressure

In the assignment that the students had three weeks to prepare, the average number of gender errors was 8.25 per page. For the test situation, there was a higher count of 11.6 errors per page. (Although dictionaries were permitted during the test writing, the time constraint limited their use.) This may indicate a somewhat greater tendency on the part of at least some of the students to look up the gender of nouns when they had time to do so. However, fewer errors in one situation than the other cannot be assumed to be an indication of better control of gender in that situation. For example, the student who avoids an error in a 'prepared' text by looking up the gender of the noun *communisme* in the dictionary does not know what francophones know automatically; i.e., that any noun ending in -isme is masculine. There is also no guarantee that the student will generalize the gender of a single item to other nouns with the same ending, or even that the gender of the single item will be remembered the next time that he/she wants to use the word.
3.5. Types of gender errors

Errors in the corpus included almost every possible part of speech that can be marked for gender: definite and indefinite articles, possessive and demonstrative adjectives, descriptive adjectives, past participles, and personal pronouns. No demonstrative or possessive pronouns (e.g., *celui, la mienne*) were found in the samples. Due to the high frequency of articles, it is not surprising that they accounted for 70.5% of the errors. Descriptive adjectives were next, at 23%. Thus, if these students were to avoid errors involving articles and descriptive adjectives, they would eliminate about 93% of gender errors from their written work.

3.6. Patterns of gender errors

Of the 261 errors in the corpus, 99 (38%) involved incorrect attribution of the masculine (e.g. *un race*) and 162 (62%) incorrect use of the feminine (e.g. *la directeur*). However, the numbers are badly skewed by the fact that the words *livre* and *film* are identified as feminine on 72 occasions. (Both writing assignments involved the discussion of a book, and one required the students to compare a novel and a film.) Three students consistently assigned feminine gender to these two words. If all but 6 (to account for the first incorrect use of each word by each of the three students) of the 72 incorrect examples of *livre* and *film* are excluded from the analysis, we are left with 195 errors, of which 99
(50.8%) involve incorrect masculine assignment, and 96 (49.2%) attribute feminine gender to masculine nouns.

This remarkably even masculine/feminine split in the overall results does not apply to individual students, however. For example, 12 out of 12 article errors by one student substitute the masculine for the feminine (e.g. "un histoire, le direction"). Another student's article errors involve substituting the feminine for the masculine (e.g. "une groupe, la seul jour"), except for a few instances where the contracted masculine forms du and au are used with feminine nouns ("du dynamite, au societe"). Many students do not indicate a strong preference for either the masculine or the feminine, since they have an almost equal split between masculine and feminine errors.

An analysis of the subsample of adjective errors reveals an almost even split (31 masculine incorrect / 29 feminine incorrect). This contrasts with Stevens' (1984) finding that students with up to six years' experience in French immersion almost always used the masculine form of the adjective. The students in the present study were just about as likely to write "il est jalouse or "des plus grandes changements" as "certains similarites" or "de meilleurs amis.

3.7. The role of natural gender

Examples from the corpus indicate that natural gender clues do not
always prevent students from making mistakes. The following were all found in
the data: 'une homme, "la directeur (referring to a man), "Mario est plus
jalouse, "il est spéciale, "la père, "son frère Carl ... la favourite, "un petite fille,
"une pêcheur.

It is also possible that natural gender can contribute to errors in
grammatical gender when a conflict occurs between the two. For example,
female characters in a novel are referred to by one student as "la personnage
and "la bébé.

3.8. Consistency of errors

Patterns vary from student to student when the same noun is used
several times in a writing sample. A student who uses il to replace l'école also
writes "un autre école. However, on another occasion, he correctly uses la
nouvelle école. The same student writes "des choses importants, "un des
 choses, "un autre chose, "un des choses importants, and "les différents
 choses, but correctly uses une chose once.3 Another student consistently (5
times) assigns feminine gender to groupe. One student uses la famille three
times, "le familie once, and "du familie twice.

3Roy Lyster (personal communication) has pointed out that the discussion of chose is
slightly problematic because the word can be either feminine or masculine (un chose being
roughly synonymous with truc or machin, and equivalent to 'thingamajig' in English). However,
the masculine form of the word is far less common than the feminine, and the student obviously
did not intend the meaning of the masculine form in the composition.
The most frequently used nouns in the corpus, *film* and *livre*, provide an interesting case. Three students show an overwhelming preference for the feminine for these words. One student uses *la livre* 17 times, *le livre* three times, and *du livre* once. *Film* is identified as feminine 18 times, masculine twice. The second student uses *la filme* 16 times, ‘correcting’ the article from *le* on one occasion. However, three unchanged examples of *le filme* also occur. This student never assigns masculine gender to *livre*. The third student uses *la livre* 10 times, *un livre* once, and *du livre* twice. *La film* is used 10 times, and *du film* three times. There appears to be a pattern to this student’s choices. The feminine definite article is consistently chosen, except in combination with *de*, when the contraction *du* is used. It is interesting that all of this student’s definite article errors involve substituting *la* for *le* (*la jour, *la monde, *la camp, *la même age, *la directeur). Except for one case (*une animal *), indefinite article errors are in the opposite direction (*un fille, *un fête, *un salle, *un réunion, *un relation, *un idée, *un raison). In fact, this student uses *une* only once, and incorrectly, in four pages of writing. However, *le* is correctly used on a few occasions (*le renard, le directeur, le père, le meilleur ami*), all of which, though perhaps coincidentally, involve natural gender.

3.9. **Gender errors with familiar and unfamiliar words**

Are mistakes more likely with unfamiliar than with familiar words? This is a difficult question to answer, since it is not possible to know from the data
exactly which words were well known to any individual. Furthermore, what does it mean to be familiar with a word? How many times does one need to hear a word, encounter it in print, or use it in speech or writing in order to be deemed familiar with it?

In spite of these questions, an overview of the entire corpus gives the distinct impression that the majority of gender errors were made with words that the students had encountered and used many times before. Words such as *livre, fille, père, restaurant, salle, animal* and *chose* were undoubtedly familiar to the students in elementary school. Some nouns, such as *livre* and *chose*, must have been encountered on an almost daily basis. The fact that errors were made with nouns such as these seems to indicate a lack of attention to gender, rather than a lack of familiarity with vocabulary.

Other words, while perhaps not dating back to the students' elementary school experience, should have been familiar to them. The noun *guerre* was used throughout their Grade 10 history course. *Film* was used in the Grade 11 course during the discussion of a film which the students compared to a novel, and appeared in print many times in questions that were handed out or written on the blackboard.

This issue is further complicated by the fact that correct gender assignment to unfamiliar nouns does not necessarily indicate that the students 'know' the gender of these words. If a word is used only once in a writing sample, the student has about a 50% chance of correctly guessing the gender.
It is also possible to check the gender of an unfamiliar noun in the dictionary, but not to remember it for future use.

3.10. **Errors with nouns whose endings are highly predictive of gender**

Tucker, Lambert and Rigault (1977) and Luce (1979) have demonstrated that many noun endings are highly predictive of gender. For example, Luce claims 100% certainty that words ending in -sme are masculine, and almost 100% certainty that words ending in -tion are feminine (exceptions are *le bastion* and *le spallation*).

Following are just a few of the mistakes found in the writing samples involving nouns with highly predictable endings: *le direction, le location, un collection, un réunion, un relation, son réputation, son réaction, la commencement, la changement, une restaurant, une délinquant, sa dessin, un influence, une animal, la billet*. Obviously, the immersion students who wrote these words have not developed the automatic control of gender assignment, even for the most predictable noun endings, that native speakers of French have.
3.11. Conclusions

The findings of the exploratory study are generally consistent with those of previous research (Tucker, Lambert and Rigault 1977; Harley 1979; Stevens 1984); i.e., that anglophone learners of French, even after many years of study, demonstrate little systematic control over grammatical gender. An awareness of consistent relationships between noun endings and gender does not appear to be integrated into the developing linguistic system of the learners. Factors such as natural gender, and the familiarity or frequency of words, do not appear to be of much help either. The high rate of errors (about 35%) indicates that the students have neither unconscious control (automaticity) over gender, nor conscious awareness of its importance, since most of them do not often verify the gender of nouns even when they have unlimited opportunity to do so, as in prepared compositions. Furthermore, there appear to be no patterns of use that are consistent for all the students, such as a preference for the masculine or the feminine. Individuals vary widely in terms of frequency, consistency and type of error. This differs from the over-generalization of the masculine typically found in studies with younger learners (Tarone, Frauenfelder and Selinker 1976; Harley 1979; Stevens 1984; Taylor-Browne 1984), and may indicate that the older learners in the study at least have some basic idea that distinctions need to be made, haphazard as their implementation of this knowledge might be. It should also be pointed out that the exploratory study dealt entirely with written samples of language, and that the patterns of the six participating students'
spoken production might be different with respect to gender control.

3.12. Implications for further research

The exploratory study was obviously very limited in nature. Only six students were involved, and only written output was analyzed. The study did, however, provide a great deal of information that was of value for designing a larger study intended not only to observe immersion students' behaviour with respect to grammatical gender in French, but also to assist them to improve their proficiency.

First, there is obviously a need for consciousness raising. The high rate of errors (35%), the frequency of errors with very commonly-used nouns (such as livre), and the often fluctuating assignation of gender to the same word, all indicate that gender is not a feature to which the students pay much attention. Psycholinguistic explanations for this lack of awareness have been summarized in Section 2.7.

The fact that errors are made with both familiar and unfamiliar words, and that students appear largely unaware of the word-ending clues to gender, suggest that a focus on noun endings (both phonological and written) should be central to any attempt to improve learners' proficiency.

Since the vast majority of errors (over 93%) involved articles and descriptive adjectives, it seems logical to focus on these features in any initial
program of intervention. Possessive adjectives are excluded not only because of their relative infrequency in the writing samples analyzed in the exploratory research, but also because they differ psycholinguistically from articles and descriptive adjectives for anglophone learners of French. This is because English possessives depend on the sex of the possessor, while French possessives depend on the grammatical gender of the modified noun. English distinguishes between *his book* and *her book*, while French uses *son livre* for both, since *livre* is masculine. Another difficulty with possessive adjectives occurs because of forms such as *mon amie*, where the normally masculine form must be used before a feminine noun beginning with a vowel. Since possessives by themselves could be the focus for a whole study, they have been omitted from the present research.

Tarone, Frauenfelder and Selinker (1976), Harley (1979), Stevens (1984) and Taylor-Browne (1984) found overgeneralization of masculine forms. However, the present exploratory study showed no overall preference for one gender or the other (although there were some individual learner preferences). The almost even split between masculine and feminine errors suggests that equal attention needs to be paid to both masculine and feminine endings in any instructional design.
Chapter 4

Methodology

4.1. Overview of the design of the study

This chapter describes the participants in the study (Section 4.2), the tests and questionnaires used (Sections 4.3 and 4.4), procedures (Section 4.5) and the experimental classroom treatment (Section 4.6).

The study set out to investigate the effect of form-focused instruction on control of grammatical gender by French immersion students in Grade 11. Pretests, immediate posttests and delayed posttests were administered in one experimental class and two comparison classes. Treatment materials which were designed specifically for the study were implemented in the experimental class over two months following the pretests. Students in the experimental class completed questionnaires designed to assess their reactions to the treatment materials and activities. Some of the teacher/researcher's observations of the experimental class during the treatment period were recorded and reported. The test results were subjected to statistical analysis in order to compare the growth over time of the three classes.
4.2. Participants

The participants in the study were 62 Grade 11 French immersion (FI) students in three classes from two schools in the same school board in the Greater Metropolitan Toronto area. The experimental class (E) consisted of 28 students and was taught by the investigator, a native anglophone who is fluent in French, with 19 years of teaching experience, ten of which were in French immersion. There were two comparison classes in the study. One of these (C2), with 17 students, was also taught by the investigator. The other comparison class (C1), also with 17 students, was taught by a native francophone, fluent in English, with 16 years of teaching experience, 10 of which were in French immersion. The experimental class and one comparison class (C2) were in the same school. The second school (class C1) was chosen for two reasons: the similarity of its French immersion program to that of the other school, and the comparability of its socio-economic intake area to that of the first school. The willingness and enthusiasm of the other teacher, when approached about participating in the study, was also an important factor. Also, it was impossible for timetabling reasons that will be explained later in this section to schedule the two classes in the same school (C2 and E) in the same semester. However, classes C1 and E did run simultaneously, though in different schools.

Before the classes were selected, permission was requested and received from the Research Advisory Committee of the school board to conduct
the study. Letters were sent home explaining the general nature of the research, but not the specific language focus, and requesting parental permission for students to participate in the language testing, including tape-recording of some students (see Appendix A). There was a 100% positive return rate for the forms. Two parents requested further clarification before returning the forms, and this was provided by telephone. Two parents added comments of support and encouragement to the forms.

The students were asked to provide information about their linguistic and educational backgrounds, including languages other than English used at home, and a detailed history of their French studies if not always with the participating board of education. Students who indicated that French was spoken at home were interviewed briefly to obtain a clearer picture of their contact with French outside the school setting. One student in Class C2 came from a francophone background and spoke mainly French at home. A student in Class E also was of francophone background, and occasionally spoke French at home with one parent, though English dominated. These two students participated in the testing, but their results were omitted from the data analysis. Thus the experimental group consisted of 27 students, while the comparison group was made up of 33 students in two classes of 17 and 16. Only one student in the three classes reported speaking any language other than English at home. She was in Class E and mentioned the occasional use of Chinese.
The majority of students had spent their entire educational background with the participating board. They began French in Grade 1 in a 50/50 early immersion program (half day French, half day English). The 50/50 split continued through elementary school to the end of Grade 8. In high school, five of the eight subjects taken in Grade 9 (French, geography, mathematics, keyboarding and art) and three in Grade 10 (French, history and mathematics) were conducted in French. In Grade 11 the students took only one immersion credit (French language and literature). The two participating schools were semestered, so each credit was taken during half the school year, either from September to January, or from February to June. Classes were 76 minutes long.

Several students in each class had begun their immersion studies with a different school board, either in Ontario or another Canadian province. In all cases, their accumulated hours of instruction in FI were greater than the total for students who had completed all their studies with the participating board. Most school boards begin with most or all of the instructional time in French, either in a half-day kindergarten program or for the entire day in Grade 1, then gradually introduce English until the two languages are about evenly split in the upper years of elementary school. Separate statistics were not kept for students with different totals of accumulated hours in FI.

The number of FI students enrolled in the researcher's school yields only one or two classes a year in any grade. When there are two sections, they are
always placed one in each semester, in order to balance the timetable. Thus it was impossible for the researcher to teach a comparison and an experimental class during the same time period. Class C2 ran from February to June 1995, and Class E from September 1995 to January 1996. The comparison class in the second school (C1) was taught in the same semester as the experimental (E) group.

4.3. Testing instruments and scoring methods

In this section, a description and rationale will be provided for each of the five evaluation instruments, and scoring procedures will be explained. All tests except the oral production task were developed by the researcher, with input from several members of the Modern Language Centre (MLC) at the Ontario Institute for Studies in Education (OISE). The oral interview was adapted from materials already available at the MLC.

The first step involved the development and piloting of the testing materials. Three tests -- a listening test (LT), a written endings test (WET), and an agreement test (AT) -- were developed (see Appendices B, C and D), then piloted in December 1994 in a Grade 11 French immersion class of 20 students. The written endings and agreement tests both had two versions (Forms A and B), thus one purpose of the pilot testing was to see if both forms yielded similar results. The second purpose of the pilot testing was to ensure that the pretest
would not produce a 'ceiling effect'. By this we mean that a substantial proportion of students score at or near the top of the measurement scales on the pretest, thus leaving little or no room for gains on the posttests. The listening and written ending tests were also administered to a group of 22 exchange students who were visiting from France for three weeks. (They were not given the agreement test, since it included translation from English to French.) The French students were 14 and 15 years old; i.e. a year or two younger than the Grade 11 immersion students. Two other testing instruments, a writing assignment (WA) and an oral production task (OPT), were developed. For the writing assignment, students were given a topic and asked to write a composition in class. The oral materials were adapted from testing instruments developed by researchers at the Modern Language Centre at OISE for a gender study involving Grade 2 immersion students.

4.3.1. Listening test (LT)

Tucker, Lambert and Rigault (1977) and Karmiloff-Smith (1979) stressed the importance of phonological word endings to native French-speakers' control of grammatical gender. According to Karmiloff-Smith, before francophones learn to read and write, they construct "a very powerful, implicit system of phonological rules" (p. 167) that enable them to correctly assign gender to the vast majority of new nouns that they encounter. The listening test was designed to measure immersion students' awareness of the predictability of noun gender
according to word-ending phonological clues.

Twelve word endings, six characteristically masculine and six characteristically feminine according to Tucker, Lambert and Rigault (1977) and Luce (1979) were chosen. The masculine endings were -ment, -age, -in, -isme, -eau/-aut and -at/-ap. The feminine endings were -tion, -ette, -ure, -esse, -ance/-ence and -erie.

For each of the 12 endings, three nouns were chosen from Le Petit Robert dictionary. The criterion was that they should be words with which immersion students would likely be unfamiliar. After the listening test was administered to the pilot Grade 11 class, the students were asked if they were familiar with, or could explain the meaning of, any of the words. Two or three students expressed some familiarity with indolence, maladresse, quincaillerie, ménagerie and existentialisme, but on the whole, the 36 nouns were almost totally unknown to the 20 students. The 36 words, grouped according to ending and to gender, are presented in Table 4.1.

A cassette recording was made of the 36 words presented in mixed order (figement, indolence, maladresse, garniture, fifrelin, etc.) but with no ending repeated in any five consecutive words. Each word was preceded by a number corresponding to the questions on the student's answer sheet (see Appendix B). Each word was pronounced twice, with a pause of about two seconds between utterances ("Un ... figement ... figement ... deux ... indolence ... indolence ... trois ... maladresse ... maladresse ... etc."). Students were
instructed to circle UN or UNE on their answer sheet for each word, according to Table 4.1:

Table 4.1: Endings and nouns used in the listening test (LT)

<table>
<thead>
<tr>
<th>Masculine</th>
<th>Feminine</th>
</tr>
</thead>
<tbody>
<tr>
<td>-ment</td>
<td>-ance/-ence</td>
</tr>
<tr>
<td>figement, soubassement, reverdissement</td>
<td>indolence, défaillance, médiasance</td>
</tr>
<tr>
<td>-in</td>
<td>-esse</td>
</tr>
<tr>
<td>fifrelin, pulvérin, paladin</td>
<td>maladresse, finesse, kermesse</td>
</tr>
<tr>
<td>-age</td>
<td>-tion</td>
</tr>
<tr>
<td>sulfitage, garnissage, finage</td>
<td>spoliation, étatisation, ossification</td>
</tr>
<tr>
<td>-at/-ap</td>
<td>-ette</td>
</tr>
<tr>
<td>potentat, scélérat, sparadrap</td>
<td>oeillette, estafette, gansette</td>
</tr>
<tr>
<td>-eau/-aut</td>
<td>-ure</td>
</tr>
<tr>
<td>panicaut, meneau, carneau</td>
<td>éclaboussure, armature, garniture</td>
</tr>
<tr>
<td>-isme</td>
<td>-erie</td>
</tr>
<tr>
<td>tabagisme, existentialisme, helvétisme</td>
<td>quincaillerie, révasserie, ménagerie</td>
</tr>
</tbody>
</table>

whether they thought it was masculine or feminine. The process and the test were identical for the pretest, immediate posttest and delayed posttest in all three classes.

Each student’s test was given a raw score mark out of 36, which was then
converted to a percentage for the data analysis and statistical comparisons.

The mean score for the Grade 11 pilot class on the listening test was 66.8%. The mean score for the 22 exchange students from France was 98%, with no consistent errors. Based on these results, it was decided that modification of the test was not necessary, since its validity was confirmed by the native speaker responses and there was plenty of room for gains by the immersion students.

4.3.2. Written endings test (WET)

The written endings test, as the label implies, focuses on the written word endings, such as -tion and -isme, that enable francophones to predict the gender of the vast majority of nouns. Tucker, Lambert and Rigault (1977) used written as well as oral clues in some of their studies of francophones' control of grammatical gender (the youngest of their subjects were in Grade 3). They found that graphic-oral presentations of unfamiliar nouns tended to lead to even greater accuracy of gender prediction by native speakers than oral clues alone. Luce (1979), in an article designed to assist second language learners to remember and predict the gender of French nouns, provided reliability ratios for written endings, such as -ie, 100:1 feminine, and -age/-ège/-oge, 50:1 masculine.

For the written endings test, six masculine and six feminine word endings
were chosen. The criterion was that they had to have at least 20:1 (95%) reliability as predictors of gender according to Luce (1979). The masculine endings were -at, -eau, -in, -isme, -ment, -age. The feminine endings were -esse, -ette, -ance, -tion, -ure, -ie. These were basically the same as the phonological endings chosen for the listening test, but with only one spelling allowed for each ending. In the listening test, for example, both panicaut and carneau were included for the phonological ending /o/, since students did not see the words. In the written endings test, only the -eau spelling was used. Also, the ending -erie was shortened to -ie for the written endings test (although three of the words still ended in -erie). Phonologically, /il/ includes not only -ie but typically masculine orthographies such as -is, -it and -il. However, the -ie spelling is about 98% reliable as a predictor of feminine gender (Tucker, Lambert and Rigault 1977, p. 112). For each of the twelve endings, six pseudonouns were invented by the researcher. Given that rare nouns had already been used for the listening test, it might have been difficult to find six more for each ending and to be sure that they were unknown to all students. It was also assumed that students would be more likely to recognize real words on the written endings test than on the listening test because they could actually see the words. The main criterion used when inventing the words was that they must adhere to the phonological rules of French (i.e. no inadmissible sequences of phonemes, such as initial ng-). All the pseudowords for each ending began with a different phoneme. This was done with the purpose of realistically reflecting the variety of the French lexicon. Both Le Petit Robert
and *Le Petit Larousse illustré* dictionaries were checked to ensure that none of the invented words actually existed. The 72 pseudonouns, grouped according to ending and to gender, are presented in Table 4.2:

**Table 4.2: Endings and pseudonouns used in the written endings test (WET)**

<table>
<thead>
<tr>
<th>Masculine</th>
<th>Feminine</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>-at</strong></td>
<td><strong>-esse</strong></td>
</tr>
<tr>
<td>éventat, serjat, placat, célibat, tabarat, névrat</td>
<td>falesse, paramesse, gouresse, doutesse, tesse, vivesse</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>-in</strong></th>
<th><strong>-ette</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>ruvin, punin, drapin, frottin, glacin, balin</td>
<td>garagette, effarette, caravette, magette, soustrette, passerette</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>-eau</strong></th>
<th><strong>-ance</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>riveau, clapiteau, gradeau, pousseau, mameau, lanteau</td>
<td>factance, brumance, galance, poutance, ermitance, mangeance</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>-isme</strong></th>
<th><strong>-tion</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>topisme, crinisme, purilisme, luminisme, minisme, dansisme</td>
<td>provation, coupation, floration, tamisation, armation, couration</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>-ment</strong></th>
<th><strong>-ure</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>salument, girement, résument, bossement, rougement, menacement</td>
<td>paladure, chômure, branchure, navure, filoture, tramure</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>-age</strong></th>
<th><strong>-ie</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>brisage, pendage, gouttage, navage, nordage, florage</td>
<td>boulerie, drapologie, membrie, lapaderie, soupinerie, trapézie</td>
</tr>
</tbody>
</table>
Two versions of the written endings test, Forms A and B, were produced (see Appendix C). In each version, three words were included for each ending, giving a total of 36 items, 18 masculine and 18 feminine. As with the listening test, the words were presented in mixed order, but with no ending occurring twice in any five consecutive words. UN and UNE were printed before each word. The students were instructed to circle one or the other to indicate what they thought the gender of each word would be.

Scoring was the same as for the listening tests. Each individual test was assigned a raw score mark out of 36, which was then converted to a percentage for the data analysis and statistical comparisons.

Ten students in the pilot class received Form A of the written endings test, and ten students received Form B. The mean score was 73.7% (SD = 13.79) for Form A and 74.4% (SD = 13.23) for Form B. The two means were compared using t-tests for independent samples, and the difference was found to be not significant (p = .89). Since the two versions of the tests appeared to be equivalent, and since the results allowed room for gains on subsequent testings, it was decided not to revise the tests. The mean scores for the 22 French exchange students were 91.4% for Form A and 91.1% for Form B. As with the listening test native speaker scores, there were no consistent errors. When combined with their listening test score (98%), the average was about 95%, which coincides with Tucker, Lambert and Rigault's (1977) finding that about 95% of the time, French speakers will correctly predict the gender of an
unknown noun. The fact that the exchange students performed somewhat better on the listening test than on the written endings test might at first appear to contradict Tucker, Lambert and Rigault's (1977) finding that graphic-oral presentations of unfamiliar nouns tended to lead to even greater accuracy of gender prediction by native speakers than oral cues alone. However, it should be pointed out that in the present study, the items on the listening test were all real words, some of which might have been known to some of the native speakers, whereas the items on the written endings test were all pseudowords.

4.3.3. Agreement test (AT)

In Section 2.1 it was noted that concord, or agreement, is considered by modern linguists to be central to the definition of grammatical gender. Corbett (1991) stated that "the determining criterion of gender is agreement" (p. 4). In other words, there is no known language that categorizes nouns according to gender without having some other parts of speech, such as adjectives and articles, change form according to the gender of the nouns with which they are associated. In French, articles and adjectives are the most frequently used parts of speech that are inflected for gender.

The agreement test was designed mainly to assess the students' ability to choose the correct masculine or feminine form of an adjective to modify a noun. As with the written endings test, two versions, Forms A and B (see Appendix D), were produced. Each form contained 20 questions. For each
question, students were asked to identify the gender of a given French noun, then to translate an English adjective into French to describe the noun. For example,

\[
\text{le / la } \underline{\text{______________}} \text{ (new) garage}
\]

\[
\text{un / une } \underline{\text{______________}} \text{ (blue)}
\]

To answer correctly, students would have to circle \textit{le} in the first example and write \textit{nouveau} in the blank. In the second example, they would have to circle \textit{une} and write \textit{bleue}. Twenty different nouns, 10 feminine and 10 masculine, were chosen for each version of the test, but the same 20 adjectives were used for both versions. If the masculine form of an adjective, such as \textit{nouveau}, was required on Form A, then the feminine form (\textit{nouvelle}) was required on Form B.

Unlike the listening and word endings tests, which contained unfamiliar words and pseudowords, the agreement test was deliberately made up of words which immersion students could be expected to know. In the pilot Grade 11 group, individual students asked for the meanings of \textit{crapaud, fourrure, bijou, couronne} and \textit{église}, but most knew all or nearly all of the words. Some students were unsure of how to translate \textit{soft} and \textit{false}, but the majority had no difficulty coming up with French equivalents of the English adjectives, so no revisions were made to the tests.

The scoring procedure for the agreement test was fairly complicated. Raw scores out of 20 were converted to percentages. Separate tallies were kept for articles and adjectives. For example, the mean score for the pilot group
on the pretest was 79.87% correct for articles and 72.34% correct for adjectives. Separate results were also kept for both article and adjective correct (70.9%), neither article nor adjective correct (12.5%), and disagreement between article and adjective (14.69%). Note that if both the article and the adjective are incorrect, they agree with each other. For example, "une nouvelle garage" is incorrect because the student attributes the wrong gender to garage, but the article and adjective, which are both feminine, agree with each other. "Un nouvel garage" and "une nouveau garage" are examples of disagreement between the article and adjective. It was the totally correct scores (article and adjective both correct), demonstrating knowledge of both gender and agreement, that were mainly used for statistical comparisons between the three classes on the agreement test. However, statistical comparisons were also conducted for articles and adjectives separately.

The mean scores for the Grade 11 immersion pilot class were 72.3% (SD = 19.75) for Form A and 69.5% (SD = 21.0) for Form B. Paired t-tests indicated that this difference was not significant (p = .746).

4.3.4. Writing assignment (WA)

A day or two after the administration of the listening, word ending and agreement pretests, the participants wrote a composition in class. The subject was Un lieu important dans ma vie ('An important place in my life'). The students were told to write about a place of importance to them, such as their
cottage or their house. Two months later (after the treatment period for the experimental class), the students were assigned a second composition in class, which required them to compare two characters that they had encountered in their literature studies during the course (from novels, plays, films, etc.). The suggested length for the compositions was about 300 words. In neither case were the students told to pay particular attention to gender. However, they were instructed to check over their work for grammar and spelling. They were also asked to use many descriptive words in their compositions. They were allowed to consult dictionaries, grammar notes, etc., during the writing session, but actually had limited time to do so, since the amount of time to complete the assignment was just over an hour. In order that the two writing sessions would be taken seriously by the students, they were told that a grade, which would form part of their written term mark, would be assigned to each composition. This mark, however, was not tied to their control over grammatical gender and adjective agreement, but was based partly on the content of their work, and partly on the overall correctness of their French.

For research purposes, each composition was given a percentage score that represented the total number of correct gender agreements divided by the total number of obligatory gender agreements for the two parts of speech, articles and descriptive adjectives, on which the treatment period focused (see Section 3.12 for rationale). Possessives, demonstratives, personal pronouns, and past participles of être verbs (elle est allée) were not included in the scoring. Separate sub-tallies were kept for articles and descriptive adjectives,
so that it would be possible to determine which of these features the students seemed to have better control over. Thus for each set of compositions there is a percentage correct score for articles, a percentage correct score for adjectives, and a percentage correct total score. It should be noted that the total score does not represent the average of the other two, since students used more articles than adjectives in their writing, and this weighting is reflected in the total score. For example, a student in the C2 class received a score of 86.36% for articles and 42.1% for adjectives on the second composition. His overall score was 68.18%, which is about four percentage points higher than an average of the two separate scores. It was this final overall score which served as the main basis of comparison between the two writing sessions and the three classes.

4.3.5. Oral production task (OPT)

Oral testing materials, based on Karmiloff-Smith (1979), had been developed at the Modern Language Centre at OISE as part of a classroom experiment in form-focused instruction designed to examine the effect of teaching grammatical gender to Grade 2 French immersion students (see Harley 1994b, in press). Despite the nine-year grade difference between participants in Harley’s study and the present study, the materials adapted well for use with the older students.

The oral test consisted of two parts. In the first part, the students were
shown a series of 13 pictures and asked to identify the object in each picture. Note that there were two identical representations of each object (see Appendix E for an example). The endings of the words made their gender predictable (couronne, matelas, biberon, tondeuse, marmite, sabot, roulotte, pansement, brouette, sauterelle, chamois, moulin, soupière). For each object, the interviewer pointed to one of the two identical pictures and asked, “Qu’est-ce que c’est?” and if the student responded correctly, the reply would be, “C’est une couronne, c’est un matelas, etc.” If the student did not know the word for the object, the interviewer would say, “Voici deux couronnes” (hence the double images), then point to one of the pictures and ask, “Qu’est-ce que c’est?”, prompting the repetition of the noun in the singular with “c’est un” or “c’est une” preceding it.

The second part of the oral production task was more open-ended. The students were shown a drawing of a country scene (see Appendix F) and simply asked to describe it. The objective was to elicit as many nouns as possible in the singular, each preceded by a definite or indefinite article. Students were told not to worry about objects that they could not identify in French. A totally correct response would be something like: “Je vois une vache, un mouton et un cheval dans un champ. Il y a un homme sur une échelle qui répare le toit. Il y a un homme qui attrape un poisson. Il y a une petite fille qui prépare une party avec un cadeau et un gâteau. Sur la table il y a une assiette, une fourchette, un couteau et une cuiller. Sur un petit lac il y a un bateau et un radeau. Je vois aussi une bicyclette, un dindon, un arbre avec des pommes, un
lapin, une voiture, et une femme qui travaille dans le jardin.” Students were not expected to identify every object in the picture, but were encouraged to continue talking if they stopped before about 20 items were mentioned.

For the first part of the oral production task, involving the 13 pictures, students were assigned a raw score out of 13, which was then converted to a percentage. The students had to be clear in their pronunciation of un or une in order to receive a point for each item. Harley (1993) has pointed out that “some immersion students appear eventually to settle for a strategy of producing a vowel sound that lies somewhere between le and la and that can disguise their lack of knowledge in oral contexts” (p. 253). A similar ambiguity for un and une was noted for some students for some (but never all) of the items. Whether the lack of clarity was due to a strategy to disguise their lack of knowledge of the gender of the objects, or simply to momentary inattention to pronunciation, is not known. In any case, unclear answers were scored the same as obviously incorrect ones. This decision was based on the principle that the grammatical gender of a noun must not just be ‘known’, but must be clearly expressed whenever there is obligatory agreement (see Sections 2.1 and 3.3). Synonyms were accepted by the scorer if the gender was correctly identified, so a student who responded “C’est une bouteille” instead of “C’est un biberon” would receive the mark.

For the picture description part of the oral production task, which was more open-ended, students’ raw scores differed not only according to the
number of correct gender identifications, but also according to the possible total of correct answers. This was because they did not all identify the same number of objects in the picture. (They were all, however, encouraged to keep talking until they mentioned at least 20 items.) Thus, for example, one student might have a score of 13 out of 21, while another received 18 out of 26. For comparison purposes, each student's raw score was converted to a percentage.

To obtain an overall score on the oral production task, the scores for the two parts were averaged. For example, a student could receive 10 out of 13 on the first part (76.92%) and 19 out of 27 on the second part (70.37%) for an overall mean score of 73.65%.

4.4. Questionnaires

The students in the experimental class were asked to complete questionnaires after they finished each of the four activities during the treatment period (see Appendix G). Each questionnaire consisted of three questions on a five-point scale. The first question asked the students to rate the activity according to interest (très intéressante, intéressante, assez intéressante, un peu intéressante, pas du tout intéressante). The second question asked them to judge the activity in terms of its utility for learning to predict the gender of French nouns (très utile, utile, assez utile, un peu utile, pas du tout utile). The third question asked about the level of difficulty of the activity (très difficile, difficile, ni
difficile ni facile, facile, très facile). These three questions were followed by two open-ended ones. The first asked the students what changes they would make to the activity, if any, and why they would make them. The second invited them to add any further comments they might wish to make concerning the activity.

4.5. Procedures

4.5.1. Pretests

Pretesting in all three classes took place during the third and fourth weeks of semester. This was late February 1995 for Class C2, and late September 1995 for Classes C1 and E. The delay gave the two participating teachers enough time to get well established with their regular program and classroom routines, and to distribute and collect the permission forms. The listening test, which took less than ten minutes, was administered first. It was followed by the word endings test and the agreement test, which were handed out at the same time. Both of these had two versions. The participants in each class were identified by a number according to their alphabetical listing. Odd-numbered students received Form A on the pretest; even-numbered students received Form B (see Appendices C and D). This would be switched for the immediate posttest. On the delayed posttest, participants would receive the same version as for the pretest. Thus odd-numbered students did versions ABA, while even-numbered students did versions BAB. All students completed the
written endings and agreement tests in less than twenty minutes, although no
time limit was given. The writing assignment was completed in class a day or
two after the listening, word endings and agreement tests.

Last in the series was the oral production task, which involved a
subsample of students who were withdrawn individually from each class. This
took place a day or two after the writing assignment. Students taking part in the
oral production task were interviewed individually in a small seminar room near
the classroom. Their responses were taped using a cassette recorder. The
interviewer for all three classes was the same person (the researcher and
teacher of Classes C2 and E). A subsample of 12 students from each class
participated in the first oral test. These students were not chosen randomly.
The researcher asked for volunteers from each class. This was because part of
the agreement with the Research Advisory Committee of the school board was
that students would not be asked to participate in any aspect of the study with
which they felt uncomfortable. Several students in each class were reluctant to
take part in an individual, tape-recorded, oral production activity. Thus it is
possible that the oral production subsample is heavily representative of the
more fluent speakers in each class, and certainly of the more self-confident
ones.

4.5.2. Treatment period.

The treatment period in the experimental class began a week after the
completion of the pretesting, and lasted for just under two months. There were four main components to the treatment activities (for description, see Section 4.6). Originally the researcher had intended to involve the students in more than just four activities focusing on gender. However, it soon became obvious that this would seriously encroach upon the regular program. In the Grade 11 course, students are required to read and discuss five or six plays, novels and films, participate in debates and class presentations, write several expository and creative compositions, and deal with language topics that often present considerable difficulty, such as the subjunctive. In total, about twelve hours of instructional time were devoted to the gender treatment activities during October and November 1995. Students were asked to complete a questionnaire (see Section 4.4 above) at the end of each of the four activities. During the entire treatment period students were constantly reminded to pay attention to gender and agreement in their written and oral assignments.

4.5.3. Posttests.

Immediate posttests were administered during the first week of the fourth month of the semester (May 1995 for C2; December 1995 for C1 and E). For the experimental class, this was the week after the treatment period ended. The same five evaluation instruments were used as in the pretesting. For the word ending and agreement tests, Forms A and B were switched, so that students were given the version that they had not already done. A different topic was
given for the writing assignment (see Section 4.3.4).

The second oral production task (identical in format to the first) was held at the time of the immediate posttests, except for six students in the experimental class, who were interviewed at the time of the delayed posttests. The rationale behind the decision to split the experimental group was to see if there were differences between the students who were tested immediately after the treatment period and those who were tested a month later. Unfortunately, there was not enough time to test the entire group twice. The group was divided on the basis of pretest scores, with the mean score for each subgroup of six being almost identical. Obviously there were methodological disadvantages to this procedure. Only a very small group of six students (those who were tested at the time of the immediate posttest) from Class E could be compared directly to the comparison classes on the second test. Statistically this is a very small number to deal with. However, the splitting of the group did yield some interesting results which will be discussed in Section 5.1.5. Two students, one from Class C1 and one from Class C2, were unavailable for the posttest oral. Also, one of the participating students from Class C2 was the francophone (see Section 4.2) whose results were omitted from the data analysis. Thus the actual number of students tested orally in each class was: \( C1 = 11, C2 = 10, E = 12 \) (6 + 6). The oral production tasks were scored by the investigator, and then a teacher with 14 years' experience in high school French immersion was asked to score 20% of the tests selected at random. The level of agreement was .91, indicating high interrater reliability.
Delayed posttesting took place one month later, in early June of 1995 for Class C2, and in early January of 1996 for Classes C1 and E. There was no writing assignment for the delayed posttest, and oral testing only for the 6 students in the experimental group who did not participate at the time of the immediate posttest. For the word ending and agreement tests, students received the same form as for the pretest. Thus for the three sessions for these two tests, half the class was given ABA, and the other half BAB.

4.6. Description of treatment materials and activities

At the same time as the testing instruments were being developed and piloted, treatment materials and instructional strategies were being devised for use with the experimental class. In this section, the four main activities in which the experimental class was involved during the eight-week treatment period will be described. The amount of time devoted to each activity varied. The total amount of time spent on form-focused instruction dealing with French gender was about twelve hours spread over eight weeks.

Activity #1. Dictionary activity focusing on written word endings. This activity took place in the experimental class the week immediately following the completion of the pretests. The class of 28 was divided into six groups, with 4 or 5 students per group. Each group was given three endings or "bundles" of endings in written form as follows:
Group 1: -age, -esse, -oul/-oil/-ois

Group 2: -isme, -ure, -aise

Group 3: -ment, -erie, -a/-as/-at/-ap

Group 4: -eux, -elle, -tion

Group 5: -eau/-o/-ot, -ette/-otte, -euse

Group 6: -ite/-ute, -ance/-ence, -in/-on/-ond/-ont

Twelve of the endings are the same as the ones chosen for the listening and written ending tests (see Sections 4.3.1 and 4.3.2). The others are endings which, according to the research of Tucker, Lambert and Rigault (1977) and Luce (1979), are highly predictive of noun gender. According to Luce’s figures, they all have at least 6:1, or over 85%, reliability. Groups of endings were sometimes bundled together because they all have the same pronunciation (e.g. -a/-as/-at/-ap), and sometimes because Luce includes them in a single category (e.g. -ette/-otte both belong to the “double consonant plus -e” group).

The groups were provided with French dictionaries and instructed to find ten nouns in each category, for a total of 30 nouns, and to indicate the gender of each noun. When they had finished, they transferred the information to chart paper. The following day the sheets were posted around the classroom. Each group of students explained their findings to the rest of the class. A chart of typically masculine and feminine endings was developed on the blackboard, with each group adding their endings to it in turn. The students were instructed
to copy down the chart and were advised to memorize one key word for each ending. At the end of the activity, which took about two hours over the two days, Questionnaire #1 (see Appendix G) was handed out.

**Activity #2. Word ending clues to noun gender: Rules and exceptions.** This activity, which took place during the second and third weeks of the treatment period, was designed to reinforce and augment the knowledge of word ending clues to gender that was begun in Activity #1. Students were asked to work individually or in pairs. Most opted to work with a partner. The activity was divided into two parts. For the first part, they were given a list of 20 nouns (see Appendix H) and access to dictionaries and Luce’s (1979) article, several copies of which were available in the classroom. They were to find out the gender of each noun, then to state whether or not it conformed to the general rule provided by Luce. For example,

- **-attitude, féminin:** suit la règle qui dit que les noms en -ade, -ude sont féminins dans un rapport de 10 à 1

- **-rêve, masculin:** exception à la règle qui dit que les noms en -ve sont féminins dans un rapport de 10 à 1.

For the second part of the activity, each individual or pair was assigned ten pages from a novel that the class was studying. They were to choose any 20 nouns occurring in the ten pages and to perform the same analysis as in the first part of the activity. When they had finished, they were to meet with at least two other individuals or pairs and compare notes. Any word-ending clues which
they had not yet encountered were to be added to the list that was begun during Activity #1. Altogether, the two parts of the activity took about two hours to complete. Questionnaire #2 (see Appendix G) was handed out when the activity was completed.

**Activity #3.** **Worksheets on adjective endings and determiner-adjective-noun agreements.** This could be described as a traditional grammar exercise. Students were given two worksheets (see Appendix I). The first focused on adjective endings, starting with the most common pattern (adding -e to the masculine to form the feminine), then working through some fairly common masculine/feminine alternations, such as -eux/-euse, -f/-ve, and finishing with irregular forms, such as blanc/blanche, vieux/vieille. On the second sheet, students were given a list of phrases such as *mon oncle favori* and *la nouvelle directrice*. Their task was to change the phrase to the opposite gender (e.g. *ma tante favorite, le nouveau directeur*). None of the material on the worksheets was new to the students. The purpose was to remind them of the importance of making sure that all elements in a noun phrase agree according to gender.

As a follow-up to the two worksheets, the students were again assigned ten pages in the novel that they were studying. This time their task was to find 20 adjectives and their accompanying nouns. The adjectives were to be classified according to the formation rules listed on the first worksheet. As in Activity #2, the students were asked to compare notes with at least two other
individuals or pairs. Both parts of the activity took place during the fourth week of the treatment period. The students completed Questionnaire #3 (see Appendix G) at the end of the activity. Both parts of the activity took place during the fourth week of the treatment period, and their total duration was about one hour and 45 minutes.

During the fifth week of the treatment period, no specific gender-focused activities were planned, since the class was busy with several other aspects of the course.

Activity #4. Student-invented games focusing on gender. This was by far the most time-consuming of the four activities, and took place over about five hours during the last three weeks of the treatment period. About half of the time was devoted to creating the games and half to playing them. Many of the students put in several hours of additional time outside of class. The students were instructed to divide into groups of their own choosing, and to devise any kind of game to practise the word-ending gender clues that had been the focus of the previous activities. Four groups of six to eight students were formed. The games that they developed will be described in Chapter 6.

Questionnaire #4 (see Appendix G) was completed when the students had had the opportunity to play each of the games.

The teacher / researcher's observations of the students' reactions as they took part in the four activities will be described in Chapter 6.
Chapter 5

Results

This chapter, which reports the results of the research project, is divided into three parts. The first part deals with the analysis of the test data. Results of the listening, written endings and agreement tests (pretest, immediate posttest, delayed posttest), and of the writing assignments and oral production task (pretest, posttest) are presented along with statistical analyses of the data. Charts and tables are included to summarize the results of each set of tests. The second part of the chapter, the qualitative analysis, provides sample data from the writing assignments and the oral production task. These give an indication of the wide range of variability among students, information that cannot be gleaned from the quantitative data. In the third part of the chapter, results of the questionnaires completed by students in the experimental class throughout the treatment period are reported.

5.1. Results of test data

The purpose of the study was to investigate the effect of form-focused,
analytic instruction on control over grammatical gender by French immersion students in Grade 11. An experimental class participated in a treatment period lasting for two months, during which time they performed several tasks designed to focus their attention primarily on the word-ending clues to the gender of French nouns, but also on other aspects of the topic, such as the importance of adjective agreement.

Five pretests were administered to the experimental class and the two comparison classes before the treatment period began. Immediate posttests and delayed posttests were given after the treatment period, using the same series of tests. The dependent variable was the students’ control over grammatical gender in French. The main independent variable was class, as the results of the experimental class were compared to the results of two comparison classes that did not receive the treatment.

Mean pretest results for all five measures were compared using the SPSSx Analysis of Variance (ANOVA) program. Significant differences were found for the listening test ($F(2, 57) = 4.81, p < .02$) and the writing assignment ($F(2, 57) = 3.17, p < .05$). On the listening test, there was a difference of about 10% between the highest (Class C1) and lowest (Class E) scores (see Table 5.1). On the writing assignment there was a difference of about 11% between the highest (Class C2) and lowest (Class C1) scores, with the experimental class about halfway between the other two (see Table 5.13). No adjustments were made to pretest scores because the fundamental comparison to be made
among the three classes was their growth over time, and all pretest results left ample room for improvement. Furthermore, the design of the repeated measures MANOVA program takes into account initial differences. Comparisons of the mean pretest scores for the written endings test, the agreement test, and the oral production task did not reveal significant differences among the three classes.

Statistical comparisons of the growth over time of the three classes were made using the SPSSx Multivariate Analysis of Variance (MANOVA) program. This procedure can be used when several groups are involved (in this case three) and repeated measures are needed (three administrations of the listening, written endings and agreement tests; two writing assignments and two administrations of the oral production task). A two-factor design was used. The between-subjects factor was class, with three levels (C1, C2, E). The within-subjects factor was time, with either two levels (pretest and posttest) or three levels (pretest, immediate posttest and delayed posttest). Separate repeated measures analyses were conducted for each of the five tests. The purpose was to compare the results of the three classes over a period of time; that is, to determine whether there were significant differences between their pre- and posttest results. Since there were only two time levels for the writing assignment and the oral production task, MANOVA yielded results that were identical to univariate tests of significance.
5.1.1. Listening test

Table 5.1 gives the means, expressed as a percentage, for all three classes on the three implementations of the listening test. The same information is presented graphically in Figures 5.1 (line graph) and 5.2 (bar graph). Table 5.2 presents the results of the repeated measures MANOVA for the six comparisons that were chosen. These were: 1) the overall time effect for all three classes; 2) the class by time effect (were there significant differences among the three classes over time?); 3) Class E vs. Classes C1 and C2 (was the growth over time of the experimental class different from the two comparison classes combined?); 4) Class C1 vs. Class C2 (did the two comparison classes differ over time?); 5) Class C2 vs. Class E (did the experimental class differ over time from the comparison class taught by the same teacher?); 6) Class C1 vs. Class E (did the experimental class differ over time from the comparison class taught by a different teacher in a different school?).

Table 5.1: Class means on Listening Test (LT)

<table>
<thead>
<tr>
<th></th>
<th>Pretest</th>
<th>Immediate Posttest</th>
<th>Delayed Posttest</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>mean</td>
<td>SD</td>
<td>mean</td>
</tr>
<tr>
<td>Class C1 (n = 17)</td>
<td>72.88</td>
<td>10.09</td>
<td>64.05</td>
</tr>
<tr>
<td>Class C2 (n = 16)</td>
<td>65.28</td>
<td>9.18</td>
<td>66.49</td>
</tr>
<tr>
<td>Class E (n = 27)</td>
<td>62.24</td>
<td>12.67</td>
<td>85.29</td>
</tr>
</tbody>
</table>
Figure 5.1: Line graph of Listening Test (LT) results.

Figure 5.2: Bar graph of Listening Test (LT) results.
The results of the repeated measures MANOVA (Table 5.2) reveal a highly significant time effect when the results for all three classes are combined 
\((F(2, 56) = 14.72, p < .0005)\). Within this overall time effect, however, there are highly significant differences among the three classes, shown by the class by time interaction \((F(4, 112) = 18.27, p < .0005)\). These differences are visually apparent in Figures 5.1 and 5.2, which show Class C2 with very little difference between the pretest, immediate posttest and delayed posttest results. Class C1 has pretest and delayed posttest scores that are almost identical, but there is an obvious dip on the immediate posttest, which is 8 to 9 per cent lower than the other two tests. Only the experimental class shows a large gain on the immediate posttest (about 23%), and this is maintained on the delayed posttest, which is slightly higher than the immediate posttest.

**Table 5.2:** MANOVA of Listening Test administered to three classes on three occasions.

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>Wilks's lambda</th>
<th>(F)</th>
<th>Hypothesis df</th>
<th>Error df</th>
<th>(p)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time</td>
<td>.655</td>
<td>14.72</td>
<td>2</td>
<td>56</td>
<td>&lt; .0005</td>
</tr>
<tr>
<td>Class x Time</td>
<td>.366</td>
<td>18.27</td>
<td>4</td>
<td>112</td>
<td>&lt; .0005</td>
</tr>
<tr>
<td>Expt. vs Comp.</td>
<td>.421</td>
<td>39.24</td>
<td>2</td>
<td>57</td>
<td>&lt; .0005</td>
</tr>
<tr>
<td>C1 vs. C2</td>
<td>.778</td>
<td>4.27</td>
<td>2</td>
<td>30</td>
<td>&lt; .04</td>
</tr>
<tr>
<td>C2 vs. E</td>
<td>.526</td>
<td>18.04</td>
<td>2</td>
<td>40</td>
<td>&lt; .0005</td>
</tr>
<tr>
<td>C1 vs. E</td>
<td>.355</td>
<td>37.33</td>
<td>2</td>
<td>41</td>
<td>&lt; .0005</td>
</tr>
</tbody>
</table>
The comparison of the experimental class with the two comparison classes also shows a highly significant difference ($F(2, 57) = 39.24, p < .0005$), indicating that Class E performed significantly better over time than Classes C1 and C2 combined. This is also true if Class E's results are compared to each comparison class individually: for Class E vs. Class C2 (taught by the same teacher), $F(2, 40) = 18.04, p < .0005$; for Class E vs. Class C1 (taught by a different teacher), $F(2, 41) = 37.33, p < .0005$. While the difference over time between Classes C1 and C2 may also be regarded as significant ($F(2, 30) = 4.27, p < .04$), it is much less so than the other comparisons, and is attributable to the dip in the immediate posttest performance by Class C1.

Table 5.3 gives the pretest, immediate posttest, and delayed posttest mean scores for the experimental class for the six masculine and six feminine endings. A number of observations can be made from these figures. First, there is little difference between masculine and feminine results on any of the three tests. This even masculine/feminine split reflects the finding of the preliminary study (see Section 3.6), and differs from the over-generalization of the masculine found in several studies involving younger learners (Tarone, Frauenfelder and Selinker 1976; Harley 1979; Stevens 1984; Taylor-Browne 1984). Second, results improved for all endings from pretest to immediate posttest. With the exceptions of -isme and -ure, the improvements were maintained, and sometimes increased, on the delayed posttest. (The very slight drops for -esse and -in represent one mistake by one student.) Third, the ranking of the endings in terms of highest to lowest score changed little. The
Table 5.3: Mean scores on Listening Test by ending (experimental class)

<table>
<thead>
<tr>
<th></th>
<th>% Correct (Pretest)</th>
<th>% Correct (Imm. Posttest)</th>
<th>% Correct (Del. Posttest)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Masc. Endings</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-eau / -aut</td>
<td>80.25</td>
<td>93.83</td>
<td>95.06</td>
</tr>
<tr>
<td>-in</td>
<td>77.78</td>
<td>98.77</td>
<td>97.53</td>
</tr>
<tr>
<td>-at / -ap</td>
<td>74.07</td>
<td>91.36</td>
<td>95.06</td>
</tr>
<tr>
<td>-ment</td>
<td>59.26</td>
<td>87.65</td>
<td>92.59</td>
</tr>
<tr>
<td>-age</td>
<td>40.74</td>
<td>80.25</td>
<td>88.89</td>
</tr>
<tr>
<td>-isme</td>
<td>34.57</td>
<td>70.37</td>
<td>61.73</td>
</tr>
<tr>
<td><strong>Avg. (all masc. endings)</strong></td>
<td>61.11</td>
<td>87.04</td>
<td>88.48</td>
</tr>
<tr>
<td><strong>Fem. Endings</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-esse</td>
<td>77.78</td>
<td>100.00</td>
<td>98.77</td>
</tr>
<tr>
<td>-ure</td>
<td>72.84</td>
<td>82.72</td>
<td>71.60</td>
</tr>
<tr>
<td>-ette</td>
<td>71.60</td>
<td>91.36</td>
<td>96.30</td>
</tr>
<tr>
<td>-erie</td>
<td>53.09</td>
<td>81.48</td>
<td>88.89</td>
</tr>
<tr>
<td>-ance / -ence</td>
<td>53.09</td>
<td>80.25</td>
<td>87.65</td>
</tr>
<tr>
<td>-tion</td>
<td>51.85</td>
<td>65.43</td>
<td>75.31</td>
</tr>
<tr>
<td><strong>Avg. (all fem. endings)</strong></td>
<td>63.38</td>
<td>83.54</td>
<td>86.42</td>
</tr>
</tbody>
</table>
endings have been listed in the chart from highest to lowest pretest scores. For the masculine, the only small change for both posttests is that -in, in second place on the pretest, moves into the top position, slightly ahead of -eau/-aut. For the feminine, the only change is that -ure moves from second to third to last place. On the whole, however, the main observation is that there is a consistent tendency for improvement across almost all endings. The generally low scores for -isme, and the apparent drop from immediate to delayed posttest, may be due to the fact that this ending, though 100% predictive of gender, tends to be associated with rather academic vocabulary that high school immersion students may not encounter frequently in everyday speech and reading. The -ure ending, with the drop from immediate to delayed posttest, and almost identical pretest and delayed posttest scores, presents an interesting but probably inexplicable aberration.

5.1.2. Written endings test

During the piloting of the testing materials, the differences obtained in the results for the two versions of the written endings test (Form A and Form B) were found to be not significant (see Section 4.3.2). It was decided, however, that test order (ABA or BAB) was potentially a between-subjects factor in the analysis of the data. Therefore the overall mean scores for the three implementations of each test were compared. For Form A, the mean was 73.24% ($SD = 12.76$), and for Form B, the mean was 73.27% ($SD = 10.91$). The
two means were compared using paired t-tests, and the difference was found to be not significant \((p = .89)\). Consequently, test order was not included in the analysis.

Table 5.4 gives the means, expressed as a percentage, for all three classes on the three implementations of the written endings test. The same information is presented graphically in Figures 5.3 (line graph) and 5.4 (bar graph). Table 5.5 presents the results of the repeated measures MANOVA for the 6 comparisons that were made (see Section 5.1.1).

**Table 5.4: Class means on Written Endings Test (WET)**

<table>
<thead>
<tr>
<th>Class</th>
<th>Pretest</th>
<th>Immediate Posttest</th>
<th>Delayed Posttest</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
</tr>
<tr>
<td>Class C1 ((n = 17))</td>
<td>68.95</td>
<td>10.31</td>
<td>70.26</td>
</tr>
<tr>
<td>Class C2 ((n = 16))</td>
<td>67.71</td>
<td>9.51</td>
<td>65.97</td>
</tr>
<tr>
<td>Class E ((n = 27))</td>
<td>69.86</td>
<td>9.64</td>
<td>80.97</td>
</tr>
</tbody>
</table>
**Figure 5.3:** Line graph of Written Endings Test (WET)

![Line graph of Written Endings Test (WET)]

**Figure 5.4:** Bar graph of Written Endings Test (WET)

![Bar graph of Written Endings Test (WET)]
Table 5.5: MANOVA of Written Endings Test (WET) administered to three classes on three occasions.

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>Wilks's lambda</th>
<th>F</th>
<th>Hypothesis df</th>
<th>Error df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time</td>
<td>.922</td>
<td>2.37</td>
<td>2</td>
<td>56</td>
<td>n.s.</td>
</tr>
<tr>
<td>Class x Time</td>
<td>.810</td>
<td>3.11</td>
<td>4</td>
<td>112</td>
<td>&lt; .02</td>
</tr>
<tr>
<td>Expt. vs Comp.</td>
<td>.753</td>
<td>9.35</td>
<td>2</td>
<td>57</td>
<td>&lt; .0005</td>
</tr>
<tr>
<td>C1 vs C2</td>
<td>.983</td>
<td>.259</td>
<td>2</td>
<td>30</td>
<td>n.s.</td>
</tr>
<tr>
<td>C2 vs E</td>
<td>.739</td>
<td>7.07</td>
<td>2</td>
<td>40</td>
<td>&lt; .005</td>
</tr>
<tr>
<td>C1 vs E</td>
<td>.806</td>
<td>4.94</td>
<td>2</td>
<td>41</td>
<td>&lt; .02</td>
</tr>
</tbody>
</table>

The results of the repeated measures MANOVA were not significant for the overall time effect ($F(2, 56) = 2.37, p = \text{n.s.}$). However, the class by time interaction was significant ($F(4, 112) = 3.11, p < .02$). This is illustrated in Figures 5.3 and 5.4, which show Classes C1 and C2 making little progress from pretest to delayed posttest, and Class C2 actually falling slightly on the immediate posttest. Class E, however, shows a large gain on the immediate posttest (about 11%), with a further small gain on the delayed posttest.

The comparison of the experimental class with the two comparison classes reveals a highly significant difference ($F(2, 57) = 9.35, p < .0005$), indicating that Class E performed significantly better over time than Classes C1
and C2 combined. The results are also significant if Class E is compared to each comparison class individually: for Class E vs. Class C2, $F(2, 40) = 7.07, p < .005$; for Class E vs. Class C1, $F(2, 41) = 4.94, p < .02)$. When the two comparison classes are compared, the result is not significant ($F(2, 30) = 0.259, p = n.s.$).

Table 5.6 gives the pretest, immediate posttest, and delayed posttest mean scores for the experimental class for the twelve written endings (six masculine and six feminine). As with the listening test (Section 5.1.1), a number of observations can be made from these figures. Again, there is little difference between masculine and feminine results on any of the three tests. With the exception of -at, results improved for all endings from pretest to immediate posttest. Results also improved, or remained almost the same, from immediate posttest to delayed posttest. While the overall order of endings in terms of percentage of correct responses appears to be somewhat less stable than on the listening test, especially for the masculine, there are nonetheless some strong tendencies. The endings -esse, -ette and -iel-erie seem to be well established as indicating feminine gender on both the listening and written endings tests. Similarly, the delayed pretest scores are highest for the masculine endings -eau/-aut, -in and -at/-ap on both the listening and written endings tests. It is interesting that the two masculine endings with the lowest scores (-age and -isme) end in -e, and the feminine ending at the bottom of the list (-tion) does not end in -e. Perhaps some students were influenced by the association between final -e and feminine gender. However, these same
Table 5.6: Mean scores on Written Endings Test by ending (experimental class)

<table>
<thead>
<tr>
<th></th>
<th>% Correct (Pretest)</th>
<th>% Correct (Imm. Posttest)</th>
<th>% Correct (Del. Posttest)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Masc. Endings</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-at</td>
<td>91.36</td>
<td>83.95</td>
<td>88.89</td>
</tr>
<tr>
<td>-in</td>
<td>86.42</td>
<td>88.89</td>
<td>95.06</td>
</tr>
<tr>
<td>-eau</td>
<td>83.95</td>
<td>92.59</td>
<td>95.06</td>
</tr>
<tr>
<td>-ment</td>
<td>72.84</td>
<td>86.42</td>
<td>81.60</td>
</tr>
<tr>
<td>-age</td>
<td>46.91</td>
<td>66.67</td>
<td>77.78</td>
</tr>
<tr>
<td>-isme</td>
<td>29.63</td>
<td>67.90</td>
<td>70.37</td>
</tr>
<tr>
<td><strong>Avg. (all masc. endings)</strong></td>
<td><strong>68.52</strong></td>
<td><strong>81.07</strong></td>
<td><strong>84.79</strong></td>
</tr>
<tr>
<td><strong>Fem. Endings</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-ette</td>
<td>95.06</td>
<td>97.53</td>
<td>100</td>
</tr>
<tr>
<td>-esse</td>
<td>87.65</td>
<td>97.53</td>
<td>100</td>
</tr>
<tr>
<td>-ie</td>
<td>69.14</td>
<td>81.48</td>
<td>88.86</td>
</tr>
<tr>
<td>-ance</td>
<td>61.73</td>
<td>74.07</td>
<td>74.07</td>
</tr>
<tr>
<td>-ure</td>
<td>56.79</td>
<td>67.90</td>
<td>65.40</td>
</tr>
<tr>
<td>-tion</td>
<td>56.79</td>
<td>66.67</td>
<td>74.01</td>
</tr>
<tr>
<td><strong>Avg. (all fem. endings)</strong></td>
<td><strong>71.19</strong></td>
<td><strong>80.86</strong></td>
<td><strong>83.72</strong></td>
</tr>
</tbody>
</table>
endings had the lowest scores on the listening test (see Section 5.1.1) as well, and the students did not see the words when taking that test. It is likely, though, that some of the students would visualize the spelling in their minds when hearing the words. Finally, the reader is reminded that all of the endings are over 95% predictive of gender for francophones, according to the statistics of Tucker, Lambert and Rigault (1977) and Luce (1979). The final scores of the experimental students ranged from about 70% to 100%. Clearly some of the endings have become better established for them as predictors of gender than others, and in this regard they still differ from native speakers.

5.1.3. Agreement test

During the piloting of the testing materials, the differences obtained in the results for the two versions of the agreement test (Form A and Form B) were found to be not significant (see Section 4.1.3). As with the written endings test, however, test order (ABA or BAB) was potentially a between-subjects factor in the analysis of the data. Therefore the overall mean scores for the three implementations of each test were compared. For Form A, the mean was 67.28% (SD = 16.93), and for Form B, the mean was 70.56% (SD = 17.78). The two means were compared using paired t-tests, and the difference was found to be not significant (p = 0.56). Consequently, test order was not included in the analysis.

Table 5.7 gives the means, expressed as a percentage, for all three classes on the three implementations of the agreement test. The same
information is presented graphically in Figures 5.5 (line graph) and 5.6 (bar graph). Table 5.8 presents the results of the repeated measures MANOVA for the 6 comparisons that were made (see Section 5.1.1).

Table 5.7: Class means on Agreement Test (AT)

<table>
<thead>
<tr>
<th>Class</th>
<th>Pretest Mean</th>
<th>Pretest SD</th>
<th>Immediate Posttest Mean</th>
<th>Immediate Posttest SD</th>
<th>Delayed Posttest Mean</th>
<th>Delayed Posttest SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class C1 (n = 17)</td>
<td>62.35</td>
<td>15.72</td>
<td>69.12</td>
<td>13.72</td>
<td>67.35</td>
<td>20.24</td>
</tr>
<tr>
<td>Class C2 (n = 16)</td>
<td>68.44</td>
<td>17.20</td>
<td>69.69</td>
<td>13.60</td>
<td>73.13</td>
<td>13.52</td>
</tr>
<tr>
<td>Class E (n = 27)</td>
<td>57.59</td>
<td>18.98</td>
<td>73.70</td>
<td>20.78</td>
<td>77.78</td>
<td>16.66</td>
</tr>
</tbody>
</table>

Figure 5.5: Line graph of Agreement Test (AT)
Figure 5.6: Bar graph of Agreement Test (AT)

Table 5.8: MANOVA of Agreement Test (AT) administered to three classes on three occasions.

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>Wilks's lambda</th>
<th>F</th>
<th>Hypothesis df</th>
<th>Error df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time</td>
<td>.752</td>
<td>9.22</td>
<td>2</td>
<td>56</td>
<td>&lt; .0005</td>
</tr>
<tr>
<td>Class x Time</td>
<td>.811</td>
<td>3.09</td>
<td>4</td>
<td>112</td>
<td>&lt; .02</td>
</tr>
<tr>
<td>Expt. vs Comp.</td>
<td>.832</td>
<td>5.75</td>
<td>2</td>
<td>57</td>
<td>&lt; .006</td>
</tr>
<tr>
<td>C1 vs C2</td>
<td>.964</td>
<td>.558</td>
<td>2</td>
<td>30</td>
<td>n.s.</td>
</tr>
<tr>
<td>C2 vs E</td>
<td>.709</td>
<td>8.22</td>
<td>2</td>
<td>40</td>
<td>&lt; .002</td>
</tr>
<tr>
<td>C1 vs E</td>
<td>.866</td>
<td>3.17</td>
<td>2</td>
<td>41</td>
<td>&lt; .053*</td>
</tr>
</tbody>
</table>

*This figure has been included because it is only marginally higher than the alpha-level of 0.05 that is used in most statistical analysis
The results of the repeated measures MANOVA reveal some significant sources of variation. First, there was an overall time effect \( (F(2, 56) = 9.22, p < .0005) \) that was highly significant. The line and bar graphs (Figures 5.5 and 5.6) indicate that all three classes improved from pretest to immediate posttest, and two classes (C2 and E) continued to improve on the delayed posttest. It is also clear that the greatest gain was made by Class E from the pretest to the immediate posttest. The class by time interaction was also significant \( (F(4, 112) = 3.09, p < .02) \). When the class by time results are further broken down, there is a significant difference between Class E and the two comparison classes combined \( (F(2, 57) = 5.75, p < .006) \), and between Class E and Class C2 \( (F(2, 40) = 8.22, p < .002) \). The results of the other two comparisons, between Class C1 and Class C2 \( (F(2, 30) = .558, p < .6) \), and between Class E and Class C1 \( (F(2, 41) = 3.17, p < 0.053) \) are not significant. However, it should be noted that the \( p \) value for Class E vs. Class C1 \( (p < 0.053) \) is only slightly higher than the generally accepted value for statistical significance \( (p < .05) \), whereas the \( p \) value comparing the two comparison classes is nowhere near the level of significance.

Since the agreement test dealt with both article and adjective agreement, separate statistical analyses for each of these were conducted. Mean scores, expressed as percentages, are given for articles in Table 5.9 and for adjectives in Table 5.10.
Table 5.9: Class means for articles on the agreement test (AT)

<table>
<thead>
<tr>
<th>Class</th>
<th>Pretest Mean</th>
<th>Pretest SD</th>
<th>Immediate Posttest Mean</th>
<th>Immediate Posttest SD</th>
<th>Delayed Posttest Mean</th>
<th>Delayed Posttest SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class C1 (n = 17)</td>
<td>80.60</td>
<td>8.45</td>
<td>80.60</td>
<td>9.15</td>
<td>79.40</td>
<td>13.90</td>
</tr>
<tr>
<td>Class C2 (n = 16)</td>
<td>81.90</td>
<td>10.80</td>
<td>79.05</td>
<td>9.55</td>
<td>77.80</td>
<td>11.25</td>
</tr>
<tr>
<td>Class E (n = 27)</td>
<td>77.05</td>
<td>13.90</td>
<td>85.55</td>
<td>12.25</td>
<td>85.35</td>
<td>12.30</td>
</tr>
</tbody>
</table>

Table 5.10: Class means for adjectives on the agreement test (AT)

<table>
<thead>
<tr>
<th>Class</th>
<th>Pretest Mean</th>
<th>Pretest SD</th>
<th>Immediate Posttest Mean</th>
<th>Immediate Posttest SD</th>
<th>Delayed Posttest Mean</th>
<th>Delayed Posttest SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class C1 (n = 17)</td>
<td>65.00</td>
<td>15.00</td>
<td>72.05</td>
<td>12.75</td>
<td>72.05</td>
<td>14.25</td>
</tr>
<tr>
<td>Class C2 (n = 16)</td>
<td>73.15</td>
<td>12.35</td>
<td>72.20</td>
<td>12.10</td>
<td>74.70</td>
<td>10.85</td>
</tr>
<tr>
<td>Class E (n = 27)</td>
<td>64.05</td>
<td>14.70</td>
<td>77.20</td>
<td>16.80</td>
<td>80.35</td>
<td>13.65</td>
</tr>
</tbody>
</table>

It is clear from Tables 5.9 and 5.10 that the means for article agreement were higher than the means for adjective agreement on all three tests for all three classes. It also appears that Class E made greater gains than either of the two comparison classes for both articles and adjectives. In fact, the two comparison classes showed declining results from pretest to delayed posttest on article
agreement. These results are illustrated by the bar graphs (Figures 5.7 and 5.8).

**Figure 5.7:** Bar graph of Agreement Test (articles only)

**Figure 5.8:** Bar graph of Agreement Test (adjectives only)
Tables 5.11 and 5.12 present the results of the repeated measures MANOVA for article and adjective agreement.

**Table 5.11:** MANOVA of Agreement Test administered to three classes on three occasions (articles only)

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>Wilks's lambda</th>
<th>F</th>
<th>Hypothesis df</th>
<th>Error df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class x Time</td>
<td>.789</td>
<td>2.31</td>
<td>6</td>
<td>110</td>
<td>&lt; .04</td>
</tr>
<tr>
<td>Expt. vs Comp.</td>
<td>.797</td>
<td>4.76</td>
<td>3</td>
<td>56</td>
<td>&lt; .006</td>
</tr>
<tr>
<td>C1 vs C2</td>
<td>.984</td>
<td>0.15</td>
<td>3</td>
<td>29</td>
<td>n.s.</td>
</tr>
<tr>
<td>C2 vs E</td>
<td>.681</td>
<td>6.10</td>
<td>3</td>
<td>39</td>
<td>&lt; .003</td>
</tr>
<tr>
<td>C1 vs E</td>
<td>.860</td>
<td>2.16</td>
<td>3</td>
<td>40</td>
<td>n.s.</td>
</tr>
</tbody>
</table>

**Table 5.12:** MANOVA of Agreement Test administered to three classes on three occasions (adjectives only)

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>Wilks's lambda</th>
<th>F</th>
<th>Hypothesis df</th>
<th>Error df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class x Time</td>
<td>.781</td>
<td>2.41</td>
<td>6</td>
<td>110</td>
<td>&lt; .04</td>
</tr>
<tr>
<td>Expt. vs Comp.</td>
<td>.832</td>
<td>3.76</td>
<td>3</td>
<td>56</td>
<td>&lt; .02</td>
</tr>
<tr>
<td>C1 vs. C2</td>
<td>.897</td>
<td>1.11</td>
<td>3</td>
<td>29</td>
<td>n.s.</td>
</tr>
<tr>
<td>C2 vs E</td>
<td>.693</td>
<td>5.76</td>
<td>3</td>
<td>39</td>
<td>&lt; .003</td>
</tr>
<tr>
<td>C1 vs E</td>
<td>.902</td>
<td>1.45</td>
<td>3</td>
<td>40</td>
<td>n.s.</td>
</tr>
</tbody>
</table>
Clearly, when articles and adjectives are analyzed separately, the results remain similar to those obtained for the agreement test as a whole. There is a significant class by time effect in both cases ($p < .04$). There are significant differences between Class E and the two comparison classes combined ($p < .006$ for articles and $p < .02$ for adjectives). The differences between Class E and Class C2 are also statistically significant ($p < .003$ for both articles and adjectives). The differences between the two comparison classes are not significant, nor are the differences between Class C1 and Class E, in spite of the apparently greater gains made by the experimental class.

5.1.4. Writing assignments

Table 5.13 presents the means, expressed as a percentage, for the results of the first (pretest) and second (posttest) writing assignments that were completed by all three classes. (The reader is referred to Section 4.3.4. for a description of the assignments.) This information is also presented graphically in Figures 5.9 (line graph) and 5.10 (bar graph). While all three classes made gains from pretest to posttest, it appears that the experimental class made greater gains than either of the other two.
Table 5.13: Class means on Writing Assignment (WA)

<table>
<thead>
<tr>
<th>Class</th>
<th>Pretest</th>
<th>Posttest</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>Class C1 (n=17)</td>
<td>62.94</td>
<td>14.73</td>
</tr>
<tr>
<td>Class C2 (n=16)</td>
<td>74.06</td>
<td>11.61</td>
</tr>
<tr>
<td>Class E (n=27)</td>
<td>69.52</td>
<td>12.19</td>
</tr>
</tbody>
</table>

Figure 5.9: Line graph of Writing Assignment (WA)
The results of the repeated measures MANOVA (Table 5.14) reveal a significant overall time effect ($F(1, 57) = 7.27$, $p < .01$), reflecting the fact that all three classes did make gains over time. However, none of the class by time comparisons are significant. Even though Class E made, on average, greater apparent gains than either of the other two classes, the difference was not statistically significant.
Table 5.14: MANOVA of Writing Assignment (WA) administered to three classes on two occasions.

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time</td>
<td>445.04</td>
<td>1</td>
<td>445.04</td>
<td>7.27</td>
<td>&lt; .01</td>
</tr>
<tr>
<td>Class x Time</td>
<td>187.02</td>
<td>2</td>
<td>93.51</td>
<td>1.53</td>
<td>n.s.</td>
</tr>
<tr>
<td>Expt. vs Comp.</td>
<td>172.13</td>
<td>1</td>
<td>172.13</td>
<td>2.85</td>
<td>n.s.</td>
</tr>
<tr>
<td>C1 vs C2</td>
<td>14.89</td>
<td>1</td>
<td>14.89</td>
<td>0.41</td>
<td>n.s.</td>
</tr>
<tr>
<td>C2 vs E</td>
<td>73.90</td>
<td>1</td>
<td>73.90</td>
<td>0.89</td>
<td>n.s.</td>
</tr>
<tr>
<td>C1 vs E</td>
<td>171.64</td>
<td>1</td>
<td>171.64</td>
<td>2.96</td>
<td>n.s.</td>
</tr>
</tbody>
</table>

When the compositions of the students in the experimental class were analyzed only for the noun endings that were the focus of the treatment activities (see Section 4.6 for a list of these endings), the students made somewhat greater gains from pretest to posttest. Their pretest score for these endings was 69.47% (almost identical to their overall pretest score, see Table 5.9). However, their posttest score for these endings was 82.47%, compared to an overall gender accuracy score of 76.67%.

5.1.5. Oral production task

It was explained in Section 4.5 that the 12 students in the experimental
class who participated in the oral production task at the time of the pretest were divided into two groups of six for the posttest. Although this presented obvious methodological problems in terms of making statistical comparisons with the two comparison classes, it was felt that it would be interesting to compare the two experimental subgroups to see if there were differences between the six students who did the second oral immediately after the treatment period (designated Class E*), and those who were tested one month later (designated Class E**). It turned out that there was very little difference between the results of the two groups. The immediate posttest group, with a mean score of 76.83 on the pretest, had a posttest score of 85.67. The delayed posttest group, with a mean score of 77.0 on the pretest, obtained a slightly higher posttest score of 87.43.

Table 5.15 gives the means, expressed as a percentage, for all three classes on the oral production task. The same information (except for Class E**) is presented graphically in Figures 5.11 (line graph) and 5.12 (bar graph). The mean pretest scores for the experimental class and one of the comparison classes (C1) were very similar, with the other comparison class (C2) starting out 7 to 8% behind the other two. On the posttest Class C1 and Class E (both E* and E**) showed gains. Class C2, however, scored slightly lower on the posttest than on the pretest.
Table 5.15: Class means on the Oral Production Task (OPT)

<table>
<thead>
<tr>
<th>Class</th>
<th>Pretest</th>
<th>Posttest</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>Class C1 ($n = 11$)</td>
<td>75.91</td>
<td>9.72</td>
</tr>
<tr>
<td>Class C2 ($n = 10$)</td>
<td>69.00</td>
<td>17.63</td>
</tr>
<tr>
<td>Class E* ($n = 6$)</td>
<td>76.83</td>
<td>13.54</td>
</tr>
<tr>
<td>Class E** ($n = 6$)</td>
<td>77.03</td>
<td>13.83</td>
</tr>
</tbody>
</table>

Figure 5.11: Line graph of the Oral Production Task (OPT)
Figure 5.12: Bar graph of Oral Production Task (OPT)

Table 5.16: MANOVA of the Oral Production Task (OPT) administered to three classes on two occasions. The experimental (E*) group contains only 6 of the 12 students who participated in the OPT pretest.

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>Sum of Squares</th>
<th>$df$</th>
<th>Mean Square</th>
<th>$F$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Time</td>
<td>176.93</td>
<td>1</td>
<td>176.93</td>
<td>2.93</td>
<td>n.s.</td>
</tr>
<tr>
<td>Class x Time</td>
<td>137.65</td>
<td>2</td>
<td>68.82</td>
<td>1.14</td>
<td>n.s.</td>
</tr>
<tr>
<td>Expt. vs Comp.</td>
<td>17.19</td>
<td>1</td>
<td>17.19</td>
<td>0.27</td>
<td>n.s.</td>
</tr>
<tr>
<td>C1 vs C2</td>
<td>120.46</td>
<td>1</td>
<td>120.46</td>
<td>2.11</td>
<td>n.s.</td>
</tr>
<tr>
<td>C2 vs E*</td>
<td>73.63</td>
<td>1</td>
<td>73.63</td>
<td>1.25</td>
<td>n.s.</td>
</tr>
<tr>
<td>C1 vs E*</td>
<td>0.52</td>
<td>1</td>
<td>0.52</td>
<td>0.01</td>
<td>n.s.</td>
</tr>
</tbody>
</table>
The results of the repeated measures MANOVA (Table 5.16) show no significant overall time effect for the three classes combined, and no class by time effect. However, if Classes C1 and E* are taken as a group, there is a significant time effect \( F(1, 21) = 9.34, p < .007 \), reflecting the gains made by these two classes from pretest to posttest.

As with the writing assignments, the oral production tests of the experimental class were analyzed for the noun endings that were the focus of the treatment activities. On the pretest, the students scored 74.4% on these endings, a result that was slightly lower than their overall score of 76.9% (groups E* and E** combined). Their posttest score for the endings that were taught in class was 91.4%, compared to an overall score of 86.6% (again, groups E* and E** combined).

It is important to remember, however, that all statistical comparisons for the oral production task are of questionable value because of the relatively small number of students involved.

5.1.6. **Summary of statistical analyses**

The class by time effect was highly significant for the listening test, and significant for the written endings and agreement tests, but not for the writing assignments and the oral production task.

Further breakdowns of the class by time effect revealed highly significant
differences between the experimental and comparison groups on the listening test, whether the comparison classes were taken individually or combined ($p < .0005$). Students in Class E made considerable gains over time, whereas the results for Classes C1 and C2 changed little from pretest to delayed posttest, and Class C1 actually registered a decline of over 7% on the immediate posttest.

The breakdown of the written endings test results revealed significant differences between the experimental and comparison groups on the class by time interaction. Of all five measures, this was the one where the three classes were closest on pretest results (only 2.1% separated the highest from the lowest). Class E made considerable gains over time (over 14%) in contrast to Classes C1 and C2, which registered very modest gains from pretest to delayed posttest, with Class C2 actually showing a slight decline from pretest to immediate posttest.

The agreement test results revealed significant class by time differences between Class E and the combined comparison classes, and between Class E and Class C2. Although Class E made greater gains than Class C1, the difference fell just beyond the limit of $p < .05$ for statistical significance. When the data were broken down into articles and adjectives, there was little change to the statistical results. However, it was clear from a comparison of the means that all three classes scored consistently higher for article agreement than for adjective agreement.
No significant class by time interactions were revealed for either the writing assignments or the oral production task, even though the gains made by the experimental class appeared to be greater than those made by either of the comparison classes. However, it was noted that when the assignments and orals of the experimental class were analyzed only for the endings that were the focus of the treatment activities, the results were almost 5.8% higher on the written posttest and almost 4.8% higher on the oral posttest than the overall scores for these two posttests.

In general, the class by time interaction was most significant in comparisons of listening test results. This was due to the considerable improvement over time shown by the experimental class. A highly significant difference was also found between the experimental and comparison groups for the written endings test, and a somewhat less but still significant difference was found between the experimental and comparison groups for the agreement test.
5.2. Qualitative analysis

This section will provide samples from the writing assignments, as well as sample data from the oral production task. The intent is not to provide an in-depth qualitative analysis of all the data collected, since that would go far beyond the purpose of the study. However, the samples will enable the reader to link scores to actual written and oral output. The samples were chosen to illustrate the wide variation of responses among students, with some making large gains from pretest to posttest, while others appear to demonstrate little or no progress.

The first two samples are from the student in the experimental class (E-23) who demonstrated the greatest progress from the pretest to the posttest writing assignment. The first segment is taken from the first writing assignment, on which she scored 57% for correct gender assignment and adjective agreement. Gender errors are underlined in the text. Other types of errors have not been corrected, nor are they highlighted in the text.

Puis nous faisons cuire la dîner sur une feu. Nous sommes avec une groupe de 9 personnes de 14 à 16 ans et 2 conseillers. Nous devons tous travailler comme un groupe si on veut arriver au destination. Pendant tout notre voyage nous n'avons pas rencontré un personne hors de la groupe. C'était si abandonné et isolé. Comme groupe, on travaille bien et c'était le plus grand aventure de nos vies et à la même temps le plus amusant. Aussi après notre voyage nous avons retourné au "Base Camp". C'est situé sur le lac [name of lake] et c'était comme une petite ville.

(E-23)

The next sample is taken from the second writing assignment, on which the
Dans le cas de Kevin, c’est l’IRA et la séparation des catholiques et des protestants en Irlande. Mais dans le cas de Julien il habite dans une petite ville en France pas loin de Paris, et c’est la deuxième guerre mondiale. Julien a été séparé de sa famille pour aller habité à l’école. Kevin a une très grande famille en Irlande. Il y a neuf enfants. Julien a seulement un frère qui va à la même école que lui. Une différence entre les deux est que Julien Quentin était l’auteur du livre. L’auteur a juste changé son nom et il a écrit un livre de sa jeunesse. Mais dans l’autre livre “Au-Delà des Barricades”, c’est seulement un roman mais l’information et les situations dans le livre sont vraies. (E-23)

In the first sample there are many examples of incorrect gender assigned to articles and adjectives. There is also inconsistency, as groupe is incorrectly identified twice as feminine and once correctly as masculine. In the second sample there are no examples of incorrect gender except for the masculine form of the possessive adjective son used with the feminine noun famille. This could possibly result from confusion due to the masculine gender of the possessor Julien, although the same mistake is not made with sa jeunesse. The reader is reminded that possessives were excluded from the treatment activities since they differ psycholinguistically from articles and descriptive adjectives and could by themselves provide the focus for an entire study (see Section 3.12). With respect to endings that were taught in class, there are errors in the pretest sample (au destination, le grand aventure). In the posttest sample, the correct gender is attributed to all nouns with endings that were the focus of instruction
This student appears to have transferred the metalinguistic knowledge and gender awareness acquired during the treatment period to her own written output.

The next two samples are from the compositions by a student in the experimental class (E-13) whose pretest and posttest scores on the writing assignment were almost identical (71% and 72%). The treatment activities appear to have resulted in no improvement in her spontaneous written output, even though her listening, written endings and agreement test scores improved by an average of about 8% from pretest to delayed posttest.

Quand on marche plus loin, les troncs devient de plus en plus épais et ils montent de plus en plus haut. Il y a quelques-uns qui sont plus vieilles que la maison, leurs branches sont forts et ils bloquent le soleil. La terre qui les entoure est brun et nu, sauf que quelques feuilles morts de l’automne précédent. Les membres (sic) de ces grandes arbres regardent comme milliers de petites bras qui essayent d’atteindre le soleil. (E-13)

L’amour n’est pas le seul cause de toutes les malheurs dans les vies des deux filles. La seule support que Ciboulette reçoive est de ses amis, une bande de contrebandiers. Le seul personne qui donne l’attention et l’amour à Fleurette est Joseph. Joseph est son frère et il est un voyou. Elles reçoivent l’amour des plus mauvaises caractères. (E-13)

What is particularly striking about these samples is the difficulty with, or lack of attention to, the gender of adjectives. The student is aware of the fact that there are masculine and feminine forms of adjectives, since several of each are used;
however, her assignment of gender to adjectives appears to be quite haphazard. It is also noteworthy that the adjectives are always correct with respect to number (singular or plural), indicating at least partial knowledge of the requirement of adjective concord in French. With respect to endings that were the focus of instruction, she attributed the correct gender to *gazon*, *ruisseau*, *mouette* and *distance* on her pretest, but was wrong with respect to *bras* and *ressemblance*. On the posttest, she correctly identified the gender of *vie*, *tragédie*, *apparence*, *voyou* and *fin* (an exception), and made no errors with endings that were the focus of instruction. This perhaps indicates increased awareness of the predictability of the endings that were studied in class, although there were really too few words in the samples to make this statement with certainty.

The next example is from a student in Class C2 on the posttest writing assignment. This student displays very little control over gender, and his second composition received a mark of just over 50% for gender assignment and agreement (as did his pretest writing assignment). Notice that even the noun *directrice*, for which there is a correlation with natural gender, is assigned the wrong grammatical gender.
Mes rêves de **la** bonheur sont **différentes** des personnages dans Zone. **Une** exemple est que je ne suis pas pauvre comme Ciboulette, alors je ne dois pas penser comment je peux sortir de **ce** situation. Aussi je ne voudrais pas acheter une musique à bouche avec beaucoup de clés comme Moineau. Ma mère est **un** directrice et mon père n’est pas un alcoolique, alors l’argent nécessaire pour vivre la mère de Passe-Partout ne peut pas être comparer avec moi. **Le seul** comparaison de **la** bonheur avec moi et les personnages dans Zone est ...

The final example is taken from a student who participated in the testing but whose results were not included in the data because of his francophone background. While his writing contains many grammatical errors, especially with respect to unpronounced plural markings and verb endings, his control over gender appears to be perfect:

L’école a de très belle facilité qui compte plusieurs salle d’ordinateur, deux salles de musique, une menuiserie, garage d’autos et plusieur autre cour et salle intéressantes. Mes passe-temps implique d’écouter de la musique avec mes amis. Mes classes ce semestre sont l’anglais, le français, les mathématiques et l’éducation physique. J’aime tué le temps la fin de semaine à me promener dans le bas de la ville, c’est là que je me suis perdu au début et j’ai appris les rues à la force de retrouver mon chemin. Dans le bas de la ville, j’aime bien rentrer dans les magasins de guitare. Ce sont des endroits que je connais bien. (C2-9)

Many other samples of student writing could be included, but they would not really add to the overall impression provided by those that have already
been cited.

It is evident from both the quantitative results of the writing assignments (Section 5.1.4) and some of the samples of students' writing examined in this section, that some students in the experimental class made considerable progress, while others did not. It therefore seemed worthwhile to compare the writing assignment results with other test results. The agreement test was chosen for comparison because it was the only other evaluation measure to deal with both article and adjective agreement. Scores on the agreement test were examined for students who made considerable progress from the first to the second writing assignment, and for students who made no progress. The arbitrary criterion chosen to define 'considerable progress' was that they improved their score by at least 10 percentage points from the first to the second writing assignment (e.g. 62% to at least 72%; 75% to at least 85%). 'No progress' was defined as having the same or a lower score on the second writing assignment than on the first one. Eleven of 27 students in the experimental class fell into the category of 'considerable progress'; eight students fell into the 'no progress' category. Ten of the 11 who had demonstrated considerable progress on the writing assignment also had posttest agreement test scores (both immediate and delayed) at least 10% above their pretest scores. Of the eight students who made no progress on the writing assignment, four made little or no improvement on the agreement test either. Three improved their agreement test results by at least 10% (both posttests). The final student had little room to improve on the agreement test,
since he scored 95% on the pretest (and 100% on both posttests). His two writing assignment scores were 69% and 67%. Although there are obviously not enough data for statistical comparisons, it is clear that for 14 of 19 students, the writing assignment results paralleled the agreement test results. The relationship is particularly striking for students who improved on the writing assignment, since 10 out of 11 of them also improved on the agreement test. While improvement on the agreement test was not necessarily accompanied by improvement on the writing assignment (since 4 of 8 who made no improvement on the writing assignment either improved, or consistently scored high, on the agreement test), improvement on the writing assignment almost always was accompanied by improvement on the agreement test. In other words, it was rare (only one case out of 11) for a student to show improvement on the writing assignment without also making gains on the agreement test. The agreement test, as we have seen, was very limited in terms of the number of nouns and adjectives used. Furthermore, the worksheets (treatment Activity #3) dealing with agreement contained a very limited range of nouns and adjectives. It seems, however, that the focus on the importance of determiner-adjective-noun agreements may have provided the stepping stone that allowed many (but not all) students to then generalize their knowledge to the more open-ended task of writing a composition.

The pattern for the oral production task was similar to that of the writing assignments. Some students demonstrated considerable progress from pretest to posttest, while others made smaller gains or none at all. Several (including
some from the experimental group) scored lower on the posttest than on the pretest.

The transcripts are taken from the picture description segment of the oral production task. The reader will remember that this was the more spontaneous of the two parts of the test, requiring the students to describe a picture of a farm setting. The first two examples are from a student who made considerable progress, scoring 61% on the pretest and 89% on the posttest. Gender errors have been underlined.

\textit{Je vois dans le champ un cheval et un mouton et \underline{un} vache. Il y a une fille près d'un table avec \underline{une} gateau sur. Et il y a \underline{un} auto. Il y \underline{une} arbre avec \underline{un} bicyclette près de l'arbre. La femme travaille avec \underline{un} rateau dans le jardin. L'homme fait la pêche et il prend \underline{une} poisson. Un homme travaille sur la maison. Le garçon regarde \underline{un} balle. \underline{Une} écureuil et \underline{un} dindon sont près du "barn". Et il y a \underline{un} cadeau sur la terre. (E*-12, pretest interview)}

\textit{Je vois une auto près de ... le lac et une fille qui fait \underline{une} party. Elle a un cadeau et un gateau. Il y a \underline{un} maison avec une échelle près de, et un homme travaille sur la maison ... sur le toit. Il y a un lac et une balle et un lapin ... et aussi \underline{une} arbre avec \underline{un} pomme sur la terre. La femme a \underline{un} rateau. Un garçon est assis. Il regarde la balle. Et le cheval et les moutons ... et une \underline{vache} est dans le champ. \underline{Un} écureuil regarde dans le jardin. (E*-12, posttest interview)}

Note that the inconsistency with respect to the [o] ending that is evident in the first excerpt appears to have been corrected in the posttest. Even the gender of
auto, which was taught in class as an exception, is correct in the second sample. Some words that were used in both tests, such as balle and écureuil, are assigned the correct gender the second time, even though they do not have endings that were taught in class. Overall, this student appears to have carried over the gender awareness acquired during the treatment period to the spontaneous production of the second oral test.

The next excerpt is taken from the pretest of a student in Class C2. His overall pretest score for the two parts of the oral production task was 59%.

C'est un ville (sic) et il y a un bateau et un personne qui fait la pêche et un voiture et un party avec une fille qui a un gateau qui ... elle a mangé un pièce, et il y a un maison avec ... qui ressemble à un maison ordinaire avec deux étapes (sic) et deux fenêtres et un porte et un homme qui travaille sur ... l'échelle. Il y a un petit fenêtre sur le toit. Il y a les animaux, un cheval il y a un petit jardinière avec un homme ... les oiseaux ont peur de lui. Et il y a un petit lapin et une mère fait ... (C2-14)

It is noteworthy that all of the errors are in the same direction, with masculine articles and adjectives used with feminine nouns. This student's posttest demonstrated the same pattern, and even included un petit fille:
Je vois un petit lac, et un homme qui a juste attrapé un poisson, un homme qui travaille sur le toit de sa maison sur un échelle. Il y a un petit ferme, un lapin qui court sur le gazon, un automobile, un coq, son territoire (?), un moufette, un petit fille, du gâteau sur le table, un cadeau, un petit bateau, une femme, le gazon, un sapin, un bicyclette et un pommier. (C2-14)

Overgeneralization of the masculine was not, however, common to all students. The errors of student E*-12 quoted above are about evenly split between the two genders. Student C2-6 overgeneralized the feminine, using une épouvantail, la petite lac, une radeau and une lapin. Similarly, all of the errors made by student C1-6 were with masculine nouns: une mouton, une bateau, une cadeau, une poisson, une lapin, une arbre.

A number of students exhibited a tendency to neutralize the phonological distinction between un [œ̃] and une [yn] by producing a blend of the two [œ̃n]. (Harley [1979] noted a similar phenomenon for le and la [see Section 2.3].) When this occurred, it was marked as an error, since it was unclear whether or not the student knew the gender of the word. The most notable case of this was student C1-13, all of whose errors were of this nature. When the articles were pronounced clearly, the correct gender was always chosen. This student appears to have developed a 'neutral' form of the article as a strategy to be used when he is uncertain of the gender of a noun.
The final excerpt is taken from the student with francophone background, whose results were not included in the statistical analysis.

Je vois une ferme, dans la cour il y a un cheval, un mouton, deux vaches... une table avec une personne avec un chapeau de fête, un petit bateau, une famille de canards, un pêcheur, une maison puis un homme qui fait de la réparation sur le toit ... une madame qui passe le rateau sur le gazon, un homme assis sur le tronc d'un arbre coupé, un petit quai qui flotte sur l'eau, et je vois un coq, puis la grange, un sac à côté de l'homme qui fait de la pêche, un lapin ... entre la maison et la grange il y a un pommier, il y a un bicycle à côté sur (?) l'arbre, il y a une mouflte, une pelle, un sapin devant la maison ... il y a un cadeau dans la gauche, une voiture en plein milieu, puis c'est à peu près tout. (C2-9)

With this student, there is no doubt or hesitation as to which gender to assign to every noun.

5.3. Results of questionnaire data

Upon completion of each of the treatment activities, students in the experimental class were asked to fill out a questionnaire (see Appendix G) rating the interest level, usefulness, and level of difficulty of the activity. A five-point scale was used for each of the three criteria, and there was also room for written comments. Table 5.17 gives the mean scores for the four activities in each of the three categories. The reader is referred back to Section 4.6 for details of the four activities, and to Section 4.4 for a more detailed description of
the questionnaires.

**Table 5.17:** Students’ ratings of the treatment activities.⁴

<table>
<thead>
<tr>
<th>Activity #1</th>
<th>Interest</th>
<th>Usefulness</th>
<th>Difficulty</th>
</tr>
</thead>
<tbody>
<tr>
<td>Activity #2</td>
<td>3.2</td>
<td>4.2</td>
<td>2.3</td>
</tr>
<tr>
<td>Activity #3</td>
<td>3.1</td>
<td>3.4</td>
<td>2.1</td>
</tr>
<tr>
<td>Activity #4</td>
<td>3.0</td>
<td>3.9</td>
<td>2.4</td>
</tr>
<tr>
<td>Activity #4</td>
<td>4.6</td>
<td>4.4</td>
<td>2.4</td>
</tr>
</tbody>
</table>

The same information is presented graphically in Figures 5.13 and 5.14. Figure 5.13 provides a visual comparison of the four activities grouped according to the three criteria of interest, usefulness and difficulty. Figure 5.14 provides an individual picture of each activity.

---

⁴The reader is reminded here that a 5-point scale was used on the questionnaires, with 5 as the highest rating and 1 as the lowest. (See Section 4.4 and Appendix G for more detail.)
Figure 5.13: Class E students' ratings of the four treatment activities, grouped according to the three criteria of 'interest', 'usefulness' and 'difficulty'.

Figure 5.14: Class E students' ratings of each of the four treatment activities according to 'interest', 'usefulness' and 'difficulty'.

Question 1 asked students in the experimental class to rate the interest level of each of the four activities on a five-point scale ranging from très intéressante (5) to pas du tout intéressante (1). Of the four activities, the student-invented games proved to be by far the most popular, with a 4.6 rating (between intéressante and très intéressante). The other three activities all scored around the mid-point of the interest scale (assez intéressante 'somewhat interesting').

For Question 2, which dealt with the usefulness of each activity for learning to predict the gender of French nouns and/or to make correct agreements, all activities scored well above 3 (assez utile 'somewhat useful') on the five-point scale, with the games (#4) and the dictionary activity (#1) scoring between 4 and 5 (utile to très utile). Question 3 dealt with the difficulty level of the activities. None was judged to be difficult, as all four scored between 2 (facile 'easy') and 3 (ni difficile ni facile 'neither difficult nor easy'). It is noteworthy that for Activities #1, #2, and #3, usefulness outscored interest (see Figure 5.14), perhaps indicating that the students recognize the need for certain types of learning tasks, even if they find them only moderately interesting.

Only a few students added written comments for Question 4, which asked them if and how they would suggest changing each activity, and for Question 5, which invited them to add any other comments. Since there were so few comments, they can all be included here. They have not been edited to correct errors in grammar, spelling, etc.

For Questionnaire #1, the following comments were made:
Tu peux faire dans un jeu.

C'est un excellent méthode pour apprendre si certains mots sont masc. ou fem. Peut-être cette activité peut être enseignée dans les grades 9 ou 10.

C'était bon en groupes.

Trouvez une autre façon de présenter les mots. J'aime pas les présentations.

One student has in fact anticipated the fourth activity by suggesting a game format. Another has suggested that Grade 11 is indeed too late to begin a form-focused approach to learning gender. The third student likes the idea of working in groups. The last comment simply expresses the student's aversion to presenting in front of the class, as each group was required to explain their findings to the rest of the class.

For Questionnaire #2, the comments were:

Quelques mots qu'on a trouvé n'étaient pas sur la liste des règles et on a dû trouver d'autres mots, ex. les mots qui terminent en -le (ex. sable) et -ort (ex. port).

Des mots que j'ai trouvé dans le roman, je peux pas trouver dans les règles.

Plus de mots qu'on utilise chaque jour dans le vocabulaire français

The first two students both appear to be objecting to the fact that the word-ending rules are not foolproof, and that there will always be some nouns that are not easily predictable according to their ending and which must be
memorized. This itself is perhaps a valuable lesson to learn. (The -ort ending is, however, included in Luce's [1979] article, but under a more general heading dealing with nouns ending in a consonant.) The third student appears to be suggesting that the activity should concentrate on common, everyday words, rather than on words that students find in a novel they are studying, which may include vocabulary that is not as frequent.

For Questionnaire #3, the following comments were made:

*Tu peux faire un jeu avec cette activité pour toute la classe.*

*Pas si long.*

*C'est assez facile et tu peux apprendre beaucoup sur masculin et féminin accords.*

Again, the suggestion is made that a game be developed, anticipating the last of the four activities. The second student appears to think that any kind of traditional grammar activity is too long! The third student finds the activity useful for learning masculine and feminine agreements.

For Questionnaire #4, which dealt with the game activity, the comments were:

*Très amusant.* (two students)

*Est-ce qu'on peut faire encore?*

*Beaucoup des groupes ont fait des jeux qui sont joué avec toute la classe. Je pense que les jeux sont meilleur si nous les jouons en petites*
groupes, plusieurs fois.
C'était très amusant et on a mémorisé les terminaisons sans savoir.
J'aimais beaucoup.

On the whole, these comments echo the positive rating which was given this activity on the five-point scale, in terms of both interest and usefulness. Even the comment asking for more group work and less whole-class work seems to imply that more exposure to the activity would be advantageous.
Classroom Observations

This chapter will describe the teacher/researcher's observations of the experimental class during the eight-week treatment period. The four treatment activities will be reviewed briefly, but the reader is referred back to Section 4.6 for more detailed descriptions. The observations will focus on the amount of time taken to complete each activity, and on the reactions of the students in the class to the activities. It is important to remember that the majority of class time during the treatment period was devoted to "regular" aspects of the course (literature study, discussions, writing, presentations, and so on) and that the descriptions reported here are limited to the activities that relate to this study.

6.1. Activity #1. Dictionary activity focusing on written word endings

This activity began in the experimental class the week immediately following the completion of the pretests. The class divided into groups of their own choosing, with four or five students per group. Altogether there were six
groups. Each group was given three written noun endings (e.g. -isme, -ure, -aise) and a French dictionary. They were instructed to find ten nouns for each ending, and to indicate the gender of every noun. When they had finished, they transferred the information to chart paper. It took about one hour for all the groups to finish. As the students saw that all or almost all of the nouns with the same ending had the same gender, they caught on very quickly to the purpose of the activity, and started to comment to each other and to the teacher about the associations between endings and gender. During the dictionary search, all the groups worked diligently, and appeared to enjoy the task. Some groups deliberately searched for rarely used words that were totally unfamiliar to them, such as gargouillement and scarification. On several occasions the teacher/researcher overheard students discussing the meanings of words as well as their gender. As is so often the case in the classroom, an instructional task designed for a specific purpose led to learning in other areas.

The following day the sheets of chart paper were posted around the classroom. Each group explained their findings to the rest of the class. For every ending or group of endings, at least eight and usually nine or ten of the nouns were of the same gender. For example, Group 1 found 10 masculine nouns ending in -age, 10 feminine nouns ending in -esse, and 9 masculine nouns ending with -ou/-oi/ois (la paroi was the exception). A chart of typically masculine and feminine endings was developed on the blackboard, with each group adding their endings to it in turn. The students were instructed to copy down the chart and were advised to memorize one key word for each ending.
During the presentations the teacher pointed out common exceptions that did not appear on the charts, such as the one-syllable feminine nouns ending in -age: *page, cage* and *plage*. Soon the students got involved in trying to think of exceptions. Everyone was attentive and appeared interested during the group presentations. About one hour was spent on the second day, making a total of two hours for the entire activity.

The charts remained posted in the classroom for the duration of the treatment period as a constant reminder to the students of the noun ending/gender correspondences. One day, after the Grade 11 class had finished, the teacher/researcher was still in the room as the Grade 10 immersion students (taught by a different teacher) were entering. One student looked at the charts and was overheard remarking to a friend that she thought that their class should do something similar to help them learn gender better.

6.2. *Activity #2. Word-ending clues to noun gender: Rules and exceptions*

This activity was designed to reinforce and augment the knowledge of word-ending clues to gender that was begun in *Activity #1*. The students were asked to work individually or in pairs. Most opted to work with a partner. The activity was divided into two parts. For the first part, they were given a list of 20 nouns (see Appendix H), a copy of Luce's (1979) article (summarized in
Section 2.5), and access to French dictionaries. They were to look up the gender of each noun in the dictionary, then to state whether or not it conformed to the general rule provided by Luce. For the second part of the activity, which took place the following week, each individual or pair was assigned ten pages from a novel that the class was studying. They were to choose any 20 nouns occurring in the ten pages and to perform the same analysis as in the first part of the activity. Any word-ending clues which they had not yet encountered were to be added to the list that was begun during Activity #1. Altogether, the two parts of the activity took about two hours to complete during the second and third weeks of the treatment period. As with the first activity, the students appeared to be interested in the work, and did not have to be reminded to stay on task. There were some discussions with the teacher when students could not find a rule in Luce's article for a noun that they found in the novel, such as système. Usually there was a logical explanation which the teacher was able to point out to the students. For example, Luce only includes 'rules' that have at least five to one (83.3%) reliability. The teacher showed the students Tucker, Lambert and Rigault's (1977) statistics, which are more detailed than Luce's. They indicate that \(-\dot{\text{e}}\text{me}\) nouns are about 82.4% masculine. The students then decided to add \(-\dot{\text{e}}\text{me}\) nouns to their list of predominantly masculine endings. This is just one small example of the type of fruitful discussion that took place during the activity.
6.3. **Activity #3. Worksheets on adjective endings and determiner-adjective-noun agreements.**

In this activity, the students were given two worksheets (see Appendix I) which focused on adjective endings and determiner-adjective-noun agreements. In addition to providing practice in making agreements, the exercise turned out to be a useful vocabulary review. Several questions such as "Quel est le masculin de ‘vache’?" were overheard. The two worksheets took about half an hour to complete.

As a follow-up to the two worksheets, the students were again assigned ten pages in the novel that they were studying. This time their task was to find 20 adjectives and their accompanying nouns. The adjectives were to be classified according to the formation rules listed on the first worksheet. As in Activity #2, the students were asked to compare notes with at least two other individuals or pairs. This part of the activity took about 45 minutes. Both parts of the activity took place during the fourth week of the treatment period.

6.4. **Activity #4. Student-invented games focusing on gender.**

This was by far the most time-consuming of the four activities, and took place over about six hours during the last three weeks of the treatment period. About half of the time was devoted to creating the games and half to playing
them. Many of the students reported that they put in several hours of additional preparation time outside of class. The students were instructed to divide into groups of their own choosing, and to devise any kind of game to practise the word-ending gender clues that had been the focus of the previous activities. Four groups of six to eight students were formed, and four very different but all very effective and entertaining games were developed. These will be described briefly.

Group 1: Concentration Game. This group developed two concentration boards. The boards consisted of 36 numbered squares, each containing a hidden word. The object was to find pairs of words with the same ending, such as *nation*/*station*, *moulin*/*dessin*, and to correctly state the gender of the words. Most of the words followed the ending ‘rules’, but a few exceptions were included to make the game more challenging.

Group 2: Card Game. The second group devised a card game derived from ‘Fish’ in which participants had to collect sets of cards containing words having the same ending, such as *-age* or *-tion*. The cards were all hand made and decorated with fish that were obviously male or female depending on the gender associated with the ending. Thus as the game was played, participants could form visual associations with the endings.

Group 3: Board Game. Group 3 adapted a commercial board game called "Boule de Feu" in which the winner is the first player to get from beginning to end of a path while attempting to avoid many hazards and delays
along the way. The students created a set of question cards that required the players to identify the gender of nouns if they landed on designated spaces. Thus it was their knowledge of word-ending clues to gender that got them safely past hazards. The majority of the nouns followed the ending 'rules', but some exceptions were included to add to the challenge.

Group 4: Gender ‘Field Day’. The fourth group organized a kind of ‘Field Day’ in the classroom with a series of events and competitions. There were relays, races, and skill contests, all of which required participants to correctly identify the gender of nouns in order to continue playing and ultimately to win.

The students obviously found the games activity very enjoyable, as their responses to Questionnaire #4 (see Section 5.3) indicate. Much thought and creativity went into the development of each game, and the students had to constantly discuss masculine and feminine endings as they were creating them. The competitive element required the participants to remember gender clues quickly and accurately, thus providing a kind of grammar drill, but in a ‘fun’ setting.
Chapter 7

Discussion and Conclusions

This chapter concludes the dissertation. It summarizes the research and its findings, and discusses their significance. Then the limitations of the study are considered in terms of the research design and the testing instruments. Finally, implications for future research and pedagogy are presented.

7.1. Summary of the study

The research set out to investigate the effect of form-focused, analytic instruction on control over grammatical gender by French immersion students in Grade 11. There were two main hypotheses. The first was that form-focused instruction can improve students' proficiency and output with respect to a specific linguistic feature, such as gender, when introduced into a primarily experiential program. The second hypothesis was that it is not too late to change linguistic behaviour even when learners may have been making the same errors for years, as appears to be the case for Grade 11 immersion students with grammatical gender. Sixty-two students took part in the study.
The experimental class consisted of 28 students and was taught by the researcher. There were two comparison classes. One of these (C2), with 17 students, was also taught by the researcher. The other comparison class (C1), also with 17 students, was taught by another teacher in a different school.

The study consisted of pretests, a treatment period lasting about eight weeks in the experimental class, immediate posttests and delayed posttests. The five tests used were: 1) listening test; 2) written endings test; 3) agreement test; 4) writing assignment; and 5) oral production task. During the treatment period, the experimental group completed several form-focused treatment activities designed to make them aware of word-ending regularities as clues to the gender of French nouns, and to provide practice in making appropriate gender agreements. Other aspects of the study included teacher observations of the experimental class during the treatment period, and questionnaires designed to elicit the students' reactions to the treatment materials. Statistical analyses, using SPSSx MANOVA, were used to compare the test results over time of the three classes.

7.2. Summary of the findings

A comparison of the means for all five tests reveals that the experimental group was the only one of the three classes to score higher on all posttests (immediate and delayed) than pretests. Furthermore, the percentage increase
from pretest to immediate posttest, or from pretest to delayed posttest, was always greater for the experimental class than for either of the comparison classes.

The results of the repeated measures MANOVA revealed a class by time interaction effect that was significant for three of the five measures used: the listening, written endings, and agreement tests. In the case of the listening test, whether the comparison classes were taken individually or combined, the gains made by the experimental class over time were significantly greater. Class E's score increased by about 25% from pretest to delayed posttest. Neither of the comparison classes showed an increase of more than 2%.

The breakdown of the written endings test results also revealed significant differences between the experimental and comparison groups on the class by time interaction. By the delayed posttest, Class E had improved by about 14%, whereas the two comparison classes had made only small gains of less than 3%.

The agreement test results revealed significant class by time differences between Class E and the comparison classes. The experimental group's score increased by about 20% from pretest to delayed posttest. Neither of the comparison classes improved by more than 5%.

For the writing assignments and oral production task, the repeated measures MANOVA revealed no significant differences over time between the experimental and comparison classes. However, a comparison of the class
means shows a tendency for Class E's scores to improve more than either of the comparison classes on both of these measures. Class C1 had by far the lowest score on the pretest writing assignment, and made very little progress by the posttest (less than 1.5%). Class C2 started out with the highest score on the writing assignment and made small gains on the posttest (about 3%). The experimental class, which started out behind C2 on the first writing assignment, improved by over 7%, and came very close to equalling Class C2's score on the posttest. On the oral production task, Class C2 started out lowest of the three classes and actually declined slightly on the posttest. Class C1 and Class E were almost equal on the pretest, and both made modest gains on the posttest, with the experimental group showing a tendency towards more improvement (10% vs. 6%). Furthermore, Class E's posttest scores on both tests were higher for the endings that were the focus of instruction in class.

While the quantitative analysis revealed no statistically significant gains over time for Class E on the writing assignments and oral production task, the qualitative analysis indicated that the written and oral accuracy of individual students in the experimental group appears to have benefitted from the treatment activities. These students made fewer gender errors on their posttest compositions and orals than on their pretests, especially with respect to endings that were the focus of the treatment activities. Even though the qualitative analysis was not an in-depth examination of all the data collected, it did reveal tendencies towards improved productive skills on the part of several experimental students.
The questionnaire data indicate that most students in the experimental class gave a positive evaluation to the treatment activities in terms of both interest and usefulness. While there is not necessarily a correlation between a positive assessment of instructional activities by students and actual learning, it is encouraging to have the students' endorsement of the form-focused methodology that was used. Clearly they did not feel that the treatment activities were either boring or a waste of time. Individual comments by some students indicated that they felt that they did gain knowledge of the relationship between noun endings and grammatical gender in French. While this metalinguistic knowledge did not necessarily translate into statistically significant improvement in oral and written spontaneous language production, it did at least provide the students with an additional tool for monitoring their output.

7.3. *Significance of the findings*

We saw in the preceding section that the experimental class made greater percentage gains from pretest to posttest (immediate and delayed) on all five of the evaluation measures than did the comparison classes. On three of these measures, the class by time interaction proved to be statistically significant, indicating reliable gains for the experimental group. It therefore appears that a form-focused approach can help to improve French immersion students' control over grammatical gender, even when they are nearing the end
of their high school French program. Their interlanguage is not necessarily fossilized, even though they may have been repeating the same gender and agreement errors, or taking a rather haphazard approach to gender assignation, for many years. It is interesting that the greatest gains were made by the experimental class on the listening test, given that the treatment activities focused almost entirely on written word endings. It would seem that for older students, the visual presentation of written noun endings carries over to phonological endings as the students make associations between the two.

It is important, however, to inject a cautionary note into the discussion. The three measures for which statistically significant differences were found (listening test, written endings test and agreement test) did not require the students to apply their knowledge to their own spontaneous written and oral production. In the case of the listening and written endings tests, the experimental students demonstrated that they recognized that a relatively small number of endings (12) were highly predictive of noun gender. In the case of the agreement test, while they improved both in assigning gender to nouns and in making adjectives agree with nouns, the number of nouns and adjectives involved was very limited. On the more open-ended and productive tests (the writing assignment and the oral production task), the experimental students tended to outperform those in the comparison classes, but the differences were not statistically significant. This indicates that while the experimental students gained a great deal of awareness of the word-ending clues to noun gender and the importance of making adjective-noun agreements, they made less progress
in terms of applying this knowledge to their own output, whether oral or written. It is also possible that limitations of the research design, as well as the amount of time that could be devoted to instruction in grammatical gender during the treatment period, may have influenced the results. These limitations will be discussed in Section 7.4.

The results of this study generally support the research that has called for the inclusion of analytic strategies to complement the essentially experiential approach of immersion programs (e.g. Swain and Carroll 1987; Lyster 1990; Harley 1993). The main thrust of the immersion program remains experiential, with the students using French to discuss issues, explore literature, appreciate media, and so on. However, explicit focus on L2 features that French immersion students find difficult (in this case grammatical gender) certainly heightens their awareness of the importance of these features, and appears to improve the output of at least some of the students. It should also be noted that it is not a question of either an experiential or an analytic approach at any given moment, but that the two are truly integrated. For instance, one of the treatment activities in the present study required the students to search for nouns in the novel that they were actually studying in class, and the games activity required them to do a great deal of discussing and planning in the target language.

The present study adds to the growing body of form-focused experimental research that has produced generally positive results (Tucker, Lambert and Rigault 1977; Tomasello and Herron 1988; Harley 1989, in press;
Specifically with respect to French grammatical gender, the present research corroborates previous findings that have shown that anglophone learners experience great difficulty with this feature, which does not occur in their native language (Spilka 1976; Tarone, Frauenfelder and Selinker 1976; Tucker, Lambert and Rigault 1977; Harley 1979; Stevens 1984; Taylor-Browne 1984; Marinova-Todd 1994). Both the exploratory study and the pretest results in the main study demonstrated that about 30 to 35% of the time, Grade 11 FI students will assign the incorrect gender to a noun or adjective (this for a feature where guessing should yield about 50% accuracy). This can be contrasted to the 95% accuracy rate demonstrated by the French exchange students on the listening and written endings tests. Difficulty with gender appears to apply to all ages of learners. The present study seems to be unique to date in focusing exclusively on senior high school students.

7.4. Limitations of the study

The small sample of 62 students used in the present study may limit the generalizability of the findings. Only three classes were involved, and two of these, together comprising 34 students, were comparison classes. The number
of students in the experimental class who were involved in the treatment activities was only 28. The results of two students (one in Class C2 and one in Class E) were not used for statistical purposes because of their francophone backgrounds. Thus the results actually involve only 60 students.

There were, however, advantages to the small sample size. The researcher and teacher were the same individual for two of the classes; thus, instructional and testing procedures were able to be carefully controlled. The sample size, though small, was made up of a relatively homogeneous population. The two schools that participated are located about two kilometres from each other, and the socio-economic composition of the areas that they serve is very similar. Apart from the two francophones whose scores were not used for statistical purposes, only one student reported speaking a language at home other than English. Thus the possibility of third-language influence on the students' performance with respect to French gender was minimal.

There were some limitations in terms of the testing instruments and procedures that need to be taken into account. It was pointed out in the preceding section that the listening and written endings tests dealt with a limited number of endings (12) that are highly predictive of the gender of French nouns. The agreement test also dealt with a very limited number of nouns and adjectives. Thus the greater gains made over time by the experimental class, which were statistically significant for these three tests, may not apply to a wider range of endings, especially those that are less reliable predictors of gender.
Nor did the testing instruments in any way measure the students' capacity to deal with exceptions to the general noun-ending rules. While common exceptions do not seem to pose a problem for francophones, the same may not be true for immersion students whose contact with French is mainly limited to classroom input, much of which is provided by other immersion students and may not be correct.

The oral production task, while somewhat more open-ended than the three tests discussed in the preceding paragraph, was still tightly controlled. The first part of the task, in which the students were required to identify the gender of the same objects in both the pretest and posttest, involved a very small number of nouns (13), all of which had endings that are highly predictive of gender. The second part, which involved the description of a farm scene, required somewhat freer production on the part of the students, but the number of objects in the picture was finite, and again the noun endings were for the most part highly predictive of gender. Neither of the two parts of the oral production task called for totally spontaneous oral production. It must also be recalled that the number of students participating in the oral testing (36) was even smaller than for the other four tests. The decision to split the experimental students into two groups of six to form immediate and delayed posttest subgroups (see Section 4.5.3) further weakened the validity of the statistical comparisons for the oral production task.

Of the five testing instruments used, the writing assignments were by far
the most open-ended. While the very nature of the listening, word endings, and agreement tests would lead the students to realize very quickly what was being tested, there was nothing in the structure of the writing assignments to indicate that gender was, in fact, the focus. The compositions were assigned as part of the students' regular work, the topics were chosen to fit in with the literature program, and the only specific instruction that the students received with respect to language was to try to include a large number of adjectives. While the gains made by the experimental class were not statistically significant when compared to the comparison group, the fact that Class E students tended to increase their scores more than students in the other classes on the most open-ended of the tests is encouraging. Also, the qualitative analysis revealed that some of the students in the experimental class improved on the second composition, specifically with respect to noun-endings that had been the focus of instruction.

One can speculate that statistically significant results might have been found for written and oral production if more time had been spent on the treatment activities, or if a greater variety of treatment activities had been developed, especially activities involving more open-ended productive tasks. However, there were overriding constraints that limited the treatment activities in terms of the length of the treatment period and the amount of time that could reasonably be devoted to grammatical gender in any week. Since the entire experiment had to fit into a semester, and a certain amount of time had to elapse between the immediate and delayed posttests, eight weeks was the
maximum length of time for the treatment period. Also, in order not to encroach on the other components of the Grade 11 program (the reader is referred back to Section 4.5.2 for details), it was necessary to strictly limit the total number of hours that could be devoted to the gender experiment. While eleven to twelve hours may not seem like much, it actually represented about 10% of the total amount of instructional time for the course.

There is also a potential Hawthorne effect. The experimental students clearly understood the reason for the repeated testing and were able to relate the tests to the activities that had taken place in class. Thus, they may have been more motivated than the comparison students to put forth their best effort during the posttesting. It may be noted that Class C1 scored lower on the delayed agreement posttest than on the immediate posttest, and that Class C2 scored higher on the first oral test than on the second. It is possible that students in the comparison classes gradually lost interest in doing the same tests over again without any apparent reason or reward. What may have started out as an interesting diversion from “normal” class activities at the time of the pretest may have become rather tedious and pointless for them, although it must be emphasized that the students in both comparison classes remained cheerful, polite and co-operative throughout the three testing sessions.
7.5. **Implications for pedagogy**

The statistically significant gains made by the experimental class on the listening, written endings, and agreement tests point to a heightened awareness of the word-ending clues to noun gender in French and of the importance of making adjective-noun agreements. However, the non-statistically significant results of the more open-ended and productive tests (the writing assignment and the oral production task) indicate that the experimental students made less progress in terms of applying this knowledge to their own oral and written output. It is therefore important to reflect upon changes that might be made to the treatment materials and activities in order to help students produce more correct output with respect to gender and agreement.

The treatment activities were heavily weighted in terms of "recognition" and "manipulation" of linguistic data. The students were required to look words up in dictionaries, match endings and gender, study and discuss word-ending "rules" and exceptions, and complete "traditional" grammar exercises. This was true even of the most open-ended and creative of the activities, the student-invented games. While it is not suggested that any of the treatment activities be eliminated, it is clear that they need to be complemented with tasks of a more productive nature. The students might have performed better on the posttest writing assignment if the treatment period had included a composition activity that focused on the correct use of gender and agreement. For example, simply writing a short composition in or out of class and then exchanging it with
another student whose task would be to proofread for gender and agreement errors might be useful.

It is also clear that the treatment activities were heavily weighted in favour of tasks involving reading and writing, rather than listening and speaking. While listening skills do not appear to have suffered, since there was carry-over from the recognition of written endings to phonological endings (see Section 7.3), oral production does not seem to have fared as well. Therefore it seems essential to develop tasks that would help students to improve their spoken output. One possibility is that small groups of students could tape record their discussions about any topic, such as a novel that they are studying, then play back the tape and listen specifically for gender and agreement. While such activities can be very time consuming, they can be integrated with other aspects of the course, such as literary analysis, so that an inordinate amount of time is not devoted to one aspect of grammar.

7.6. Implications for future research

The findings of the present study provide qualified support for the hypothesis that a form-focused approach to the presentation of grammatical gender can improve high-school immersion students' control over this feature of French, and therefore that incorrect linguistic behaviour with respect to gender is not fossilized at this level. The statistically significant results obtained on
three of the tests (listening, word endings and agreement) point to a heightened awareness on the part of the experimental students of the importance of word endings as clues to noun gender, and an ability to generalize their knowledge to new nouns. While the results of the writing assignments and the oral production task were not statistically significant, the fact that the experimental students tended to outperform the comparison students on both of these measures provides some support for the assertion that the students in Class E improved with respect to their productive control over gender as a result of the form-focused treatment activities. This tentative finding needs to be investigated in research projects designed with testing instruments and treatment activities that place a greater emphasis on oral and written productive skills than was the case in the present study. Some of the changes that could be made to the testing instruments and research design have already been outlined in Section 7.4.

The study needs to be repeated, incorporating the modifications suggested in Section 7.5, with many more participants. As was pointed out in Section 7.4, the small number of subjects in the present study limits the generalizability of the findings. It would also be useful to conduct studies at different grade levels. Harley (in press) has already investigated the role of form-focused tasks in promoting the second language acquisition of children in Grade 2 immersion, specifically with respect to the formal clues of gender. That study and the present one deal with grades near the beginning and the end of the immersion program. It would be interesting to investigate the effects of a
form-focused approach at an age level near the middle of the spectrum, such as Grade 6 or 7. A comparison of the various studies could provide insights into the appropriate sequencing of activities related to grammatical gender throughout the elementary and secondary school years of the French immersion program. For example, at what point should the focus on endings that are predictive of gender begin? Which endings should be introduced at an early stage and which should be left for later? When should adjective agreement be emphasized? When is it appropriate to insist on gender accuracy in written output? When should exceptions be introduced? These are just a few of the questions that need to be addressed.

The generally positive results that have been obtained in this and other form-focused studies (the reader is referred back to Section 7.3 for references) are encouraging. As we have seen (Section 2.7), the overall assessment of French immersion programs has been generally positive; however, concerns have been expressed regarding the accuracy of students' output, even after many years of exposure to the target language. If studies continue to yield positive results when form-focused components are incorporated into FI courses, then the case for the integration of analytic and experiential language teaching strategies will be strengthened.
BIBLIOGRAPHY


APPENDIX A:

LETTER REQUESTING PARENTS’ OR GUARDIANS’ CONSENT

Note: The letters were sent home on school letterhead.
Dear Parents or Guardians:

I am conducting a research project which involves the development of curriculum materials to improve the French proficiency of students in French immersion programs. To assist in the development of these materials, I will be administering three brief tests in FIF 3A1 (Grade 11 immersion) classes at [name of school] this semester. I am therefore requesting your permission to have your son or daughter participate in the testing.

The results of the testing are for the purpose of curriculum development and will not be used to calculate the student’s mark in French. In order to ensure confidentiality, students will not be identified by name but by a numerical code. Your son/daughter may request to withdraw from the testing at any time.

In order that I can obtain oral as well as written data, one of the activities will involve tape-recording some of the students during the testing. Again, no identities will be disclosed.

This project has been approved by the Research Advisory Committee of the [name] Board of Education and has the support of [name], principal of [name of school]. The research is part of my doctoral studies at the Ontario Institute for Studies in Education. My main goal is the improvement of instruction in French immersion programs through research in curriculum development. If you have any questions about the study, I can be reached at [name of school and phone number] or at home (416-767-7885).

Yours truly,

Michael Warden
Department of Modern Languages

________________________________________________________________________

Please return by [date]

I give permission for _______________ to participate in the French language testing.

Yes _______ / No ____________

I give permission for _______________ to be tape-recorded during the oral portion of the testing.

Yes _______ / No ____________

Signature of parent or guardian: ____________________________________________
APPENDIX B: STUDENT ANSWER SHEET FOR THE LISTENING TEST

Instructions: Si vous pensez que le mot que vous entendez est masculin, encerclez UN; si vous pensez que c'est féminin, encerclez UNE. Chaque mot sera prononcé deux fois.

| 1. | UN  | UNE  | 19. | UN  | UNE  |
| 2. | UN  | UNE  | 20. | UN  | UNE  |
| 3. | UN  | UNE  | 21. | UN  | UNE  |
| 4. | UN  | UNE  | 22. | UN  | UNE  |
| 5. | UN  | UNE  | 23. | UN  | UNE  |
| 6. | UN  | UNE  | 24. | UN  | UNE  |
| 7. | UN  | UNE  | 25. | UN  | UNE  |
| 8. | UN  | UNE  | 26. | UN  | UNE  |
| 9. | UN  | UNE  | 27. | UN  | UNE  |
| 10. | UN | UNE  | 28. | UN  | UNE  |
| 11. | UN | UNE  | 29. | UN  | UNE  |
| 12. | UN | UNE  | 30. | UN  | UNE  |
| 13. | UN | UNE  | 31. | UN  | UNE  |
| 14. | UN | UNE  | 32. | UN  | UNE  |
| 15. | UN | UNE  | 33. | UN  | UNE  |
| 16. | UN | UNE  | 34. | UN  | UNE  |
| 17. | UN | UNE  | 35. | UN  | UNE  |
| 18. | UN | UNE  | 36. | UN  | UNE  |
APPENDIX C: STUDENT ANSWER SHEET FOR THE WRITTEN ENDINGS TEST (FORM A)

Instructions: Voici une série de mots inventés. Encerclez “un” ou “une” pour indiquer si vous pensez qu’ils seraient du genre masculin ou féminin.

1. un    une  douteuse
2. un    une  navage
3. un    une  luminisme
4. un    une  navure
5. un    une  bossement
6. un    une  lapaderie
7. un    une  pousseau
8. un    une  minisme
9. un    une  tamisation
10. un    une  tramure
11. un    une  poutance
12. un    une  frottin
13. un    une  cétifat
14. un    une  rougément
15. un    une  mangeance
16. un    une  soupinerie
17. un    une  mameau
18. un    une  magette
19. un    une  tabarat
20. un    une  armation
21. un    une  menacement
22. un    une  nordage
23. un    une  soustrette
24. un    une  filoturre
25. un    une  glacin
26. un    une  névrat
27. un    une  passerette
28. un    une  lanteau
29. un    une  tesse
30. un    une  dansisime
31. un    une  couration
32. un    une  florage
33. un    une  balin
34. un    une  trapézie
35. un    une  ermitance
36. un    une  vivesse
APPENDIX C: STUDENT ANSWER SHEET FOR THE WRITTEN ENDINGS TEST (FORM B)

Instructions: Voici une série de mots inventés. Encerclez "un" ou "une" pour indiquer si vous pensez qu'ils seraient du genre masculin ou féminin.

1. un   une   caravette
2. un   une   topisme
3. un   une   provation
4. un   une   boulerie
5. un   une   falesse
6. un   une   factance
7. un   une   riveau
8. un   une   paladure
9. un   une   évventat
10. un   une   crinisme
11. un   une   salument
12. un   une   brisage
13. un   une   serjat
14. un   une   paramesse
15. un   une   brumance
16. un   une   ruvin
17. un   une   girement
18. un   une   garagette
19. un   une   branchure
20. un   une   gouresse
21. un   une   pendage
22. un   une   clapiteau
23. un   une   drapologie
24. un   une   résument
25. un   une   punin
26. un   une   coupation
27. un   une   purilisme
28. un   une   membre
29. un   une   galance
30. un   une   chômure
31. un   une   gradeau
32. un   une   floration
33. un   une   drapin
34. un   une   placat
35. un   une   gouttage
36. un   une   effarette
APPENDIX D: STUDENT ANSWER SHEET FOR THE AGREEMENT TEST
(FORM A)

Instructions: Encerclez "le" ou "la", ou "un" ou "une" pour indiquer le genre du nom dans chaque phrase, puis traduisez le mot entre parenthèses en français.

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1. le / la</td>
<td>reine</td>
<td>(dead)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. le / la</td>
<td>nuage</td>
<td>(grey)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. le / la</td>
<td></td>
<td>(old) voiture</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. un / une</td>
<td>garçon</td>
<td>(aggressive)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. le / la</td>
<td>climat</td>
<td>(mild)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. le / la</td>
<td></td>
<td>(first) numéro</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. le / la</td>
<td></td>
<td>(new) machine</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. un / une</td>
<td>sucette</td>
<td>(delicious)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. un / une</td>
<td></td>
<td>(beautiful) jardin</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. le / la</td>
<td>mitaine</td>
<td>(white)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. le / la</td>
<td></td>
<td>(last) minute</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12. un / une</td>
<td>oncle</td>
<td>(happy)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13. un / une</td>
<td>journal</td>
<td>(provincial)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14. le / la</td>
<td>réponse</td>
<td>(false)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15. un / une</td>
<td>femme</td>
<td>(Canadian)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16. le / la</td>
<td>drapeau</td>
<td>(blue)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17. un / une</td>
<td>danseuse</td>
<td>(active)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18. un / une</td>
<td>estomac</td>
<td>(round)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19. un / une</td>
<td>église</td>
<td>(traditional)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20. un / une</td>
<td></td>
<td>(good) chien</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX D: STUDENT ANSWER SHEET FOR THE AGREEMENT TEST (FORM B)

Instructions: Encerclez "le" ou "la", ou "un" ou "une" pour indiquer le genre du nom dans chaque phrase, puis traduisez le mot entre parenthèses en français.

1. le / la ________________ (new) garage
2. un / une ________________ (old) crapaud
3. un / une homme ________________ (active)
4. le / la fourrure ________________ (soft)
5. un / une occasion ________________ (happy)
6. le / la ________________ (last) accident
7. le / la roi ________________ (dead)
8. un / une chemise ________________ (blue)
9. un / une soldat ________________ (Canadian)
10. le / la ________________ (first) chandelle
11. un / une élection ________________ (provincial)
12. un / une chapeau ________________ (white)
13. le / la directrice ________________ (aggressive)
14. le / la poubelle ________________ (round)
15. un / une ________________ (false) bijou
16. un / une ________________ (good) tante
17. le / la baleine ________________ (grey)
18. le / la fromage ________________ (delicious)
19. le / la christianisme ________________ (traditional)
20. un / une ________________ (beautiful) couronne
APPENDIX E: ORAL PRODUCTION TASK, PICTURE IDENTIFICATION (SAMPLE)
APPENDIX F: ORAL PRODUCTION TASK, PICTURE DESCRIPTION
APPENDIX G:

QUESTIONNAIRES COMPLETED BY STUDENTS IN THE EXPERIMENTAL CLASS
QUESTIONNAIRE #1

Activité #1: Recherche de mots selon leur terminaison.

1. Est-ce que cette activité était______________?

<table>
<thead>
<tr>
<th>5</th>
<th>4</th>
<th>3</th>
<th>2</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>très intéressant</td>
<td>assez intéressant</td>
<td>un peu intéressant</td>
<td>pas du tout intéressant</td>
<td></td>
</tr>
</tbody>
</table>

2. Avez-vous trouvé cette activité utile pour apprendre à prédire le genre des noms en français?

<table>
<thead>
<tr>
<th>5</th>
<th>4</th>
<th>3</th>
<th>2</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>très utile</td>
<td>assez utile</td>
<td>un peu utile</td>
<td>pas du tout utile</td>
<td></td>
</tr>
</tbody>
</table>

3. Est-ce que vous avez trouvé le travail difficile ou facile à faire?

<table>
<thead>
<tr>
<th>5</th>
<th>4</th>
<th>3</th>
<th>2</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>très difficile</td>
<td>ni difficile</td>
<td>facile</td>
<td>très facile</td>
<td></td>
</tr>
</tbody>
</table>

4. Est-ce que vous changeriez l'activité? Comment et pourquoi?

5. Autres commentaires:
QUESTIONNAIRE #2

Activité #2: Recherche des mots dans l'article et dans le roman.

1. Est-ce que cette activité était _______________?

très intéressante     assez intéressante     un peu intéressante     pas du tout intéressante

2. Avez-vous trouvé cette activité utile pour apprendre à prédire le genre des noms en français?

très utile     assez utile     un peu utile     pas du tout utile

3. Est-ce que vous avez trouvé le travail difficile ou facile à faire?

très difficile     ni difficile     facile     très facile

4. Est-ce que vous changeriez l'activité? Comment et pourquoi?

5. Autres commentaires:
QUESTIONNAIRE #3

Activité #3: Exercices écrits

1. Est-ce que cette activité était______________?

<table>
<thead>
<tr>
<th>5</th>
<th>4</th>
<th>3</th>
<th>2</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>très intéressante</td>
<td>assez intéressante</td>
<td>un peu intéressante</td>
<td>pas du tout intéressante</td>
<td></td>
</tr>
</tbody>
</table>

2. Avez-vous trouvé cette activité utile pour apprendre le genre et les accords?

<table>
<thead>
<tr>
<th>5</th>
<th>4</th>
<th>3</th>
<th>2</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>très utile</td>
<td>assez utile</td>
<td>un peu utile</td>
<td>pas du tout utile</td>
<td></td>
</tr>
</tbody>
</table>

3. Est-ce que vous avez trouvé le travail difficile ou facile à faire?

<table>
<thead>
<tr>
<th>5</th>
<th>4</th>
<th>3</th>
<th>2</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>très difficile</td>
<td>ni difficile</td>
<td>facile</td>
<td>très facile</td>
<td></td>
</tr>
</tbody>
</table>

4. Est-ce que vous changeriez l’activité? Comment et pourquoi?

5. Autres commentaires:
QUESTIONNAIRE #4

Activité #4: Les jeux

1. Est-ce que cette activité était______________?

5 4 3 2 1
très intéressante assez un peu pas du tout intéressante intéressante intéressante intéressante

2. Avez-vous trouvé cette activité utile pour apprendre le genre et les accords?

5 4 3 2 1
très utile assez utile un peu utile pas du tout utile

3. Est-ce que vous avez trouvé le travail difficile ou facile à faire?

5 4 3 2 1
très difficile ni difficile facile très facile
difficile ni facile facile

4. Est-ce que vous changeriez l'activité? Comment et pourquoi?

5. Autres commentaires:
RÈGLE OU EXCEPTION?

Indiquez si les mots suivants sont masculins ou féminins. Donnez le numéro de la règle et indiquez si le mot suit la règle ou s'il est une exception.

Exemple: LA cloche, numéro 20, suit la règle

1. attitude ____________________________
2. rêve ____________________________
3. scie ____________________________
4. blessé ____________________________
5. zone ____________________________
6. activité ____________________________
7. silence ____________________________
8. orage ____________________________
9. facture ____________________________
10. cynisme ____________________________
11. monopole ____________________________
12. peigne ____________________________
13. cerise ____________________________
14. grammaire ____________________________
15. auberge ____________________________
16. miracle ____________________________
17. bombe ____________________________
18. sauterelle ____________________________
19. montre ____________________________
20. dictionnaire ____________________________
APPENDIX 1: STUDENT WORKSHEET USED IN TREATMENT ACTIVITY #3

MASCULIN ET FÉMININ DES ADJECTIFS

1) Généralement, pour former le féminin, on ajoute ______ à la forme masculine:

Exemple: ___________________(m.) ______________________ (f.)

2) -eux devient _________

Exemple: ___________________(m.) ______________________ (f.)

3) -on et -en deviennent _________ et _________

Exemples: ___________________(m.) ______________________ (f.)

____________________ (m.) ______________________ (f.)

4) -el et -ell deviennent _________________ et _________________

Exemples: ____________________(m.) ______________________ (f.)

____________________ (m.) ______________________ (f.)

5) -f devient ____________

Exemple: ___________________(m.) ______________________ (f.)

6) -er devient ______________

Exemple: ___________________(m.) ______________________ (f.)

7) -ou devient ______________

Exemple: ___________________(m.) ______________________ (f.)

8) -eau devient ______________

Exemple: ___________________(m.) ______________________ (f.)

Il y a quelques adjectifs que, malheureusement, il faut mémoriser, comme blanc, doux, faux, favori, frais, gros, et vieux. Donnez le féminin pour ces adjectifs.
APPENDIX I: STUDENT WORKSHEET USED IN TREATMENT ACTIVITY #3

LE GÉNRE DES NOMS ET DES ADJECTIFS

Donnez le masculin ou le féminin, selon le cas, des groupes suivants:

1. mon oncle favori ____________________________
2. la nouvelle directrice ____________________________
3. la reine morte ____________________________
4. ce frère doux ____________________________
5. le chien noir ____________________________
6. ton ami italien ____________________________
7. une vache blanche ____________________________
8. un boulanger heureux ____________________________
9. la belle femme ____________________________
10. un acteur actif ____________________________
11. Cher Monsieur ____________________________
12. un garçon compliqué ____________________________
13. une patronne sèvre ____________________________
14. la poule brune ____________________________
15. cette enfant ridicule ____________________________
16. le prince Jean ____________________________
17. un professeur fou ____________________________
18. son père gentil ____________________________
19. ma nièce menteuse ____________________________
20. un poisson mâle ____________________________
21. une souris femelle ____________________________