uOttawa

L'Université canadienne
Canada's university

FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTOCTORALES

uOttawa

L'Université canadienne
Canada's university

FACULTY OF GRADUATE AND POSDOCTORAL STUDIES

Hazel Curties

AUTEUR DE LA THÈSE / AUTHOR OF THESIS

M.A. (Translation)

GRADE / DEGREE

School of Translation and Interpretation

FACULTÉ, ÉCOLE, DÉPARTEMENT / FACULTY, SCHOOL, DEPARTMENT

The Bilingual Canadian Dictionary: What? How? Why?

TITRE DE LA THÈSE / TITLE OF THESIS

Roda P. Roberts

DIRECTEUR (DIRECTRICE) DE LA THÈSE / THESIS SUPERVISOR

CO-DIRECTEUR (CO-DIRECTRICE) DE LA THÈSE / THESIS CO-SUPERVISOR

EXAMINATEURS (EXAMINATRICES) DE LA THÈSE / THESIS EXAMINERS

Lynne Bowker

Aline Francoeur

Gary W. Slater

Le Doyen de la Faculté des études supérieures et postdoctorales / Dean of the Faculty of Graduate and Postdoctoral Studies

THE BILINGUAL CANADIAN DICTIONARY: WHAT? HOW? WHY?

Hazel Curties

Supervised by

Roda P. Roberts, Ph.D.

School of Translation and Interpretation

Thesis submitted to the

Faculty of Graduate and Postdoctoral Studies

In partial fulfillment of the requirements

For the MA degree in Translation

School of Translation and Interpretation

Faculty of Arts

University of Ottawa

Your file Votre référence
ISBN: 978-0-494-25758-6
Our file Notre référence
ISBN: 978-0-494-25758-6

ACKNOWLEDGEMENTS

Firstly, I would like to thank my thesis director, Dr. Roda Roberts, for her guidance, generosity, and support in my academic endeavours. I am very grateful to her for giving me the opportunity to extend my knowledge and experience of bilingual lexicography by working on the Bilingual Canadian Dictionary Project.

I would also like to thank my former colleagues on the Dictionary Project for their encouragement along the way and my friend and current colleague Marie-Claude Charland for editing the Résumé.

I thank my family for the steady support and encouragement they have given me throughout this undertaking.

Finally, my sincere thanks go to my partner, Ken Hitchmough, who has further inspired my love and admiration by his understanding, his unfailing willingness to discuss my problems, and his apparently infinite supply of solutions.

ABSTRACT

Although Canada is an officially bilingual country, in which French is stated to be the mother tongue of almost 6.8 million people and English of 17.5 million people (2001 Canadian census), there is not a single general bilingual French-English dictionary on the market of Canadian origin.

This means that Canadians are forced to use European-produced bilingual dictionaries of English and French such as the *Collins-Robert* and the *Oxford-Hachette*, which, since they are intended principally for the European and U.S. markets, do not include many elements of English and French that are used in Canada. When these elements – these "Canadianisms" – are included in European dictionaries, their treatment is often unsatisfactory as they may be presented in an unsystematic, sometimes incomplete and even confusing manner.

The Bilingual Canadian Dictionary Project aims to remedy this situation by providing Canadians and more particularly Canadian writers, editors, translators and interpreters, as well as advanced second-language learners with a linguistic tool specially designed to meet their particular needs.

This thesis describes the Bilingual Canadian Dictionary Project and its methodology, as well as the Dictionary itself, its source materials, especially its electronic corpora, and the computer tools used in its compilation. The thesis then discusses the particular features that distinguish Canadian French and Canadian English and summarizes the significant events in the evolution of Canadian unilingual and bilingual lexicography. Finally, examples of regional usage from Canada, North America, France and Britain are presented and their entries in two European dictionaries analyzed and compared with their entries in the Bilingual Canadian Dictionary.

RÉSUMÉ

Bien que le Canada soit un pays officiellement bilingue où presque 6,8 millions de personnes se disent de langue maternelle française et 17,5 millions de langue maternelle anglaise (recensement canadien de 2001), il n'existe sur le marché aucun dictionnaire général bilingue anglais-français d'origine canadienne.

Cela signifie que les Canadiens sont obligés de se servir de dictionnaires bilingues anglais-français rédigés en Europe, tels que le *Collins-Robert* et le *Oxford-Hachette* lesquels, puisqu'ils sont destinés principalement aux marchés de l'Europe et des États-Unis, ne comprennent que peu d'éléments propres à l'anglais et au français du Canada. Lorsque ces éléments – ces « canadianismes » – figurent dans les dictionnaires européens, leur traitement s'y avère souvent insatisfaisant à cause d'une présentation peu systématique, parfois incomplète et même ambiguë.

Le projet du Dictionnaire canadien bilingue se donne comme objectif de remédier à cette situation en offrant aux Canadiens et plus particulièrement aux écrivains, aux réviseurs, aux traducteurs et aux interprètes canadiens, ainsi qu'aux étudiants de langue seconde de niveau avancé, un outil linguistique conçu spécialement pour répondre à leurs besoins particuliers.

Cette thèse a tout d'abord pour but de décrire le projet du Dictionnaire canadien bilingue et sa méthodologie, ainsi que le Dictionnaire lui-même, ses sources, en particulier son corpus électronique et les outils informatiques servant à sa compilation. La thèse propose ensuite une discussion des traits distincts du français canadien et de l'anglais canadien ainsi qu'un résumé des événements importants de l'évolution de la lexicographie unilingue et bilingue au Canada. En dernier lieu, il s'ensuit une présentation d'exemples des usages régionaux provenant du Canada, de l'Amérique du Nord, des États-Unis et de la Grande-Bretagne, une analyse des entrées correspondantes dans deux dictionnaires européens et une étude comparative de ces mêmes entrées dans le Dictionnaire canadien bilingue.

TABLE OF CONTENTS

LIST OF TABLES

INTRODUCTION

Since its launch in 1988, the Bilingual Canadian Dictionary (BCD) Project has inspired many studies[1], each focusing on a specific aspect of the work undertaken. As the Project is drawing to a close, it is now appropriate to present an overall evaluation. This thesis will examine the BCD Project in its entirety, providing answers to the three-part question of the title: *The Bilingual Canadian Dictionary: What? How? Why?*

Presentation of the Topic

The BCD Project has as its primary objective the preparation of the first comprehensive Canadian French-English English-French general dictionary. Containing between 40,000 and 50,000 lexical items, the BCD will be a large, single-volume dictionary aimed at translators, writers, editors, teachers and advanced second-language learners. The BCD will differ from other English-French dictionaries in that it will present the English and French languages as they are used in Canada. Canadian usage consists of two different types of linguistic items: those that are shared with other English- and French-speaking nations and those that are particular to Canada ("Canadianisms"). The BCD includes both types of items, making it, therefore, a dictionary of overall Canadian usage, rather than a mere dictionary of Canadianisms.

Specific Reasons for this Study

It is the coincidence of two major factors that has motivated me to undertake this study of the BCD:

a. As the BCD Project is nearing completion, now is an appropriate time to present a more global and up-to-date account of the Project, examining to what extent it has achieved its goal of presenting the English and French languages as they are used in Canada.

b. I consider that I am in a very good position to undertake this study as I have worked in bilingual English and French lexicography for virtually the whole of my professional career to date, working for three years as lexicographer and anglophone reviser on the BCD Project at the

---

[1] See the works given in the bibliography (page 133) by Clas, Cormier, Hamilton, Langlois, Roberts, Roberts and Bossé-Andrieu, Roberts and Megginson, and Roberts and Montgomery, all of which deal with different aspects of the Project.

Project centre at the University of Ottawa. Before moving to Canada, I had always worked on bilingual dictionaries in a unilingual environment, that is, in England and France. Working on an English-French dictionary in an officially bilingual country where both English and French have equal standing has greatly enriched my knowledge as a lexicographer and my involvement with the Project has proved to be an enjoyable and highly rewarding experience.

Objectives of this study

The objectives of this research project are

a. to present an account of the BCD Project, describing the product of this Project, the Dictionary itself, and its evolution in terms of its macrostructure, microstructure, source material and the computer tools used in its production, and to provide an analysis of a variety of BCD entries;

b. to show why there is a need for the BCD and

c. to show that the BCD provides the bilingual reference tool that Canadians need.

Thesis Outline

As the title of this thesis indicates, this study will answer the three questions: What? How? and Why? Chapters 1 and 2 provide answers to the first two questions by describing respectively *what* the Bilingual Canadian Dictionary Project is and *how* the Dictionary is produced. The first part of the answer to the question of *why* a bilingual Canadian dictionary is needed is provided in Chapter 3 through an examination of the particular features that distinguish Canadian French and Canadian English and a summary of the significant events in the history of Canadian unilingual and bilingual lexicography. The answer to this third question is completed in Chapters 4 and 5, firstly through the establishment of a framework of analysis that will be used to evaluate bilingual dictionary treatment of regionalism and secondly, through a comparative assessment of sample entries from the BCD and two other bilingual English-French dictionaries. Presented below is an outline of each chapter.

Chapter 1: What? What is the Bilingual Canadian Dictionary Project?, describes the BCD Project, the Dictionary itself and its evolution in terms of its macrostructure and microstructure.

This chapter explains that the BCD presents its material in a descriptive rather than a prescriptive way.

Chapter 2: How? Production of the Bilingual Canadian Dictionary, describes the methodology of the BCD and how this has evolved during the life of the Project. The source materials used, especially the corpora, as well as the different computer tools used on the Project are described.

Chapter 3: Why? The Need for the Bilingual Canadian Dictionary, first details the particular features that distinguish Canadian French and Canadian English and then, secondly, summarizes the significant events in the history of Canadian unilingual and bilingual lexicography.

Chapter 4: The Treatment of Regionalism in Dictionaries, firstly details the framework of analysis that will be used to evaluate bilingual dictionary treatment of regionalism. The criteria that make up the framework are presented and each of these criteria described with reference to both theoretical works and to dictionary prefaces and lexicographic policy documents.

Chapter 5: The Treatment of Regionalism in Dictionaries According to Bilingual Dictionary Entries, begins by presenting the five English and five French regional lexical items to be studied. Their presence or absence in unilingual and bilingual dictionaries is verified. The main part of the chapter consists of a comparative analysis of the treatment of the 10 sample lexical items in two European-produced bilingual dictionaries and in the BCD.

There is one appendix containing the dictionary entries themselves.

Methodology

In the preliminary chapters 1–3, the primary research method is descriptive. The description of the Project and the Dictionary is based on BCD documentation and personal experience. The last two chapters, which focus on one specific aspect of bilingual English and French lexicography, namely, the treatment of regionalism, are more analytical. A framework of analysis based on theoretical works, dictionary prefaces and policy documents is used to compare the treatment of regional usage in the entries of two bilingual English and French dictionaries published in Europe with the treatment of regional usage in BCD entries.

By presenting an overall description of the BCD Project and a comparative analysis of BCD entries, it is hoped that this study, one of a large number generated by the Project, will contribute to the development of expertise in bilingual lexicography in Canada.

1. WHAT? WHAT IS THE BILINGUAL CANADIAN DICTIONARY PROJECT?

In order to understand what the BCD Project is and what it involves, some historical and background details on the establishment of the BCD Project will be presented, followed by an examination of the macrostructure of the Dictionary itself and the microstructure of a typical BCD entry.

1.1 The Bilingual Canadian Dictionary Project: The Early Years

In a country where bilingualism is not only an official government policy but also a routine way of life for thousands of people, there is an obvious need for an up-to-date bilingual dictionary. The only such dictionary ever produced in Canada, Jean-Paul Vinay's *Canadian Dictionary* published in 1962, has been out of print as of the 1970s. In the mid-1980s, the idea of producing a new bilingual Canadian dictionary was floated and a period of committee discussion followed between various interested parties. In 1988, seed funding for a comparative lexicography project of French and English in Canada was awarded jointly by the universities of Ottawa and Montreal and the Social Sciences and Humanities Research Council of Canada (SSHRC) and the BCD Project was launched. The Project centre is at the University of Ottawa with two other centres, one at the University of Montreal and the other at Laval University in Quebec City. The Project Director since its launch and one of the two chief editors is Dr Roda P. Roberts at the University of Ottawa. The other chief editor is Dr André Clas at the University of Montreal.

Unlike most large-scale dictionary projects, the BCD Project is firmly based in the academic rather than the commercial world. The Dictionary was conceived from the first as a scholarly project that would provide interdisciplinary links between linguistic theory and practical lexicography. Thus, the research team set to work on the preliminary planning stages for the Dictionary, examining and expanding existing concepts and frameworks on dictionary organization put forward by scholars such as Al-Kasimi (1977: 109-112) and Wiegand (1984: 13-30).

A fundamental aspect of this planning was the identification of the users of the Dictionary: the BCD is aimed at translators and interpreters, writers and editors, teachers and advanced second-language learners whose mother tongue is either English or French. These groups were later canvassed in surveys to establish their specific needs and wants (Roberts, 1994).

Given the Dictionary's target audience and the specifically bilingual environment of Canada where at least a reading knowledge of the second language is desirable, the content of the BCD would be focused more on the written than the spoken language.

Since the intended dictionary users would consult the proposed dictionary primarily for the purposes of understanding or writing non-literary texts, the important decision was made during this early stage to collect the majority of the source material for the dictionary from journalistic rather than from literary sources. This decision followed the tradition established by Noah Webster who, in his *American Dictionary of the English Language* of 1828, was the first dictionary editor to move away from the normal lexicographic practice at the time of using only literature as dictionary source material. Webster took examples of usage from contemporary newspapers, journals and magazines (and was much criticized for his pains!). The BCD researchers realized that a journalistic corpus would provide far more valuable source material for its purposes than a literary corpus, as a single newspaper contains many articles on a variety of topical subjects, all written by different people in registers which, while somewhat different, do not stray very far from the neutral, whereas a novel or monograph has usually one author and therefore yields examples only of that single person's use of the language, and of registers that may fluctuate considerably depending on the characters and situations portrayed.

In order to be easily exploitable, the corpus had to be in electronic form. However, in the 1980s, the first electronic corpora for language analysis were just beginning to be compiled[2] and there were few electronic corpora of Canadian texts apart from the Strathy Corpus[3] and a few

---

[2] The first edition of the *Collins Cobuild English Dictionary*, published in 1987, was compiled using a 20-million-word corpus (*Collins Cobuild English Dictionary*, 1995: viii).

[3] Funded by a bequest from an alumnus, J.R. Strathy, the Strathy Language Unit was established in the English Department of Queen's University, Kingston, Ontario in 1981 with the aim of studying standard English usage. Today, the Strathy Corpus

other more limited ones such as the small private corpora maintained by some language professors.

The BCD team thus made the bold decision to establish its own electronic database of Canadian texts in English and French in order to provide authentic examples of usage. This was by no means an easy task in those early years as relatively few texts were available in electronic form. By selecting material from various CD-ROMs and from computer tapes of different newspapers, the BCD corpus of unilingual texts in English and French was compiled between 1988 and 1995. In order to adapt the material from these different sources to the specific needs of the BCD, a certain amount of streamlining was required. One example of this streamlining was the removal of the classified sections from the newspapers. It was decided that, as these "small ads" are made up almost entirely of abbreviations and rarely contain full sentences, they would not provide much useful material. Such tailoring was kept to a minimum, however, in order to reduce costs.

The selection of Canadian texts in French posed a special problem because of a specifically Canadian phenomenon, namely, the large number of French texts in existence that are translations from English. In Canada, it is often very difficult to ascertain whether a specific text was written originally in English or French or whether it is a translation. Efforts were made to identify and delete possible translations from the French unilingual sub-corpus, for example, by eliminating articles whose source was Reuters, although it is unlikely that all possible translations were deleted.

Translated texts, along with their originals, were nevertheless retained as part of a separate database. In fact, since translators and interpreters constitute one of the primary user groups targeted by the BCD, it was always intended that translated texts would provide a part of the source material for the Dictionary. Thus, the research team decided to acquire a corpus of aligned translated texts called TransBase. A detailed examination of these two BCD corpora is given in Chapter 2.

---

contains more than 57 million words of written and spoken Canadian English (May 2006 <http://post.queensu.ca/~strathy/topics/about_the_unit.html>).

Having started thus to constitute its corpus, the BCD team then needed a corpus-analysis tool that would allow query and manipulation of the data. The Department of Computer Science at the University of Waterloo had been working on a project to digitize the *Oxford English Dictionary* and put it onto CD-ROM and, as part of this work, had developed a tool for text analysis called Pattern Analyzer of Text or PAT, for short. As this text-searching system met the majority of the BCD's needs and in the interests of saving time and money by not developing the Project's own system, a decision was made to purchase PAT. Although this tool was not always especially user-friendly, the Project continued its use until it was superseded by a new version, which, at C$50,000, was prohibitively expensive and more sophisticated than the BCD required. As technical support for the older version of PAT was soon to be withdrawn, a solution had to be found. A computer software analyst and designer joined the BCD team and, in 2002, selected a freely downloadable Web-based search and concordance-generating software called Lucene to adapt to the needs of the BCD. This tool has proved extremely successful and its features are discussed in greater detail in Chapter 2, section 2.3.1.

As an essential part of this initial planning phase, the research team also studied a large number of existing bilingual English and French dictionaries, comparing their nomenclature, the different types of entries and different entry formats. This detailed analysis was designed to ensure that the BCD would improve on existing bilingual dictionaries both at the level of the macrostructure and the microstructure. It led to a number of important decisions. Firstly, the types of lexical items to be included in the BCD were identified. Because of the overall orientation of the Dictionary towards Canadian usage, the nomenclature of the BCD, unlike that of other bilingual dictionaries, would include a large number of Canadianisms, while excluding lexical items considered less pertinent in a Canadian context, for example, the names of African animals such as *aardvark* and *aardwolf*. In order to better meet the needs of the sophisticated users targeted, it was decided that the BCD would include more technical terms, multi-word items[4] and collocations[5] than many general bilingual dictionaries. Secondly, from the point of

[4] Multi-word items include orthographically separated compounds such as, for example, *accusé de réception* and *swing bridge*, as well as hyphenated compounds, such as, for example, *porte-malheur* and *tongue-tied*. Multi-word items also include fixed expressions, which are defined in the BCD Project as word combinations whose meaning cannot be deduced from the sum of their parts (e.g. *vivre d'industrie*), proverbs (e.g. *cœur qui soupire n'a pas ce qu'il desire*) and sayings (e.g. *nature abhors a vacuum*).

view of entry content, the decision was made to include more target-language equivalents and semantic and stylistic information than other general bilingual dictionaries, so as to provide a greater number of target-language choices and to facilitate their selection.

The huge scale and long-term nature of the Project meant that the preparation of dictionary entries had of necessity to begin as soon as possible and before all the required tools and structures were fine-tuned and perfected. Starting in a modest way in 1988, therefore, dictionary entries were prepared, although in simple fashion, using WordPerfect, a readily available word processing program. But the entry structure was not finalized until 1995, when the BCD team moved to using Standard Generalized Markup Language (SGML) for its entries.

1.2 Objectives of the Bilingual Canadian Dictionary Project

By 1993, the initial planning stages were complete and the Project was successful in being awarded longer-term funding by SSHRC. The four major objectives of the BCD Project had by then been fixed.

(i) to produce a Canadian English-French, French-English Dictionary.

The BCD Project has as its primary objective the preparation of a comprehensive Canadian French-English and English-French general dictionary. The targeted user group consists of two main categories: translators, interpreters, writers and editors on the one hand and, on the other, teachers and advanced second-language learners. The Dictionary is designed, therefore, to make a major contribution to the improvement of the quality of translation in Canada and of the communication skills of those studying English and French as a second language.

(ii) to establish an electronic corpus of Canadian texts in English and French

As Bowker and Pearson make clear (2002: 48), in order to draw statistically valid conclusions from the analysis of a corpus, researchers need to have adequate data samples at their disposal. Corpora used for lexicographical purposes need to be particularly large. Examples of large corpora are the British National Corpus, containing more than 100 million words (July 2006

[5] A collocation is defined in the BCD Project as a commonly recurring idiomatic word combination, for example, fâché contre, au mur and after all, to eat heartily, etc.

<http://www.natcorp.ox.ac.uk/>) and the Collins Cobuild Bank of English, which, at 524 million words (July 2006 < http://www.collins.co.uk/books.aspx?group=153>), is the largest corpus in existence in any language. To put these figures into perspective, an average-sized paperback book of 400 pages contains about 132,000 words. The authentic, quantifiable evidence of a large corpus ensures that no major sense is overlooked and provides reasonable justification for omitting certain senses. The BCD planned to establish a corpus of several million words in English and French. Given its Canadian orientation, its corpus would consist primarily of Canadian texts, but would also include some texts from the United States and from France, so that lexicographers could better identify what is specifically Canadian usage[6]. It was envisaged from the start that the corpus could be used in the future in the compilation of other Canadian dictionaries as well as to produce language-teaching materials.

(iii) to produce a multiple-purpose dictionary database

The BCD planned to set up a computerized dictionary database, which would serve several purposes. It would be used for the compilation and revision of entries and for the storage of data. This database would have the potential to become a national linguistic resource and make a major contribution to the future of bilingual lexicography in Canada. It could be used to compile future dictionaries, for example, a dictionary of Canadianisms proper or specialized field dictionaries. The database could also be used to supply data for linguistic research.

(iv) to develop Canadian expertise in bilingual lexicography

The Canadian Official Languages Acts of 1969 and1988 enshrined bilingualism as the official government policy and in so doing also guaranteed fertile ground for the flourishing of what has come to be known as the Canadian school of terminology. Much of current terminology practice was first developed in Canada in the 1970s and 80s and Canada has developed great expertise in this field. Canadian expertise in bilingual lexicography has, however, always lagged behind. The launch of the BCD Project in 1988 was intended to improve significantly the research environment in dictionary making in Canada and foster metalexicography in a variety of

---

[6] A collection of British texts in electronic format was not easily available for transfer to the BCD corpus at the time of its establishment. The Collins Cobuild Bank of English only became available by subscription in 1991 and the international edition of the British National Corpus only became available for worldwide distribution in 2001. So the BCD used the Internet as a general-purpose tool for checking examples of British regional usage.

different ways. The BCD Project centres at the universities of Ottawa, Montreal and Laval would provide three of a very small number of educational establishments and private organizations worldwide where practical lexicography training is available and where lexicography may be studied as a scholarly discipline. Most of the lexicographers would be students – both undergraduate and graduate – who would be trained in dictionary making through formal sessions and on-the-job training. A lexicographical aptitude test for potential student assistants was developed and the initial version of the *Methodology for Research Assistants*, containing an explicit set of instructions for dictionary compilers, was prepared.

1.3 Macrostructure of the Bilingual Canadian Dictionary

Before serious work could begin on the BCD, its editors had to establish a preliminary list of lexical items in French and English to be included in the dictionary and decide how these items would be presented. In establishing the nomenclature of their dictionaries, all lexicographers make a selection from the millions of words available and, as Al-Kasimi underlines, the editor should "select his entries in accordance with the purpose of the dictionary" (Al-Kasimi, 1977: 22). The purpose of the BCD is to meet the needs of the target user group, as detailed above, in the bilingual context of Canada. Thus, it is intended that bilingual individuals who use their first and second languages in their daily work will consult the BCD as both an active and a passive dictionary to encode and decode both English and French, predominantly in written form. In contrast with the majority of other general bilingual dictionaries, the BCD is aimed at a more homogeneous user group with a high overall standard of language skills and working in a professional or academic environment. When consulting a bilingual dictionary, not only do these sophisticated users hope to find more linguistic information than the average member of the general public, but they also have the ability to handle a more complex presentation of the material. As Howard Jackson notes (Jackson, 2002: 76), "coverage and accessibility" are two of the major concerns of dictionary editors and users alike and they are inevitably the two aspects that lead to some of the most extravagant claims in dictionary promotional material.

1.3.1. Selection of the Word List

The selection and organization of the word list of the BCD was determined, therefore, by the type, purpose and size of the planned dictionary. As indicated above, the BCD is a dictionary whose aim is to present the English and French languages as they are used in Canada. Canadian usage consists not only of linguistic elements exclusive to Canada, but also of those that are shared with other English- and French-speaking nations. The BCD covers both types of elements. Given the specifically Canadian orientation of the BCD, the following regional variants are not included in the word list:

- words and senses restricted to francophone countries other than France
- words and senses restricted to certain regions of France, the United Kingdom and the United States
- words and senses restricted to Commonwealth countries other than Canada and the United Kingdom.

Given the sophisticated user-group targeted (see section 1.1. above), the BCD includes more specialized vocabulary[7], more multi-word items and collocations than many other bilingual dictionaries of comparable size. Since the focus of the BCD is on the written rather than the spoken language, the number of colloquialisms, generally more typical of speech, is restricted. As the BCD aims to be a dictionary of contemporary usage, it includes few archaisms. As a linguistic dictionary, the BCD includes little information of an encyclopedic nature and only a limited number of proper names having particular importance in a Canadian context.

As was mentioned in the Introduction, it was planned that the Dictionary would include a total of between 40,000 and 50,000 lexical items, divided more or less equally between the French-English and English-French sections. It should be borne in mind, however, that as the

[7] It is very difficult to estimate the number of terms (specialized vocabulary) contained in a general dictionary because not all terms included in these dictionaries are identified by field labels (Josselin-Leray, 2005: 193). Studies have produced estimates of the number of terms included in general unilingual dictionaries varying from 25% to as high as 73% of the nomenclature (Landau, 2001: 241), depending on the size of the dictionary: the larger the nomenclature, the higher proportion of terms it usually contains (Béjoint, 1988: 360-361). It has been noted by metalexicographers that the number of terms included in dictionaries is increasing (Josselin-Leray, 2005: 198) and their number has been described as "a fair proportion" (Béjoint, 1988: 355), which may represent on average as much as "soit la moitié ou même plus de la moitié de la nomenclature du dictionnaire" (Clas, 2001: 232). No study has so far attempted to estimate the total number of terms (as opposed to the terms in one specific domain) in a general bilingual dictionary.

Dictionary is currently in the final stages of revision, these numbers may be liable to change, especially as some embedded items may be assigned headword status and vice versa.

With these overall principles to guide them, BCD researchers began the task of determining the specific lexical units to be included in the nomenclature. The initial word list was compiled from an analysis of a range of unilingual and bilingual dictionaries of similar size to the planned BCD. The principal unilingual dictionaries used to establish the French nomenclature were *Le Petit Robert* (1987) and the *Lexis* (1987) and, for the English nomenclature, *Webster's Ninth New Collegiate Dictionary* (1987) and later the *Random House Webster's College Dictionary* (1992). To complement these and in order to cover specifically Canadian usage, the principal dictionaries used were, for the French nomenclature, the *Dictionnaire du français plus* (1988) and later the *Dictionnaire québécois d'aujourd'hui* (1992) and, for the English nomenclature, the *Gage Canadian Dictionary* (1983) and the *Penguin Canadian Dictionary* (1990). Two bilingual dictionaries of comparable size to the planned BCD, the *Oxford-Hachette French Dictionary* and the *Robert-Collins*, were also consulted during the establishment of the nomenclature.

Corpora were also used to determine the lexical items to be treated in the BCD. During establishment of the initial nomenclature, because the BCD's own bilingual comparable corpus was not yet fully assembled (see section 1.1 above), the frequency list of words found in the parallel corpus, TransBase, was used to confirm the choice of items to be included, although given the nature of the texts it contains (the edited transcriptions of proceedings in the Canadian House of Commons), this source had to be used prudently. Modifications to this initial nomenclature were made during entry preparation on the basis of analysis of the BCD corpus TEXTUM and especially on the basis of the frequency of occurrence of lexical units in that corpus.

1.3.2. Organization of the Word List

Not all the lexical items selected for inclusion in the dictionary would have headword status, as the planned first edition of the BCD would be in one volume and size was a factor in determining the number of entries. Headword status is assigned, therefore, to:

- simple lexical items (for example, *cambriolage, femme, eject, payable*, etc.)
- agglutinated compounds written as single words (for example, *contredire, schoolboy*, etc.)
- abbreviations (for example, *AB* for *Alberta, dep* for *depanneur*, etc.) and
- acronyms (for example, *cégep, NATO*, etc.)

Multi-word items such as orthographically separated compounds (for example, *accusé de réception*), hyphenated compounds (for example, *tongue-tied*) and fixed expressions (for example, *to beat around the bush* and *le silence est d'or*) are not generally given headword status, but are embedded within the appropriate entry.

Homonyms are not treated as separate entries: a single form is given a single entry, that is, it is treated as polysemous. While this is contrary to the conventional technique used in unilingual dictionaries, which attribute separate headword status to homonyms and differentiate them with a superscript numeral, bilingual dictionaries have not on the whole attempted to distinguish polysemy and homonymy. In the BCD, therefore, homonyms such as, for example, the English word *date* (meaning both the day of the month and the fruit of the date palm), are included as separate senses under one single headword.

As a general language dictionary, the BCD follows the conventional lexicographic practice of listing headwords in alphabetical order, contrary to many specialized dictionaries which opt for some kind of systematic or thematic ordering of their word lists (Bowker, 2003: 157-158).

The BCD lexicographic database permits modifications and additions to be made to the nomenclature in the light of subsequent information. Any material removed from the word list for the first edition will be retained in the database for possible future use.

1.4 Microstructure of an Entry in the Bilingual Canadian Dictionary

As has already been mentioned above, the scholarly nature of the BCD Project aims to provide links between linguistic theory and lexicographic practice. The format of individual BCD entries was designed, therefore, in line with what Hausmann and Wiegand (1989: 348-349)

call the "complete obligatory microstructure", and a four-part entry was fixed on as being the clearest and most flexible format. These four parts are (i) the introductory zone; (ii) the semantic divisions; (iii) the fixed expression section; (iv) the compound section. The total set of possible items in each zone may be presented as an abstract microstructure in linear form:

Introductory Zone

headword > (spelling variants) > (feminine form for French adjectives and for French and English nouns; irregular plural form for French adjectives and nouns and English nouns) > grammatical category > (irregular verb forms)

Sense division

Number of division > (register label) > (geographical label) > (field label) > (currency label) > (commentary label) > (preposition and other remarks) > (sense) > (actants) > equivalents > (preposition and other remarks – TL) > (free combinations – SL + TL) (collocations – SL + TL) > (collocation cross-references)

Fixed expression section

(exp) > (SL + TL) > (cross-references)

Compound section

(cmp1) > (SL when the first element is the headword + TL) > (cross-references)

(cmp2) > (SL when the first element is not the headword + TL) > (cross-references)

The elements not in parentheses are those that will be present in every entry that is not simply a cross-reference. Those that are in parentheses will not necessarily be found in every entry. This initial microstructure has been developed and expanded over the years to become more complex, but the basic structure remains the same. The following section illustrates each part of the four-part entry with examples in French and English taken from the BCD.

The first part of the entry, the introductory zone, begins with the lemma, that is, the uninflected form of the headword, plus any spelling variants, for example:

bleuet *nm or* **bluet**

The headword is always followed by the grammatical category, for example:

dietitian *n or* **dietician**

rapidement *adv*

There then follow the feminine forms for French adjectives and for French and English nouns, for example:

paresseux *adj, f* **paresseuse**

boursier *nm, f* **boursière**

chairman *nm, f* **chairwoman**

There then follows the irregular plural form for French adjectives and for French and English nouns, for example:

national *adj, pl* **nationaux**

feu *nm, pl* **feux**

phenomenon *n, pl* **phenomena**

There then follow the irregular verb forms of any English verb that does not add the simple ending of -s, -ed or -ing. The third person singular of the present and past tenses and the past and present participles are given, for example:

dance *vtr* **(dances, danced, danced, dancing)**

The second part of the four-part entry is the sense division. Each sense division is numbered and the numeral may be followed by one or more usage labels indicating register, geographic origin, specific field of activity and currency, if the use of a lexical item in a given sense is restricted in any way.

The two register labels used in the BCD are *formal* and *informal* in the English-French part of the Dictionary and *soutenu* and *familier* in the French-English section.

The five geographic labels used in the BCD[8] are:

CD Canada

FR France

GB Great Britain

US United States

NA North America

[8] Geographic labelling in the BCD is presented in greater detail in Chapter 4, section 4.2.2.3.

A large number of field labels, for example, *Admin*, *Biol*, *Mus*, *Sport*, are available for use in the BCD, each label indicating that a word or sense is current only or especially in a particular field of activity.

Currency labels are a type of usage marker used to limit the use of a given word, sense or equivalent to a certain temporal period. However, since the BCD reflects primarily modern usage and omits most archaic usage, currency labels are not generally used. There are two principal exceptions to this rule and, in both cases, the currency label is combined with a geographical label:

- Words or senses that are current in Canadian French may be archaic or out of date in France. In that case, the temporal label *vieux* or *vieilli* is combined with the geographical label, for example: **tapon** *nm (CD) (FR: vieilli)*. This combination of labels indicates implicitly that the word or sense is current in Canadian French.
- Since the BCD is giving particular attention to Canadian usage, it will include some historical Canadianisms. What is meant in the BCD Project by a historical Canadianism is either a term denoting a Canadian reality that longer exists, such as *wintering partner* (a partner or officer in the Hudson's Bay Company stationed at trading posts), or older words that, although they represent Canadian realities that still exist, are no longer used, such as *Brûlé* (a 19th-century term for a mixed-blood Indian or a Métis). In these cases, the historical label # is combined with the geographical label as *(CD#)*, for example: **Brûlé** n . . . **3** *(CD#)* (half-breed, Métis).

The following commentary labels are used in the BCD to indicate a social restriction of usage: iro (ironic)
pej (pejorative)
vulg (vulgar)
neg[9] (negative)
pos (positive).

[9] The commentary labels *neg* and *pos* usually occur together in the BCD, being attached to contrasting equivalents that express fine semantic nuances within the same sense division. The English adjective *forward*, for example, is labelled *often neg* in the sense of "presumptuous" and *pos* in the sense of "forthright", both within the same sense division for the general meaning of "bold".

At this point in the microstructure of the BCD entry, there then follows information on prepositions collocating with the headword, if any, for example: **abuser** *vi* (**de**).

The next element of the BCD entry is the sense indication. In the BCD, senses are not only clearly separated and numbered, but sense indications are normally provided for every sense of every headword. The purpose of the sense indications in the BCD is not to fully define the word (that is, to distinguish it from all other words), as unilingual dictionaries do. Rather, it is to allow the user to pinpoint more exactly and more quickly the exact equivalent he needs. At the entry *bleuet*, for example, the user can choose between *bleuet* the flower, *bleuet* the fruit, or *bleuet* the inhabitant of the Saguenay-Lac-Saint-Jean region. Sense indications are given in French for a French headword and in English for an English headword and can take several forms: a short definition, a synonym, a related word (often a generic), etc.

Short definition or explanation:

motoneige *nm* (*petit véhicule pour circuler sur la neige*)

Synonym:

admettre *vt* . . . **3** *(tolérer)*

abduct *vt* **1** (*kidnap*)

Related word:

bleuet *nm* **1** *(fruit, plante)* . . . **2** *(fleur)* . . .

maple *n* **1** (*tree, wood, flavour*)

The next section in this second zone of the four-part entry contains one or more examples of usage of the headword with their translations. In the BCD methodology, these illustrative examples are called free combinations in order to distinguish them from the more fixed structure of collocations, fixed expressions and compounds:

besoin *nm* **1** . . . ***les besoins des étudiants ont changé** the needs of students have changed.

There then follows the section of the entry containing those lexical items that have been identified as collocations by the lexicographic methodology of the BCD, along with their translations and any cross-references:

dismiss *vt* ****to dismiss a claim** rejeter une demande, rejeter une prétention, rejeter une revendication, (*Droit*) rejeter une requête, rejeter une plainte, nier une allégation, démentir une allégation, contester une allégation.

The third zone of the four-part entry contains those lexical items identified as fixed expressions by the lexicographic methodology of the BCD, along with their translations and any cross-references:

audacieux, *nm, f* **audacieuse** . . . **(exp) la fortune sourit aux audacieux** fortune favours the brave.

The final zone of the four-part entry contains those lexical items that have been identified as compounds by the lexicographic methodology of the BCD, along with their translations and cross-references. This zone is further subdivided into two sections. Most compounds are placed in the first section, named **(cmp1)**, containing compounds whose first element is the headword, for example:

air *n* . . . **(cmp1) air conditioning** *n.*

However, in some compounds the first element is semantically "weak". These compounds are, therefore, placed in the entry for the strong element. The compound *bel âge*, for example, is placed at the headword *âge*, but in a second section, named **(cmp2)**, thus differentiating it from compounds such as *âge adulte*, which are placed in the **(cmp1)** section:

âge *nm* . . . **(cmp1) âge adulte** *nm* . . . **âge légal** *nm* . . .
(cmp2) bel âge *nm* . . . **grand âge** *nm* . . .

In order to show how some of the elements discussed above fit together, presented below are two complete entries from the BCD, the first a French Canadianism and the second an English Canadianism:

tapon *nm* 1 *(boule de qch) (CD) (FR: vieilli, familier)* wad, ball.
* **un tuyau bloqué par un tapon de cheveux** a pipe blocked by a wad *etc.* of hair.
des tapons de poussière dust balls, dust bunnies *(familier).*
2 *(tas de choses semblables, grande quantité) (CD)* heap, pile, bunch *(familier).*
***un tapon de vêtements sales** a heap *etc.* of dirty clothes.

des tapons de nuages noirs masses of dark clouds.
3 *(morceau, masse) (CD)* lump; *(neige, terre)* clump; *(butter)* lump; *(dans une sauce)* lump.

(exp) en tapon *(CD) (FR: vieilli)* balled up, bunched up, bundled up, in a heap.
***des vêtements** *laissés* **en tapon sur une chaise** clothes left in a heap on a chair.

tuque *n (CD) or* **toque** *(CD)* 1 *(knitted cap)* tuque *f (CD)*, bonnet *m.*
* **he pulled on a tuque** il a mis une tuque.
a Montreal Canadiens' tuque une tuque des Canadiens de Montréal.

1.5 Conclusion

This chapter has answered the question "What is the Bilingual Canadian Dictionary?" by first giving some historical and background details of the establishment of the BCD Project and explaining its scholarly nature and its four main objectives. The principal product of the Project, the Dictionary itself, has been described in terms of both its macrostructure and its microstructure. Once the macrostructure and microstructure of the BCD had been established, work on entry preparation commenced in earnest. The following chapter describes how this work was carried out.

2. HOW? PRODUCTION OF THE BILINGUAL CANADIAN DICTIONARY

This second chapter explains the methodology of the BCD, firstly, in terms of the source material and computer tools used and, secondly, in terms of the preparation of a Dictionary entry as it moves through the different production stages, from compilation through revision to its placement in the final sequence of all Dictionary entries.

2.1 Methodology

At the end of the 1980s, as the BCD Project was getting underway, the vast majority of dictionaries produced were entirely commercial undertakings rather than scholarly achievements. Dictionary making was viewed as an essentially practical activity, an art or a craft, and there was little evidence in published dictionaries of any close relation between linguistic theory and lexicographical practice.

The BCD Project, on the other hand, is an undertaking of "scientific lexicography" (Hausmann, 1989), the methodology of which is based on principles derived from the disciplines of theoretical lexicography, lexicology and English and French linguistics and from the results of its own research in dictionary use and dictionary criticism. As a scholarly rather than a commercial undertaking, the BCD Project has benefited from having at its disposal resources that have allowed the continual reappraisal of this methodology in the light of specific problems encountered at each stage of the preparation and revision of dictionary entries. The methodology has been modified in the light of this frequent re-evaluation and these modifications have, in their turn, provided feedback for the reinterpretation and reworking of the theoretical aspects of lexicography on which the methodology is based. Characterized, therefore, by a constant interaction between theoretical aspects and practical applications, the BCD Project seeks to bridge the gap that has existed for so long between the theory of lexicography and its practice. It is envisaged that through this process the BCD Project will make a major contribution to a methodology of bilingual dictionary making that could become "as widely known and accepted

as that of bilingual terminology" (Roberts, "Methods of bilingual dictionary making: the Canadian experience", 1992: 113).

The initial methodology of the BCD was developed in a collaborative process between faculty members in the three universities of Ottawa, Montreal and Laval. An excellent starting point for the BCD methodology was provided by Zgusta (1971: 223), who identified four stages of dictionary making:

(1) the collection of material;
(2) the selection of entries;
(3) the construction of entries; and
(4) the arrangement of the entries.

These four stages that Zgusta termed the "articulation of work" (ibid.) were expanded on by the BCD Project and detail added to give a process of ten stages (Roberts, "Methods of bilingual dictionary making: the Canadian experience", 1992: 95-96):

(1) selection of headwords
(2) collection of material
(3) analysis of material and selection of data for individual entries
(4) construction of individual entries
(5) preliminary overall revision of individual entries
(6) revision of source-language material in individual entries
(7) revision of target-language material in individual entries
(8) consultation of specialists for a number of entries or parts of entries
(9) arrangement of entries in the dictionary
(10) final editing of the dictionary.

There are two reasons why the BCD methodology reverses the first two stages of Zgusta's suggested procedure. The first reason is that, in the BCD, the kinds of texts collected as source material would depend to a great extent on the type of headwords selected. Since, for example, it was decided to exclude most archaic usage from the BCD (see Chapter 1, section 1.4), it was not necessary to gather any archaic source material. The second reason is that

Zgusta's *Manual* was designed to be "particularly useful to the development of lexicography in Africa and Asia" (1971: 18) where the standard national languages may not be well documented in lexicographical form and where the only way to establish a list of potential headwords would be by first gathering textual material from which they could be extracted. This is obviously not the case in English-French lexicography.

2.2 Source Material

Dictionary making requires the use of two types of documentation, lexicographic material (dictionaries) and non-lexicographic material (texts).

2.2.1. Lexicographic Material

A whole library of unilingual and bilingual English and French dictionaries (Roberts, 1997: 7-14) has been assembled since the launch of the BCD Project, comprising not only general language dictionaries, but also, among others, specialized technical dictionaries, combinatorial dictionaries and dictionaries of fixed expressions and proverbs. BCD researchers have access to the most up-to-date editions of these dictionaries, both in hard copy and on CD-ROM, permanently loaded on their computers. The research team also has unlimited access to the two great Canadian term banks, *Termium* and the *Grand dictionnaire terminologique.*

However, despite the fact that the Project is preparing a Canadian dictionary, many, if not most, of the dictionaries it uses are European or American. This is especially true of general dictionaries, since, as indicated in Chapter 1, until relatively recently there was no good general dictionary available covering Canadian usage in its entirety. In the 1980s, when the Project was launched, Canadian usage was documented only in either dictionaries of Canadianisms in the strict sense of the word, such as Dulong's *Dictionnaire des canadianismes* (1989), or in dictionaries of regional Canadianisms such as *the Dictionary of Newfoundland English* (1982) and the *Dictionary of Prince Edward Island English* (1988). Other dictionaries claiming to present Canadian usage, such as the *Dictionnaire du français plus* (1988), the *Dictionary of Canadianisms on Historical Principles* (1967) and *Funk & Wagnall's Canadian College Dictionary* (1989), were originally partially Canadianized versions of existing French and American dictionaries.

Of these Canadianized versions of American English dictionaries, the *Gage Canadian Dictionary* (1984) was held for many years to offer the most comprehensive and authoritative treatment of Canadian English. This dictionary was one in a range entitled the *Dictionary of Canadian English* series of which the first appeared in 1962. Based on the Thorndike-Barnhart American dictionaries and Canadianized for spelling and pronunciation, these dictionaries aimed to present not only Canadianisms, but Canadian English as a whole. However, even the revised and expanded edition of the *Gage Canadian Dictionary*, published in 1997, does not cover as much material as most collegiate American dictionaries.

Because it was produced entirely in Canada and was not an adaptation of an American dictionary, the *Penguin Canadian Dictionary* was a significant arrival on the Canadian English dictionary scene in 1990. However, given the fact that it is a concise dictionary intended for high-school students, it offers only a partial coverage of Canadian English.

It was only in 1998, therefore, with the publication of the *Canadian Oxford Dictionary*, that a more extensive coverage of Canadian English usage was eventually provided. With 130,000 entries, of which 2,000 are Canadian words or senses, the *Canadian Oxford* is accepted as the most thorough and up-to-date account of Canadian English recorded in dictionary form. A second revised edition was published in 2005.

The first French dictionary conceived specifically for a Canadian audience, the *Dictionnaire québécois d'aujourd'hui*, was published a few years earlier in 1992. This dictionary uses Canadian French as the language of reference and includes in its 40,000 entries a large number of Canadian words, senses and expressions.

More details of the evolution of French and English lexicography in Canada are given in section 3.4. below.

2.2.2. Non-lexicographic Material

Because of this initial lack of any exhaustive reference work treating either Canadian English or Canadian French, the linguistic evidence provided by a corpus of authentic Canadian

texts would be absolutely essential to the dictionary-making process. But although it was relatively easy for the Project to obtain lexicographic material – albeit very little covering Canadian usage – gathering the non-lexicographic material, the texts that would constitute the BCD corpus, was, however, not such an easy task. In 1988, comparatively few texts were available in computerized form and it took a considerable amount of time end effort to gather material in sufficient quantity and of the appropriate text types and tailor them to the specific needs of the BCD (see section 2.2.2.1.1 below).[10]

Size is a very important factor when judging whether a corpus is truly representative of the way language is used. The first major computerized corpus, the Brown Corpus of American English, contained just over one million words and, although this was considered large enough to be representative at the time of its compilation in 1963–64, it has since been shown that many common words and phrases may occur rarely or not at all in a corpus of that size. It is now generally agreed that in order to be representative enough to cover the broad spectrum of language required of a general dictionary, a corpus intended for lexicographic research needs to be much larger than one used for purely linguistic research, such as for a syntactic study, for example. As Sinclair explains, "For any recurrent pattern to be identified as a unit of meaning there must be a sufficient number of occurrences for the regularity to be observable through the variation" (Sinclair, "Corpus processing", 2003: 185). For dictionary purposes, a corpus of several hundred million words is now considered essential. Whereas the first edition of the *Collins Cobuild English Dictionary*, published in 1987, was compiled using a corpus of more than 20 million words (*Collins Cobuild English Dictionary*, Introduction, p. viii), the latest 2001 edition has been revised using the Bank of English, a very large corpus of more than 500 million words (www.collins.co.uk 10 December 2004). The total number of words available to BCD researchers in the two databases the Project uses runs to more than 600 million words. A detailed breakdown of BCD corpora is given in section 2.2.2.1.1 and 2.2.2.1.2 below.

One of the greatest advantages of using a large electronic corpus in dictionary making is that lexicographers can see many examples of the lexical item they are researching in a very

[10] Work on the Project had to proceed, nonetheless, so during this interim period, BCD researchers used illustrative examples from different unilingual dictionaries, from certain key-word-in-context (KWIC) lists and from their own general reading for analysis of lexical items.

wide variety of different contexts. A larger corpus provides a larger number of examples to choose from, providing authoritative evidence of the most frequent language patterns of naturally occurring running text and thus enabling lexicographers to make decisions about the meanings, uses and patterns of words with much greater confidence. Accurate detail can be extracted from the corpus about the collocates of a query term – the particular words that are commonly used with the headword – and the fixed expressions and compounds in which the headword occurs. A large corpus also provides more instances of the less common words so that their senses can be clearly identified. In addition, a large corpus allows for easier identification of the pragmatic uses of language, those instances where language is used for a specific purpose, such as to emphasize, to express emotion, to insult or to invite. These uses often only become obvious when a large number of examples can be viewed together. In all these aspects, a dictionary produced using a large corpus will offer a far higher degree of reliability.

As the BCD is produced using large databases in both English and French, this same level of reliability is achieved for both source- and target-language items. The Project has successfully integrated corpus use into all stages of bilingual dictionary making and the resulting Dictionary entries are undoubtedly enriched by this resource.

The main problem when working with a large corpus is making sensible and balanced selections from the vast quantity of data available. While the use of a large electronic corpus greatly increases the reliability of the material presented in a dictionary, it inevitably also increases the length of time taken in entry preparation and revision.

2.2.2.1 The BCD Corpora

Researchers on the BCD Project have access to two computerized databases, a bilingual comparable corpus and a parallel corpus.

2.2.2.1.1 The Bilingual Comparable Corpus

One of the four major objectives of the BCD Project (see Chapter 1, section 1.2) has been fulfilled in the establishment of an electronic corpus of Canadian texts in English and French. Originally set up on a Unix computer at the University of Montreal, this bilingual comparable

corpus is called TEXTUM, a title formed from "text" and the initial letters of "University" and "Montreal". The database was later transferred to the University of Ottawa and is accessed via the BCD network that unites the three research centres at Ottawa, Montreal and Laval universities.

The majority of the texts in the database are Canadian but, as was mentioned in Chapter 1, texts from the United States and France are also included so that Canadianisms and Canadian usage can be verified by checking their absence in non-Canadian texts. The database consists of over 310 million words distributed over 11 sub-corpora: five French, giving a total of around 100 million words and six English, totalling 210 million words. Three of the French sub-corpora are Canadian, the other two being from France, and four of the English corpora are Canadian, the two others being from the United States[11]. TEXTUM is made up predominantly of newspaper and magazine articles, but also contains academic texts, government documents and fiction. The texts cover a wide range of subjects and although most were written for the general public, rather than for a specialist readership, some scientific and technical material is also included. As the texts were produced between 1987 and 1995, the corpus presents a synchronic snapshot of the English and French languages rather than a historical, diachronic record. The BCD bilingual comparable corpus is not lemmatized or annotated, for practical reasons related to the time and effort this would require. Presented below is a table giving details of the corpus.

[11] See also footnote 6, page 10.

Table 1

TEXTUM, the BCD Bilingual Comparable Corpus

ENGLISH	SIZE (in millions of words)	FRENCH	SIZE (in millions of words)
English Canadian Press (N+M, G, CD)	129.0	Presse canadienne française (N+M, G, CD)	77.0
Wall Street Journal (N, G+ST, US)	41.8	Le Monde (N, G, FR)	17.1
Gazette (N, G, CD)	6.7	Ouest-France (N, G, FR)	4.9
Queen's (N+M+F, G, CD)	5.0	Leméac (F, G, CD)	0.9
Department of Energy (GD, ST, US)	27.2	ACFAS (M, ST, CD)	0.13
Canadian Geographic (M, G, CD)	0.3		
TOTAL	210.0		100.03

Key: N = newspaper M = magazine, journal ST = scientific/technical

F = Fiction G = general GD=government documents

CD = Canadian US = American FR = France

Working with a "closed" or "finite" corpus of this kind has inevitably meant that gaps have become evident as the Project has progressed. This is particularly obvious in the area of neologism in technical fields. For example, the BCD corpus contains only a handful of

references to wireless technology, a field that has developed very rapidly over the last couple of years and the terminology of which lexicographers would definitely wish to consider for inclusion in the Dictionary. This lack is, however, far from being the problem it appears to be initially, as any lack of material in the corpus can be supplemented from a variety of non-lexicographic sources, such as the official Web sites of a multitude of Canadian federal and provincial organizations and government departments. The World Wide Web is a huge source of information, a vast collection of texts in electronic format from which ad hoc collections of documents can be compiled to provide linguistic information for particular purposes (Varantola, 2003 and Zanettin, 2002). These collections of Internet documents have been described as "DIY corpora" (Zanettin, 2002: 241) and can be disposed of when they have served their purpose. Additional linguistic information is also supplied in the BCD Project from other non-lexicographic sources such as the on-line and paper copies of a wide range of newspapers, magazines and specialist journals and the personal expertise of the BCD Project's team of experienced lexicographers and language and subject-field specialists.

2.2.2.1.2 The Parallel Corpus

The vast majority of the linguistic analysis carried out by BCD researchers is conducted on the bilingual comparable corpus, but the BCD Project also uses TransBase, a parallel corpus comprised of the edited transcription of debates in the Canadian House of Commons and consisting of source texts aligned with their translations. This type of corpus is sometimes referred to as bitext (Bowker, 2002: 45), a term said to have been coined by Brian Harris (1988). Most of the "millions of sentences that have been translated between English and French" (Introduction to TransBase, July 2006 <http://www.tsrali.com/help.cgi?topic=FAQ-what&UTLanguage=en&userName=>) in TransBase are translated from English to French, but some are translated from French to English.

A version of TransBase was made available to BCD researchers in 1995 under licence from the Centre for Information Technologies Innovation (CITI), an agency of Industry Canada. That initial version of TransBase consisted of three years of Hansard and contained almost a million aligned segments totalling 45.7 million words, of which 21.6 million were in English and 24.1 million in French. This version, although limited in the number of words and with a less

user-friendly concordancer, specified the source language of each segment. The BCD Project now uses an updated version of TransBase, which has been available as an on-line subscription service since 2001. This version of the database contains Hansard reports dating from April 1986 to the present and totalling 235 million words. It also contains documents drawn from the collected decisions of the Supreme Court of Canada, the Federal Court of Canada and the Tax Court of Canada. The documents from these sources also date from 1986 to the present and total 88 millions words. This Web-based version of TransBase, although much larger in size, is consulted using a much simpler concordancer which does not allow for specification of the source and target languages. Since December 2003, the database has been administered by Terminotix Inc., an Ottawa-based company specializing in computer tools for translation (<http://www.terminotix.com>). TransBase is not annotated with part-of-speech information.

2.3 Computer Tools

Storing and retrieving the large quantities of corpus and lexicographic data used and generated by the BCD Project requires a range of computer tools. This section gives details of the computer tools used and how they have been developed and fine-tuned to meet the evolving needs of the Project.

2.3.1. Searching the Bilingual Comparable Corpus

As was detailed in the previous chapter (Chapter 1, section 1.1), during the early years of the BCD Project, TEXTUM was queried using the concordance-generating program PAT (Pattern Analyzer of Text), which was well suited to lexicographic purposes. Using the KWIC[12] index system, PAT could search for significant patterns occurring after (but not before) the queried word. Concordances were displayed alphabetically on the basis of the word following the queried word. PAT retrieved exact matches, but wild cards were available to find all contexts containing a given string. It was possible to perform proximity searches, use Boolean operators and intersect the results of two queries.

[12] KWIC stands for "keyword in context" and describes a computer indexing system that generates and displays lines of text (concordances) in which the search term occurs, and presents them in an easy-to-read format.

In 2002, the BCD Project adopted a different corpus-analysis tool, a Web-based search and concordance-generating software called Lucene. The basic version of the software, written in Java, can be downloaded free of charge from the Jakarta Apache site (<http://jakarta.apache.org/lucene>). The version of Lucene used on the BCD has been tailored to the specific needs of the Project and permits more sophisticated query patterns than did the original search tool. A basic search of the BCD corpora using Lucene allows users to find occurrences of the query term and displays the KWIC concordances in which the query term appears in the centre of each line of surrounding context. The user can specify the number of characters that precede and follow the query term and the number of concordances displayed on each page. The query term is displayed as a live link which, when selected, opens the whole article containing the query term, thus displaying a maximum amount of context for researchers to examine. The normal default search is weighted, giving priority to articles where the search term appears more often. This helps to focus on examples of the search term in fewer, but perhaps more specific contexts. There is also a "distributed search" option available that allows the user to specify a search that is evenly distributed through the database, thus providing a greater selection of articles.

Lucene will find exact matches of a single word or of a phrase composed of two or more words (e.g. "coffee bar", "tasse de café", etc.). Searches can be expanded considerably using the question mark and the asterisk as wild cards. These wild cards may be placed at the end or in the middle of a search string. The question mark may represent one character (?) or a specified number of characters, for example, two (??), three (???), four (????), etc. The asterisk is used to represent zero or more characters. So, while the query term "ben?" will produce all occurrences in the corpus of "ben" plus one other character, for example, "bend", "bent", "Benz", "ben*" will return concordances of "ben" plus one or more characters, so "bend", "bent", "Benz" will be returned in addition to all the other occurrences of this string in the corpus, from "bench" through "benedictions" to "benzodiazepine". Sophisticated searches combining both wild cards can save the lexicographer a great deal of time. For instance, to examine all possible forms of the verb *se démener*, a researcher could use the search pattern "dém?n* ". In addition to the verb forms "démenaient", "démené", "démène", "démènerait", etc. that are of interest to the researcher, this query would also produce unwanted material such as "démon", "démons",

"démonstration", "démontrer", etc., but, since Lucene also permits the use of Boolean operators, this "noise" could be eliminated by adding "NOT démon*" to the query string. In this way, the lexicographer can exclude unwanted material and obtain a very useful set of returns with one single search.

Lucene is also capable of performing a "fuzzy" search which finds words containing groups of letters that are similar, but not identical, to the query term, The insertion of a tilde after the query term, for example, "roam~", will retrieve fuzzy matches such as "roams", "road", "foam", etc.

In addition, Lucene allows the user to carry out a proximity search by using the tilde followed by a numeral to specify the number of words between the search items. For example, "close contact"~4 will find the term "close contact" itself plus, in addition, those two words, in either order, with one, two, three or four words between them.

Lucene also allows a frequency query to be performed. This produces a list of all the words that follow (but not precede) the query term, arranged in order of decreasing frequency. A frequency query for "crash" in the *English Canadian Press* section of the comparable corpus takes several minutes to perform as the search tool scans all 129 million words. As no frequency threshold can be specified, every single differing word that follows the search term in the corpus is listed, along with the number of times that each particular combination occurs. It is not possible to vary the parameters of a frequency query. Thus it is possible, for example, to identify two-word compounds in which *crash* is the first element (e.g. *crash course*), but not those in which *crash* is the second element (e.g. *car crash*). The frequency query in English does allow the identification of an intransitive verb collocating with the query term (e.g. *crash occurred*). Overall, the more sophisticated querying techniques described above produce more useful information than the frequency query. In French, the frequency query produces less useful results as the first word following the query term may often be a preposition (e.g. *accident de voiture*) or a pronoun e.g. *accident se produit*. Presented on the next page are the first 18 lines of the word-frequency list produced for the query "crash" in the *English Canadian Press*.

Table 2

Word-frequency List for the Query "crash" in the *English Canadian Press* section of the BCD Corpus

Following Word	Number of Occurrences
in	364
of	275
and	243
the	223
that	190
on	173
site	160
was	144
at	119
course	118
Vegas	92
occurred	90
a	89
test	73
near	72
into	67
but	60
landing	58

In this case, the first six words that most frequently follow the search item are all function words. The first two nouns to co-occur with the search item are "site" with 160 occurrences and "course" with 118 occurrences. Using this list as a starting point, lexicographers make further investigations to determine whether the lexical items *crash site* and *crash course* are to be included in the Dictionary and, if so, whether they are to be included as a compound or collocation or in a free combination.

2.3.2. Searching the Parallel Corpus

TransBase is queried using a bilingual concordancer (also known as a bi-concordancer) called TransSearch, a corpus-analysis tool developed at the University of Montreal by the computer-assisted translation group in the RALI (Recherche appliquée en linguistique informatique) laboratory (<http://rali.iro.umontreal.ca>). The concordancer described here is the one used with the Web-based version of TransBase. A TransSearch query can be a word or a phrase in either English or French. By default, TransSearch looks for the given expression without regard for language, but it is possible to specify the language to distinguish, for example, between the French and English words *tape*. The system retrieves occurrences of the query term in context along with the translation of each occurrence and displays the results as pairs of aligned sentences. The user can change the default display of 10 pairs of sentences per page to 5 or 25 and can view longer contexts. By adding the operator "+" to a search term, the results are lemmatized so that the query string "fly+", for example, will find and display all occurrences of all possible forms of *fly* (e.g. *flew*, *flies*, *flying*, etc.). TransSearch uses the ellipsis (". . .") as an operator to perform a proximity search, which looks for a specific word in the vicinity of another specific word. The operator "..." represents any number of characters between two words occurring in the same sentence in the order specified. A restricted version ("..") of the ellipsis operator limits the range between the two words specified to 25 characters. It is possible to combine these operators to narrow the focus of a query. As an illustrative example, the following table presents the results of a search in TransBase.

Table 3

Results of the Query "dump+ steel ... market" in the Parallel Corpus, TransBase

Elle est d'ailleurs soumise à d'importantes pressions depuis que les États-Unis ont imposé, tout en excluant le Canada, des sanctions à certains pays contre lesquels il a été prouvé qu'ils effectuaient du dumping d'acier sur leur marché	It has been under a great deal of pressure since the United States imposed penalties on countries, excluding Canada, that had been proven to **dump steel on their market**.
Cependant, quand les Américains viennent vendre leurs produits d'acier ici et font du dumping dans notre marché, que fait le gouvernement?	However, when the American products come across the border, when the Americans **dump steel into our market**, what do the government cronies do?
Mais nous avons été surpris du jugement du Tribunal canadien du commerce extérieur, qui a décidé que certains pays faisaient du dumping de tôle d'acier sur notre marché, mais que ce n'était pas le cas des États-Unis.	But we were surprised when our own Canadian International Trade Tribunal decided that other countries are **dumping steel plate into our market** but the Americans are not.
Pour autant que je sache, jamais nous n'avons écoulé à perte d'acier sur le marché américain.	We were not and have not, to the best of my knowledge, ever **dumped steel on the American market**.

In order to better meet the needs of the targeted user group, as was discussed in Chapter 1, the BCD will, where appropriate, provide a larger number of target-language equivalents than other bilingual dictionaries. A query of the Project's parallel corpus may yield a number of equivalents, which, because they are the product of the work of a group of experienced professional translators, will probably not appear in any dictionary or occur to a junior lexicographic assistant. A BCD lexicographer using TransBase may include one or more of these as an equivalent or in the translation of an illustrative example. In this way, BCD entries are enriched by use of the parallel corpus and offer the sophisticated user a greater choice of target-language alternatives taken from real-life contexts.

It is well known that translated texts can often prove to be unreliable documentation sources because of the insidious influence of the source language on the translation. This is a particular problem in Canada's bilingual environment where the two official languages are always in close contact and where there is, therefore, a greater risk of one language interfering with the other, resulting in "translationese". BCD lexicographers, like any terminologists, try to

avoid using translated texts as source material, although it is recognized that this is an ideal which it is not always possible to meet (L'Homme, 2004: 126). Nonetheless, the translated texts in TransBase are not used as a primary source when researching equivalents for the BCD. The primary source of equivalents in the BCD is the comparable corpus and the parallel corpus is only used at the end of the research process to ensure that no useful equivalents have been overlooked.

2.3.3. The Document Type Definition

In the early years of the Project, Dictionary entries were compiled in accordance with the BCD entry structure using succeeding versions of WordPerfect, although it was realized from the start that this was an unsatisfactory method both for the preparation of entries and for their future dissemination. Because of the potential complexity of a dictionary entry and the problems associated with handling its recursive and optional elements, it was difficult for the BCD to find a ready-made program that met its entry compilation needs, although the team took many years examining the various technologies available. Finally, it enlisted expert assistance to prepare a program appropriate for its needs, using as its basis a Document Type Definition (DTD) for the BCD entry structure, which would allow entries to be formatted in SGML. A DTD is a formal description in the markup language syntax of a particular type of document, in this case, of a dictionary entry. When XML (eXtensible Markup Language), the abbreviated version of SGML, became available in 1998, the Dictionary's SGML files were converted to this format. The DTD has been handled throughout using WordPerfect as a text editor.

2.3.4. The Lexicographic Database

The XML-encoded entries are stored in a multi-purpose lexicographic database, the creation of which was one of the four major objectives of the BCD Project (see Chapter 1, section 1.2). This database is managed using the Comet WebFileManager, a freely downloadable program that allows users in the three BCD centres to share XML files in a secure environment (<http://cwfm.sourceforge.net>). It is from this database that the printed dictionary will be extracted. By either abridging or supplementing the material stored, the database could also be used in the future to produce a smaller dictionary, such as, for example, a dictionary of Canadianisms proper, or a larger, two-volume general bilingual dictionary.

The BCD lexicographic database can also be queried using Lucene as a search tool. In this way researchers may retrieve both complete and incomplete entries to verify cross-references, for example, or to check if a particular illustrative example has already been used. The search can be limited to retrieve data from specific fields within an entry, for example, from the headword field, the sense-indication field, the compound-section field, and so on, or from any combination of fields. This is possible because of the XML tags embedded within an entry, which can be searched to identify and retrieve specific information.

2.3.5. The Workflow Database

It is obviously important for the Project to keep track of entries completed or in progress. This is done using a separate information storage system, the Workflow Database. Since, in the BCD Project, work is done not on individual headwords, but on a group of morphologically and semantically related headwords, these related entries are grouped together, not only in hard copy in one single folder, but also in the BCD Workflow Database. Thus, the file *admettre* contains the verb itself and, in addition, the derivatives *admis*, *admissibilité*, *admissible*, *admission*, *inadmissibilité*, *inadmissible*, *réadmettre* and *réadmission*. This does not in any way imply that these entries have to be embedded in one single entry in a published version of the BCD, but it will afford greater flexibility when planning future editions and formats of the Dictionary. As a family of words proceeds through the various stages of entry preparation and revision (for example, initial compilation, source-language revision, target-language revision, professorial revision, etc.), its progress is logged in the Workflow Database, thus providing information on the current status of a particular entry and on the number of entries at each stage.

Not only is the progress of work logged in the Workflow Database, but a record is also kept of key elements included in each entry. These records can be searched in order to identify and count, for example, all English headwords or all French compounds or all fixed expressions or, of particular importance for the BCD, all Dictionary entries containing a Canadianism or an example of Canadian usage.

2.4 Compilation of a BCD entry

In the ten-stage process of dictionary making listed above (2.1), stages 3 and 4 are concerned with the compilation of individual entries. Before this can begin, however, headwords have to be allocated to individual lexicographers. In the BCD Project, the headwords assigned are in the lexicographer's second language, following the usual translational convention of working from the second language into the native language. Since groups of related headwords are treated together on the BCD (see section 2.3.5 above), each lexicographer is assigned a family of entries to compile. The family *ice*, for example, includes not only the noun and verb, but also several derivatives, such as *iced*, *icily* and *icy*, all of which would usually be treated by the same lexicographer. In addition, headwords are assigned in such a way as to match, where possible, their complexity with the ability, experience and special knowledge or interest of each lexicographer.

A BCD entry contains all the standard information expected of a good desk-sized bilingual dictionary: grammatical information about the headword, equivalents, illustrative examples, usage labels and function indications. In addition, a BCD entry may include the following special features that do not figure systematically in other bilingual dictionaries of comparable size and of which some are distinctly innovative:

(i) clear separation of different senses of the headword into numbered categories
(ii) meaning indication for each major sense of the headword in the form of a partial definition or synonym
(iii) spelling variants of the headword
(iv) large number of equivalents
(v) context words to disambiguate equivalents
(vi) large number of illustrative examples
(vii) illustrative examples separated into two types: free combinations and collocations
(viii) separate section for fixed expressions
(ix) separate section for compounds containing the headword.

These innovative features are reflected in the microstructure of a BCD entry as detailed in Chapter 1, section 1.4.

The compilation of a bilingual dictionary entry comprises four parts: source-language analysis, transfer from source language to target language, synthesis and entry input. Although each stage involves certain specific tasks, a lexicographer may often work on more than one stage concurrently. He may, for example, select an example based on a partial semantic and grammatical analysis, translate it, decide to retain it and input it into the DTD, while then continuing with further analysis on other aspects of the entry. However, each stage will be treated separately below.

2.4.1. Source-Language Analysis

The first stage in the compilation of a BCD entry, the analysis of the source-language material, is often the most time-consuming because a number of different kinds of detailed research have to be carried out: investigation of variant forms, identification and ordering of the senses and grammatical and structural analyses.

The source-language research begins with an examination of information found in the major unilingual dictionaries. The lexicographer then makes an initial query of the headword in the BCD comparable corpus, adjusting the parameters to produce one-line contexts for each headword. When working with a large corpus of this kind, care must be taken to ensure that the enormous volume of available material does not overwhelm researchers, especially the more junior and inexperienced. For this reason, the number of concordance lines analyzed in the initial query of the headword is limited to a maximum of 100. This 100 can be made up of, for example, 80 occurrences from the *English Canadian Press* sub-corpus and 20 concordances from the *Wall Street Journal* sub-corpus. In this way, both Canadian and U.S. usage is examined. On the French side, the 100 examples of usage could be made up of, for example, 80 concordances from the *Presse canadienne française* sub-corpus and 20 concordances from the *Le Monde* sub-corpus, thereby providing examples of French usage in both Canada and France. More refined searches can be performed and longer contexts requested as necessary. Both the concordances and the relevant dictionary sections are analyzed in printed form (using photocopies of dictionary pages when necessary), so that they can be annotated and relevant sections highlighted using a colour-coded system. In this way, each potential element of a BCD

entry – free combination, collocation, fixed expression, compound – is clearly identified and a permanent record of the research is available for revisers.

2.4.1.1 Analysis of Variants

Variant forms are identified and verified by crosschecking between the unilingual dictionaries and the comparable corpus. The lexicographic and non-lexicographic sources may provide evidence of variation in spelling (e.g. *canoé* and *canoë*), hyphenation (e.g. *pot belly* and *pot-belly*), capitalization (e.g. *Maple Leaf* meaning the Canadian flag and *maple leaf*), feminine forms (e.g. *professeure* and *professeur*) or verb forms (e.g. *got* and *gotten*), etc.

In the majority of cases, the preferred spelling of a headword is determined by the frequency of its occurrence in the corpus. If the number of occurrences warrants its inclusion, the less common form is placed after the preferred form, thus:
dietitian *n or* **dietician**

2.4.1.2 Semantic Analysis

The lexicographer begins semantic analysis by drawing up a sense chart covering all the major senses of the headword, as presented in the unilingual dictionaries. He then analyzes the corpus concordances in order to determine the senses of the headword that are actually used and their approximate frequency of use. This semantic analysis of the corpus is often one of the most exciting aspects of BCD research, for it can reveal important differences between the number and order of senses in the unilingual dictionaries and the way the word is used in reality.

An analysis of the corpus evidence may reveal, for instance, that the headword is used in a sense not covered in the dictionaries. If the corpus provides sufficient evidence to justify it, an additional sense division will be added. If it appears that a particular sense may be restricted to Canadian usage, this can be crosschecked by comparing concordances from the different regional sub-corpora, for example, the *English Canadian Press* against the U.S. *Wall Street Journal* for English headwords and the *Presse canadienne française* against *Le Monde* and *Ouest-France* for French headwords. This is especially important given the Canadian orientation of the Dictionary.

Once the senses have been identified, they need to be ordered. Throughout the history of lexicography, four basic methods of ordering senses have been employed: chronological ordering, ordering by frequency, logical ordering and ordering by sense dominance (Roberts, "Organization of information in a bilingual dictionary entry", 1992: 224-225). As each of these methods has its limitations, the BCD, in common with many other dictionaries, uses a combination of the following criteria to determine the order of the semantic divisions:

(a) more common usage before less common usage;

(b) standard and widely used senses before more specific Canadian senses;

(c) Canadian senses before senses restricted to continental French or British English;

(d) modern usage before more outdated usage (Roberts ibid.).

A common-sense interpretation of these four criteria is required, with each headword examined on an individual basis. In most cases, senses are ordered according to their frequency in the comparable corpus (see (a) above). But given the specifically Canadian orientation of the BCD, a Canadian sense may often be placed before other senses. For example, the word *bleuet* is far more frequent in the French BCD corpus with its Canadian sense of "blueberry" (a fruit) than with its international meaning of "bachelor button" (a flower). In the BCD entry for *bleuet*, therefore, the Canadian sense precedes the international sense.

An additional criterion used at the start of the Project for determining the order of senses was to place concrete senses before abstract. However, as work progressed, it became apparent that this criterion led to frequent contradictions with the criterion of putting more common usage before less common. An abstract usage of a word is sometimes more common than the concrete usage, which may often be a technical sense and therefore less common. For example, *cœur*, meaning "bois le plus ancien au centre d'un arbre" is a concrete, technical sense that would, in BCD methodology, be placed after the more common but abstract sense of *cœur* meaning "courage". Given this type of contradiction, the criterion of concrete before abstract sense was not retained.

Being both contemporary and regionally diverse, the separate sub-corpora that make up TEXTUM provide the evidence required to order senses according to the four BCD criteria retained today.

2.4.1.3 Grammatical Analysis

Grammatical analysis in unilingual dictionaries and of the corpus concordances allow the lexicographer to identify and confirm the part of speech of the headword, especially if it is a homonym such as *jaune* (noun and adjective) or *spit* (noun and verb). Analysis of the corpus evidence may also reveal additional grammatical information, such as the use of an English noun as a modifier, for example, *grocery*, in *grocery store*, *grocery chain* and *grocery bag*. Careful analysis of the corpus concordances can sometimes highlight discrepancies between the grammatical category attributed to a word by the dictionaries and the way it is used in reality. Furthermore, dictionaries may not always agree on the grammatical category of a word. For example, the form *unifamilial(e)* is noted as being an adjective in the *Dictionnaire du français plus* and the *Dictionnaire nord-américain de la langue française*, as a noun in *Richesses et particularités de la langue écrite du Québec*, and as a noun and an adjective in the *Dictionnaire québécois d'aujourd'hui* (1993) and in the *Dictionnaire des canadianismes*. Analysis of corpus examples of this word reveals its use as both an adjective and a noun, with noun usage predominating. In such cases, it is the information obtained from the corpus that best represents authentic usage and therefore, when such decisions have to be made in the BCD Project, it is the evidence of the corpus that takes precedence (in the vast majority of cases) over the information contained in dictionaries.

2.4.1.4 Structural Analysis

The lexicographer analyzes the corpus concordances to determine how the headword functions in sentences and the structures in which it commonly occurs. Recurring linguistic structures are noted, such as, for example, a preposition commonly co-occurring with a particular noun (e.g. *a love of*) or verb (e.g. *s'avancer vers*), or a transitive verb being used frequently in the passive voice (e.g. *être débordé*). Further concordances can be generated for these and the numbers of occurrences examined, together with lexicographic evidence, to determine whether these linguistic structures should be included in an illustrative example in the entry. Other

recurrent word combinations consisting of a group of lexical morphemes are also noted and identified as either collocations (e.g. *d'avance*), fixed expressions (e.g. *c'est pas d'avance*) or compounds (e.g. *avance rapide*), along with their sense or senses, for possible inclusion in an appropriate section of the entry.

2.4.1.5 Selection of Illustrative Examples

As the lexicographer examines the corpus evidence during the semantic, grammatical and structural analyses of the headword, he also assesses the potential of each concordance line to be included as an illustrative example in the entry. Most larger contemporary dictionaries, both unilingual and bilingual, contain examples of usage, but their inclusion often appears haphazard and unsystematic. One of the most innovative aspects of the BCD is that the headword is routinely illustrated by an example, which may be of one of two types:

free combination: a phrase where the headword is used without any special syntactic or semantic restraints;

collocation: a commonly recurring idiomatic word combination, for example, *after all*, *to eat heartily*, *au mur*, *fâché contre*, etc.

The vast majority of free combinations given in the BCD are lifted verbatim from the texts in the corpus. They provide examples of language use in authentic texts "composed independently in the respective language communities" (Laffling, 1991: 81). However, in order to work well, a dictionary example must not be too long or unwieldy, so occasionally a lexicographer may cut material from the beginning or end (but not from the middle) of a particular corpus excerpt. The lexicographer signals the cut by marking the example with his initials, thereby drawing the reviser's attention to it for checking. This is the only type of modification that BCD lexicographers make to corpus material they select as an illustrative example.

There may sometimes be a specific aspect of the usage of the headword that makes the inclusion of a free combination redundant. For example, a free combination may be omitted for a specialist sense whose field label identifies its area of usage so precisely that no further clarification is needed, for example:

arme . . . **4** *nfpl* (*Escr*) fencing.

Similarly, the inclusion of a free combination is not essential if a given sense is illustrated by a collocation.

2.4.2. Transfer from Source Language to Target Language

In the preparation of a bilingual dictionary entry, the transfer of the source-language material into the target language is a major step, unique to bilingual lexicography. It is at this stage that the lexicographer must establish target-language equivalents for each of the senses of the headword and for any compounds as well as to translate all the free combinations that have been retained along with any collocations and fixed expressions.

BCD methodology recommends that, before examining the target-language equivalents proposed by the main bilingual dictionaries and term banks for the headword in question, the lexicographer runs through the concordances from the comparable corpus and mentally translates each one, making a note of each equivalent that comes to mind. This exercise, useful at all times, is essential where no equivalent is given in the bilingual dictionaries, for example, when a particular sense is not included, or in the very frequent cases where Canadian usage is not covered. Lexicographers can check any intuitive ideas that they may have for an equivalent in either the target-language side of the comparable corpus or in the parallel corpus or in both to ensure that the proposed equivalents are indeed used in the same sense as the source-language word. Those equivalents that do appear in the bilingual dictionaries and term banks are generally validated in any case in the comparable corpus and, if more information is required, in the parallel corpus also.

Where several interchangeable, synonymous equivalents exist for a given headword, the normal practice on the BCD is to present these equivalents in the order of their frequency in the corpora. The comparable corpus is also used to determine the most frequent spelling of an equivalent. This is important as the BCD gives only the most common spelling of an equivalent.

Having gathered the source-language material for the illustrative examples and determined which are to be classified as free combinations and which as collocations, the

lexicographer then turns his attention to the translation of this material, producing his own translations as required and verifying translations that already exist in the bilingual sources (dictionaries and term banks). Because the BCD seeks to give users a greater choice of target-language choices, possible alternatives are noted. For example, at the French headword *armature*, four possible translations of a free combination are offered:

> **les fermes collectives constituent l'armature économique de ce pays** collective farms make up this country's economic framework = collective farms make up this country's economic infrastructure = collective farms form the basis of this country's economy = collective farms are the backbone of this country's economy.

All translations of illustrative examples are carefully checked at the target-language revision stage (see section 2.5.3 below) when the reviser ensures that no important alternative is excluded and that no redundant or repetitive material remains.

2.4.3. Synthesis

Synthesis is the third phase of entry compilation in which the lexicographer reviews all the information retained and, combining the results of the decisions taken at the two earlier stages of source-language analysis and translation, aims to make the final selection of material to be included. If, for example, additional equivalents have come to light during the translation of the illustrative examples, some adjustment to the entry may be needed to add these new equivalents to the preliminary list. It is at this stage also that sense divisions, established during source-language analysis, may need to be adjusted. This adjustment is sometimes necessary because the target language makes distinctions that the source language does not, or because the target language does not make a distinction where the source language does. On the whole, unilingual dictionaries divide senses more finely than bilingual dictionaries. If two or more senses of a single headword are translated by the same target-language equivalent(s), the lexicographer may choose to combine a number of separate senses in one sense division. For example, for the headword *chaudière*, the *Dictionnaire québécois d'aujourd'hui* and the *Dictionnaire de la langue québécoise* distinguish the general "seau en métal, en bois ou en plastique" and the specific "petit seau servant à recevoir la sève des érables". Since, in both cases, the equivalent is "bucket" or "pail", it is not necessary to subdivide the senses.

2.4.4. Entry Input

The final phase of compiling is entry input where a first draft of the Dictionary entry is entered into the DTD on the computer. Some decisions may still remain to be made at this stage, concerning for example, the most appropriate order for the illustrative examples or the possibility of presenting collocational material through the use of actants.

2.5 Revision

As detailed above (2.1), the production of the BCD involves ten stages. Following preparation of the entry (stages 3 and 4), stages 5, 6 and 7 are concerned with its revision. Revision focuses on all aspects of the dictionary, from the division of the senses and the sense indication provided for each sense division, to the choice and translation of illustrative examples and the use of labels. Revisers may review all the corpus evidence used by the lexicographer in the preparation of the source- and target-language elements of the entry and, if necessary, perform further searches in either corpus or in both BCD corpora. All the documentation used in the compilation of the entries is kept together in one paper folder, which greatly simplifies the revisers' task as they can verify more quickly and efficiently the sources consulted. Revisers may also wish to consult additional lexicographic sources such as specialized dictionaries, for example, of information technology, law, medicine, proverbs or dialect. Revision itself involves a number of stages, which are discussed below.

2.5.1. Preliminary Revision

In the preliminary revision, the aim is to check that the information provided is complete and accurate overall and that it is placed correctly within the entry, respecting the conventions of the DTD format. Thus, revisers check that the different senses of the headword have been clearly identified and that free combinations and collocations have been inserted in the appropriate sense division and in the proper order within that division. This preliminary revision stage also ensures that suitable examples have been selected and that their translations are acceptable. A note is made of any elements requiring further verification. The preliminary revision is carried out in the BCD Project by a full-time lexicographer, usually a source-language reviser, who has been "closely involved in the formative phase of the Project and who is therefore fully cognizant of

the Project's aims, philosophy, and methodology" (Roberts, 1992: 107). This preliminary revision, as well as the other stages of revision, is carried out on hard copy. BCD revisers have found it more practical and convenient to make their notes and suggestions in the margin of a printed copy than to try and annotate a dictionary entry on-screen. Following preliminary revision, the first draft of the entry is returned as quickly as possible to the lexicographer who compiled it so that he can make the corrections to the XML file. This correction stage and accompanying discussions with the reviser provide valuable feedback to less experienced research assistants and form an important part of the BCD training.

2.5.2. Source-Language Revision

Whereas the preliminary revision examines both source- and target-language elements of the entry together, the next two stages of revision focus on these elements individually. In the source-language revision, the reviser examines all the source-language material more closely, checking that all the required information has been included and that it is accurate. The source-language reviser focuses in particular on checking the accuracy of the sense divisions of the headword, reviewing and, if necessary, shortening or modifying the sense indications. Examples may be added or amended or unnecessary examples deleted. Usage labels, if present, are checked or they are added where necessary. After the source-language revision, the original compiler of the entry incorporates any corrections required and the entry then moves on to the next stage, the target-language revision.

2.5.3. Target-Language Revision

Target-language revision is designed to ensure that the equivalents and translations of examples given in the entry are accurate and idiomatic. As was discussed in Chapter 1, the BCD aims to provide its sophisticated users – translators, bilingual editors, government employees, journalists, second-language teachers and advanced second-language learners – with a greater number of target-language choices than is currently the norm in bilingual dictionaries. There is, therefore, no fixed limit to the number of equivalents and translations that may be included in a BCD entry. Assessing the target-language choices included by the lexicographer, the reviser will ensure not only that no important alternative equivalent or translation has been omitted but also that redundant or repetitive elements are excluded.

Revisers will also ensure that equivalents have been disambiguated where necessary and that the necessary usage labels have been included. This revision will also check that any relevant grammatical information about target-language items, such as the gender of French nouns, is accurate and complete.

2.6 Consultation of Specialists

It is at this point (stage 8 in section 2.1 above) that subject-field experts or language specialists may be consulted for a number of entries or parts of entries. Given the specifically Canadian orientation of the BCD, it may be necessary to consult language specialists regarding particular points of Canadian usage in French and English and incorporate their recommendations into the entries concerned.

2.7 Arrangement of Entries

As was discussed in Chapter 1 (1.3.2), the arrangement of lexical items and information pertinent to them in the first version of the BCD (i.e. which lexical items are treated as headwords and which nested within an entry) was decided at an early stage.

However, depending on the particular version or format of the Dictionary under consideration, it may be necessary at a later stage to embed more entries or conflate some entries to form a single entry. At the moment, for example, separate headword status is assigned to all English adverbs. In a case such as *defiant* and *defiantly*, where both adjective and adverb cover the same areas of meaning, the adverb could be nested under the adjective as a space-saving device in a more abridged dictionary published at a later date.

2.8 Final Editing

In the final editing of the BCD, all entries will be re-read, in their final format and sequence and all information reviewed for accuracy and consistency. It is at this point that cross-references, spelling and typographical features can be reviewed to ensure consistency across all the compiled entries. It is also appropriate at this stage to revise particular types of entry

together, for example, the days of the week, numerals, prepositions, etc., in order to achieve maximum consistency in coverage and layout.

2.9 Conclusion

This chapter has examined in detail the methodology of the BCD, how an entry is produced and the tools used. There is no doubt that the use of corpora has greatly enhanced the quality of BCD entries, giving lexicographers the confidence to make authoritative choices regarding both source- and target-language lexical items based on their analysis of examples of authentic language in a range of contrasting sub-corpora. BCD Project leaders have succeeded in identifying and developing a range of technological solutions to provide an efficient and flexible working environment for the production of the Dictionary.

3. WHY? THE NEED FOR THE BILINGUAL CANADIAN DICTIONARY

Although there are many adequate and even excellent French-English, English-French dictionaries on the market today, none of them completely meets all the needs of the Canadian user. The primary reason for this is that the French and English of Canada are not identical to the varieties of these languages used elsewhere and recorded in existing bilingual dictionaries. This chapter will first detail the particular features that distinguish Canadian French and Canadian English.

Although it is to be expected that bilingual dictionaries produced outside Canada have not given extensive coverage to Canadian usage in French and English, it is harder to understand why so few bilingual French and English dictionaries have been produced in Canada itself. The second part of the chapter will seek to explain this situation by summarizing the significant events in bilingual lexicography in Canada.

An analysis of the treatment of Canadian usage in bilingual French-English, English-French dictionaries will be given in Chapters 4 and 5.

3.1 Regionalism

This study focuses primarily on the one aspect of the Bilingual Canadian Dictionary that distinguishes it clearly from other bilingual dictionaries: its treatment of Canadianisms. But before establishing a definition of a Canadianism, it is useful first to position the Canadianism in a larger grouping as one of a number of different regionalisms within the French and English languages.

Halliday, Bell, and others defined linguistic variation according either to the *use* of language or to the *user* of language. Whereas use-related variation is seen in differences of register (formal, informal, colloquial, casual, etc.), what Bell terms "user-based variation" (1991: 184) is realized as dialect, in which the language of an individual carries indications of age,

social class, or geographic origin. Regionalism is linguistic usage that is restricted according to the geographic origin of the language user.

Basing his findings on "diverses expériences dictionnairiques", Jean-Claude Boulanger, editor-in-chief of the *Dictionnaire québécois d'aujourd'hui*, estimates that the proportion of French vocabulary common to all regional varieties of the language is approximately 80% (Boulanger, 1998: 176). Others have suggested that the proportion may be as high as 90% (De Villers, 1998: 60). As for the vocabulary of the English language, while not attempting to put a precise figure on the proportion, David Brazil states that there is "a central core of ordinary, most frequently used words that all geographically distinguished varieties of the standard language share" (Brazil, 2000: xxvii). It is the small remaining portion then that constitutes that part of the vocabulary of a language that is particular to any one geographic region, that part which is constituted of regionalisms.

The definition of the word "regionalism" requires some clarification as it varies according to sources consulted and may be interpreted in a variety of different ways depending on the criteria applied. A regionalism can be defined according to two distinct pairs of parameters: firstly, according to whether it is restricted to a national geographic region or a local geographic region and secondly, according to whether it is a lexical item of a purely linguistic nature or a lexical item denoting a culture-specific phenomenon.

As far as the geographic restriction is concerned, dictionaries define a regionalism in general terms as, for example, " a word or phrase associated with a particular locality where a language is used" (*Dictionary of Lexicography*) or " a word or an expression whose use is restricted to a specific geographical area within a linguistic community" (*Translation Terminology*). But the *Dictionary of Lexicography* makes a further and valid distinction between regionalisms that are typical of "a national variety, e.g. Americanisms or Briticisms", and those that are typical of "a geographical dialect or topolect, e.g. the French of Paris or Québec".

From a synthesis of the defining features of "regionalism", the following definition has been constructed: a lexical item used either in one or more (but not all) countries where the

language in question is used, or in a specific area (or areas) of a single country. In the first category fall, for example, both North Americanisms in English, which are used in Canada and the U.S., but not in other English-speaking countries, and Canadianisms, which are used in Canada, but not in the U.S. Also in this first category fall lexical items that are typical of any of a large number of regional varieties of both French and English that are aligned with the national boundaries of the countries in which either of the two languages is used (e.g. "belgicismes", "helvétismes", Australianisms, etc.). Within each of these national regional varieties are usages typical of a more restricted geographic area such as those examples of usage peculiar, for example, to Acadian French or to the French language used in any of the regions of France, such as Brittany or Provence, and in English those examples of English usage identified in dictionaries by such markers as "especially Ontario", "Northern English dialect", "Southwest US", etc.

In the second pair of parameters used to define a regionalism, De Villers (1998) identifies firstly, words, expressions, and senses that are specific to the language of a particular geographic region and secondly, those lexical items "qui désignent des réalités régionales" (De Villers, 1998: 159). The first type of regionalism (which will be known in this study as a "linguistic regionalism") may simply be an orthographic form that differs from that (or those) used in other geographic areas. For example, the orthographic forms *centre*, *honour*, *realise*, are used in some parts of the English-speaking world and the variant forms *center*, *honor*, *realize*, in others. This first type of regionalism is also characterized by the existence in different regions of different lexical units to denote the same reality. Thus, to denote, for example, a "navire spécialement aménagé pour faire traverser des passagers, des véhicules d'une rive à l'autre d'un cours d'eau" (*Dictionnaire québécois d'aujourd'hui*), Canadian French uses *traversier*, whereas in France the word *ferry-boat* is used. Other pairs of this kind, where different lexical units exist in different regions to identify the same notion, include the Canadian French *chiropratique* and *bande publique*, known in France as *chiropraxie* and *canal banalisé*, and in English such North American/British differences as *gas/petrol*, *pants/trousers*, and *sidewalk/pavement*. This first type of regionalism also includes what may be termed "semantic regionalisms" in which a word acquires an additional sense only in a particular geographic region. Semantic regionalisms in Canada include, for example, in French, *dépanneur* (in the sense of "épicerie de quartier"),

souffleuse (in the sense of "véhicule ... qui permet de projeter la neige à distance"), and *tabagie* (in the sense of "établissement commercial où l'on vend surtout du tabac") and in English, *Confederation* (in the sense of "the federal union of provinces and territories forming Canada"), *riding* (in the sense of "a district whose voters elect a representative member to a legislative body"), and *warden* (in the sense of "the head of a county council").

In contrast to the "linguistic regionalism" defined above, the second type of regionalism denotes "des notions purement locales ou institutionnelles" (Office québécois de la langue française, 1990: 118) and includes those lexical items identified by Zgusta as "culture-bound words" (1971: 294) and by the *Dictionary of Lexicography* as "culture-specific vocabulary", that is, "words and phrases associated with the 'way of life' of a language community" (ibid.). This type of regionalism (which will be known in this study as a "cultural regionalism") refers to exclusive features of the culture, the climate, the flora and fauna, etc. of a particular geographic region. In Canada, *carte d'assurance-maladie*, *cégep*, *épluchette*, and *réserve amérindienne* in French and *to bush the ice*, *Grit*, *quinzhee*, and *snow route* in English are typical of this kind of regionalism in which both "le signifiant et le signifié sont tous deux régionaux" (Office québécois de la langue française, 1990: 175).

3.2 The French and English Languages in Canada

Canadianisms such as those presented above have resulted from the historic development of French and English in Canada. In 1534, the French explorer, Jacques Cartier reached the Gulf of St Lawrence and in 1583 Sir Humphrey Gilbert established the first permanent English settlement in North America in St. John's, Newfoundland. Within seven years of each other, between 1603 and 1610, Frenchman Samuel de Champlain and Englishmen Henry Hudson and John Guy claimed the land that is now Canada for their respective kings. Over the succeeding centuries, generations of Frenchmen and Britons followed in their footsteps, firmly establishing the French and English languages "en terre laurentienne" *(Dictionnaire québécois d'aujourd'hui*: 1993, ix). Cut off from their native roots in their respective homelands, transported to a world of new realities, and influenced by the languages and regional dialects of their neighbours close by, the French and English of Canada developed distinctive features.

These distinctive features are what constitute Canadian usage. Attempts to define the precise nature of Canadian usage have often led to argument about the exact definition of the term "Canadianism". In order to be considered a Canadianism, does a lexical item have to originate in Canada? Does it, moreover, have to be used solely in Canada? In his preface to the *Dictionary of Canadianisms on Historic Principles* of which he was editor-in-chief, the celebrated Canadian scholar and lexicographer, Walter Avis, defined a Canadianism as "a word, expression, or meaning which is native to Canada or which is distinctively characteristic of Canadian usage though not necessarily exclusive to Canada" (Avis, 1967: xii). Although Avis sought to define only the English Canadianism, his definition serves equally well to define the French Canadianism. So that the Canadianism label *CD* may be attached accurately and consistently throughout the BCD to both source- and target-language items, a firm definition of Canadianism in the two languages is required. Having considered Avis's definition along with many others, BCD editors defined a Canadianism, in the context of the BCD, as a feature of the French or English languages that is either used solely in Canada or that is used far more frequently in Canada than in other French- or English-speaking countries, regardless of whether it originated in Canada or not. The next two sections will show how Canadianisms developed in this country's two official languages.

3.2.1. Development of Canadian French

Although the French settlers arriving in Canada in the 17th-century came from many different provinces of France, each group speaking the local patois of their region, it did not take long for a common unified speech to evolve. As protection against the hostile environment, the new settlers lived in closely gathered communities where a common language was essential. The result of this close interaction was that, by the middle of the century, a more homogeneous French language had replaced the individual dialects the settlers had brought with them. French visitors to Canada at this time often commented on the purity of the French language heard in Canada. For example, an envoy sent by Louis XV to assess the state of French possessions in North America reported that "nulle part ailleurs on ne parle plus purement notre langue" (Orkin, 1971: 9).

But the second half of the 18th century brought two significant events that changed forever the nature of the French language in Canada. The first of these was the Treaty of Paris of 1763, by which Canada was ceded to Great Britain, effectively putting an end to French immigration and influence in North America. As the French language was then cut off from its roots, subsequent linguistic developments in France had very little effect on the French spoken in Canada.

The second significant event was the War of American Independence (1776-1783), during and after which large numbers of English-speakers, loyal to Britain, moved north across the border into Canada, hoping to find there a way of life more compatible with their loyalist sympathies. Not only did anglophones increasingly outnumber their French-speaking neighbours, but they also led the way in Canada in the development of educational, economic, social, political, and commercial systems. A knowledge of English became a valuable asset for francophones, bringing with it social acceptance and a potential for success and prestige.

The inevitable result of two languages existing side by side in close proximity is the linguistic phenomenon of borrowing in which elements of each language penetrate the other. Given the numerical and cultural dominance of English- over French-speakers in Canada, this phenomenon has been largely, but not exclusively (see section 3.2.2 below), a one-way process, with English exerting a far greater influence on the French language than vice versa. This penetration is seen most clearly in the large number of lexical items that have found their way into Canadian French from English, either by borrowing without any change of form (e.g. *balance*, *chum*, *fun*, *spa*), by borrowing with a change of form (e.g. *bines* from *beans*, *contracteur* from *contractor*, *drave* from *drive*) or by loan translation (e.g. *autobus scolaire* from *school bus*, *crème glacée* from *ice cream*, *personne-ressource* from *resource person*, *surtemps* from *overtime*).

From the beginning of the 19th century right up to the present day, Canadian French has been criticized for being too affected by the English language and for departing too far from the norms of the French language as spoken in France. Although, given the peculiar circumstances of the French language in Canada, these phenomena are, in linguistic terms, entirely inevitable,

language "purists" have continued to insist that Canadian French should be purged of all anglicisms and should conform to the educated or "correct" French of metropolitan France. These have been the twin concerns over the years of many French Canadians, a number of whom have supported the imposition of some measure of control, of "dirigisme linguistique", upon the French language in Canada, seeking to halt what they see as its degradation and to protect it from further corruption. It was indeed these concerns that led to the establishment in 1961 of l'Office de la langue française whose aim, while recognizing francophones as "un groupe solidaire dans le monde" (*Répertoire des avis linguistiques et terminologiques*: 179), is to improve the standard of written and spoken French in Quebec by establishing "une norme de la langue française au Québec" (ibid. 172). These long-running battles over the standards of the French language in Canada have led to what Bouchard has called the "insécurité linguistique" of French-speaking Canadians (Bouchard, 1998: 72).

The distinctiveness of Canadian French consists not only in its early homogeneity and its susceptibility to the influence of English, but also in the new words and senses that it added. While neologism is a natural linguistic process common to all languages that are cut off from their origins, the special conditions of North American life inspired the creation of a significant number of new words and the addition of new senses in Canadian French. This aspect of Canadian French will be discussed in greater detail in section 3.3.1.4.

3.2.2. Development of Canadian English

Although Canadian English does not seem to be as distinctive as Canadian French, it is nevertheless different to a certain degree from both British and North American English. Because far and away the majority of English-speaking settlers and immigrants in Canada have come from Britain and from what is now the United States, the variety of English that has developed in Canada is a blend of British and American English. This blend, while containing elements of both British and American English, is also recognizably different from both. Throughout the course of its development, Canadian English has been affected to a greater or lesser extent at different periods by both British and American English, depending on the major immigration pattern at the time. But especially since the latter part of the 18th century, the American influence on Canadian English has been enormous. At the end of the Revolutionary

War in 1783, for example, large numbers of Americans fled the United States for Canada, bringing with them their American English. All in all, Canadians have been exposed for a long time to the powerful influence of the American English of the numerous cross-border settlers, immigrants, and visitors, and of the American English that they hear in movies and in TV and radio programs, and that they read in novels, magazines, and newspapers.

As well as being a blend of British and American English, however, Canadian English also contains certain features that are peculiar to it alone. Most prominent among these are a number of lexical items denoting specifically Canadian realities, be these natural phenomena such as the *chinook arch* (a cloud formation seen on the Western prairies) or social, economic or political features, such as the *allophone* (an immigrant whose first language is neither French nor English), the *loonie* (the Canadian dollar), or the *Grits* (the Liberal Party).

Like Canadian French, Canadian English has also been subject to the phenomenon of borrowing although, not surprisingly, given the fact that English speakers have formed the significant majority in Canada for over 200 years, the degree of penetration of English by the languages of its neighbours has been relatively small. Examples of borrowing into Canadian English from Canadian French as well as from the native languages of Canada are given in section 3.3.1.3 below.

3.3 Distinctive Features of Canadian French and Canadian English

Canadianisms in French and English fall into one of four main categories: lexical, grammatical, orthographic and phonetic. As phonetics is a highly specialized field of linguistic study and given the problems associated with recording the wide variation of pronunciation of both French and English across Canada, it was decided not to include phonetic transcriptions in the first edition of the BCD. Phonetic Canadianisms will not, therefore, be discussed here.

3.3.1. Lexical Canadianisms

Lexical Canadianisms are words or senses that are typically Canadian and it is these items that are by far the most numerous of the four different categories of Canadianism mentioned above (3.3). The four principal ways in which lexical Canadianisms are formed are by

archaism, by semantic extension, by borrowing, and by neologism. These four linguistic processes will now be presented.

3.3.1.1 Archaism

Cut off from its roots in the native country, any language follows its own natural development and, being subject to different influences in its new home, does not evolve in the same way as the language does in the mother country. Both Canadian French and Canadian English have retained a number of lexical items that are now considered archaic outside Canada, but as the terms of the Treaty of Paris of 1763, under which Canada was ceded to Great Britain, effectively put an end to French immigration and, consequently, to all linguistic influence from France, Canadian French has a higher number of archaisms than Canadian English.

Here is a selection of just a few archaisms in French, with their definitions in parentheses. (These and all subsequent definitions are taken from the 1993 edition of the *Dictionnaire québécois d'aujourd'hui*, unless a different source is quoted.): *abrier* or *abriller* ("recouvrir" or "cacher"), *appartement* ("pièce"), *ennuyant* ("ennuyeux"), *grafigner* ("égratigner"), and *en temps* ("à temps").

Some examples of lexical items retained in Canadian English that are archaic elsewhere are *fall* in the sense of "autumn", *mad* meaning "angry", and *I guess* in the sense of "I suppose". (These and all subsequent definitions are taken from the 2005 edition of the *Canadian Oxford Dictionary*, unless a different source is quoted).

3.3.1.2 Semantic Extension

Canadianisms may also be formed when a lexical item acquires one or more additional senses particular to Canada. For example, in all parts of the French-speaking world, including Canada, *blonde* denotes a female with blond hair, whereas in Quebec this word has taken on an additional sense of "femme, jeune fille, qu'on fréquente ou avec laquelle on vit maritalement". Other examples where a Canadianism has been produced by semantic extension (with the additional sense given in parentheses below) include *babillard* ("tableau d'affichage"), *casse-croûte* ("restaurant spécialisé dans la préparation de plats rapides"), *frasil* ("masse plus ou moins

dense de cristaux ou de fragments de glace") from *fraisil* ("cendre de charbon de terre", *Le Nouveau Petit Robert*), *gratte* ("chasse-neige"), *poudrerie* ("neige fine et sèche"), and *avoir de l'allure* as in *ton idée a de l'allure* ("ton idée est intéressante").

Canadianisms have also frequently been produced in English by semantic extension. An example of this process is *chesterfield*, a word originally used to denote a specific style of "padded sofa with arms the same height as the back", but which acquired an additional sense in Canadian English to mean "any couch or sofa". Thus, in Canadian English, *chesterfield* now has two senses, whereas in British and American English it has only one. Similarly, the collocation *by acclamation* acquired an extra sense in Canadian English ("by virtue of being the sole candidate"), but, by contrast with the previous example, it then lost the original meaning of "by an overwhelming majority without a ballot" (*Collins English Dictionary* 2000), which is the sole sense retained in British and American English. The words *band* ("an Indian community officially recognized as an administrative unit by the federal government"), *concession* ("a tract of surveyed farmland"), *confederation* ("the federal union of provinces and territories forming Canada"), and *reeve* ("the elected leader of the council of a town or other rural municipality") have undergone similar processes of semantic extension in Canada.

3.3.1.3 Borrowing

The process of linguistic borrowing takes place whenever different language communities come into contact with one another. Borrowing consists of the introduction of a lexical item from one language into another, usually in order to fill a lexical gap – to provide a name for a new object or phenomenon for which no name exists in the host language (Yule, 2006: 52). There has been a considerable amount of borrowing between the French and English languages in Canada, but, as anglophones have outnumbered francophones in Canada since the second half of the 18th century (see section 3.2.1 above) and have been in a stronger political, economic and social position than francophones, it is not surprising that the number of borrowings from Canadian English into Canadian French is far greater than those occurring from Canadian French into Canadian English.

At its most basic, an anglicism in the Canadian context is simply a "mot anglais employé en français" (*Dictionnaire québécois d'aujourd'hui*, 1993: xi), but these lexical items exhibit a variety of characteristics once they are transferred from one language to the other. Some loan words, for example, retain both the same spelling and meaning as in English: *caucus* meaning "réunion", *cheap* meaning "bon marché", and *has-been* meaning "personne . . . qui a fait son temps". Others retain the English spelling but change their meaning: for example, *bloke*, a pejorative adjective and noun in Canadian French meaning "anglophone", denotes "a man, a fellow" in British English, although it is no longer current in Canadian English. In their move from English, some words are adapted to French orthography, morphology, or pronunciation, but retain the same meaning, for example, *aréna* meaning "amphithéâtre sportif", *bécosses* meaning "toilettes" (from *backhouse*), *bruncher* meaning "prendre le brunch" (from *brunch*), and *matériel* meaning "tissu". Others acquire both a new form and a new meaning, for example, *poutine* (from *pudding*).

There is a further stage in the borrowing process where the English lexical item is translated literally into French, creating a loan translation or calque (Office québécois de la langue française, 1990: 177), for example, *carré aux dates* from "date square", *centre d'achats* from "shopping centre", and *fin de semaine* from "weekend".

Yet another category, described by Jean Darbelnet (1976: 47) as "anglicismes de maintien"[13], comprises forms which are now archaic outside Canada but that have been retained in Canadian French because of their similarity to an English word of the same or similar meaning. This group includes such words as *appartement* ("pièce"), *barbier* ("coiffeur"), *brassière* ("soutien-gorge"), *breuvage* ("boisson"), and *mitaine* ("gant").

The anglicisms that usually cause the greatest concern to those who seek to preserve the integrity of the French language in Canada are those that are perceived as totally redundant because a perfectly valid way of expressing the same meaning already exists in French. Such

[13] Darbelnet defines "les anglicismes de maintien" as "des mots authentiquement français qui sont tombés en désuétude en France et se maintiennent au Canada à cause du voisinage du mot anglais de même famille" (1976: 47).

"anglicismes critiqués" include *gazoline* for *essence*, *gradué* for *diplômé*, *rafting* for *radeau* ("descente sportive, en canot pneumatique, de cours d'eau"), and *régulier* for *ordinaire*.

Although far fewer, as explained above, there are, nonetheless, a number of borrowings from the French language as used in Canada into Canadian English. *Depanneur* is a particularly interesting example, where English-speakers in Canada have adopted the French Canadianism *dépanneur* (meaning "épicerie de quartier"), which was itself created by semantic extension from an existing French word, to denote exactly the same kind of "convenience store". Other borrowings from Canadian French into Canadian English include *Fameuse* ("a variety of apple grown esp. in Quebec"), *gaspereau* (a type of herring), *Metis* ("a person of mixed Aboriginal and European descent"), *portage* ("the carrying of boats"), *prairie*, and *rapids*. A few loan words, *babiche*, for instance, (meaning "strips of rawhide or sinew used as laces . . . etc.") and *caribou*, have undergone two stages of borrowing, firstly, from a native language into French and thence into English.

Borrowings from French into Canadian English are often termed "gallicisms" when some measure of disapproval is implied. As with Canadian French, disapproval is often expressed when it is perceived that a perfectly valid term already exists in the host language. The use, for example, of the Canadianism *to francize* (from *franciser*) instead of the existing English verb *to gallicize*, is the kind of borrowing that attracts criticism.

A small amount of borrowing from native American languages into both Canadian French and English has occurred, mostly of words denoting elements of the natural environment. Borrowings from Amerindian languages into Canadian French include names of plants and trees (*atoca*, *chicoutai*, *pimbina*), fish (*achigan*, *touladi*, *maskinongé*), birds (*cacaoui* or *kakawi*), and animals (*carcajou*, *caribou*, *ououaron, pékan*). Borrowings from Amerindian languages into Canadian English are of similar types, including names of plants (*saskatoon*), trees (*tamarack*), fish (*kokanee*, *maskinonge*, *muskellunge*), and animals (*pekan*, *wapiti*).

Borrowing from the Inuit language into French and English has been even more limited than from the Amerindian languages and is restricted to lexical items, mostly nouns, denoting0

only a very few aspects of Inuit life. Canadian English and Canadian French have borrowed many of the same words from the Inuit languages including *igloo*, *kayak*, *mukluk* and *parka*, and the less well-known *qiviut* ("fine, soft wool from the underbelly of a muskox") and *umiak* ("large, open, flat-bottomed boat").

3.3.1.4 Neologism

Finding themselves in a new country, surrounded in their daily life by physical realities for which no words existed in their own language, the French and English settlers in Canada coined new words. Many of these neologisms refer to aspects of the natural environment – plants, trees, fish, animals, and birds – or to the exploitation of this environment in agriculture, fishing, the fur trade, the lumber industry, etc. Such neologisms, for example, as *acériculture* ("culture et exploitation industrielle d'une érablière"), *atocatière* ("terrain où poussent les atocas"), *bleuetière* ("terrain où abondent les bleuets"), *coureur de bois* ("aventurier qui s'occupait de la traite des fourrrures"), and *pourvoirie* ("entreprise qui loue aux chasseurs et aux pêcheurs des installations et des services") in French, and *annual allowable cut* ("the volume of wood which may be cut each year in a specified area"), *made beaver* (a unit of exchange, a coin, or a token formerly used among fur traders), and *prairie crocus* (*Anemone nuttalliana*, the floral emblem of Manitoba), in English, are lexical Canadianisms denoting phenomena of special relevance in the Canadian context.

As Canadian society developed, new realities were created in all areas of social activity, referents for which no word existed in either language. In order to supply the designation required in this kind of situation, language takes a number of routes, one of which is to use the existing language to create a completely new form. Examples of such creations in Canada include *cégep* ("collège d'enseignement général et professionnel"), *lacrosse* ("sport d'équipe"), and *motoneige* ("petit véhicule d'hiver . . . qu'on enfourche . . . pour se déplacer sur la neige") in French, and *friendship centre* ("an institution established in a predominantly non-Aboriginal community to provide counselling and social services etc. to Aboriginal people"), *humidex* ("first used by the Toronto weather office in 1965"), *loonie* ("the Canadian one-dollar coin"), and *toonie* ("the Canadian two-dollar coin"), in English.

In other cases, the referent also exists in other parts of the French- and English-speaking worlds, but in Canada, a new and different word is coined to designate it, for example, in French, *croustilles* ("pommes de terre coupées en minces rondelles puis frites", = in France, *chips*), *débarbouillette* ("petit carré de tissu éponge avec lequel on fait sa toilette" = in France, *gant de toilette*), *lave-auto* ("établissement équipé d'une installation automatique pour le lavage des autos" = in France, *portique de lavage automatique*), *podiatre* ("médecin qui soigne les pieds" = in France, *podologue*), and *téléroman* (= in France, *feuilleton télévisé*); and in English, *beverage room* ("a lounge, bar, etc. where alcoholic drinks are sold"), *ice pad* (ice rink), *Red River cart* ("a sturdy two-wheeled wooden cart pulled by oxen or horses"), and *summerfallow* ("agricultural land left fallow in the summer").

3.3.2. Grammatical Canadianisms

As has been discussed, lexical Canadianisms account for the vast majority of the features that distinguish Canadian French and English from the varieties of these languages spoken elsewhere. By contrast, scholars have identified only a few slight differences between the grammar of French and English used in Canada and the grammar of these languages used elsewhere (Hamilton, 1997). Given the small number of these differences and their minor nature, I will provide only two illustrations:

(i) a small number of French nouns, mostly loan words, have a different gender, for example, *gang*, *job*, and *minestrone* are feminine in Canada, but masculine in France; *mozzarella* is masculine in Canada and feminine in France; *pamplemousse*, *parka*, and *thermos* are masculine in Canada and either masculine or feminine in France; and *radio* is either feminine or masculine in Canada and feminine in France;

(ii) as noted above, the distinctive feature of English Canadian usage is that it is divided between North American usage and British usage. In the case of certain verbs having an alternative form of the past tense, for example, there are some cases where Canadian usage firmly follows American usage (*leaned* rather than *leant, learned* rather than *learnt*, *spelled* rather than *spelt*), some cases where Canadian usage is divided between American and British usage (*dwelled/dwelt*, *gotten/got*, *smelled/smelt*, *spilled/spilt*), and other cases where corpus analysis indicates that Canadian usage may prefer a different form to that preferred by both American and British English (*dove* rather than *dived*).

3.3.3. Orthographic Canadianisms

There seem to be no spelling variants that are unique to Canadian English, although, as with grammatical Canadianisms, there are several preferred orthographic variants that Canadian English shares with American English and which distinguish both varieties from British English, typified by such examples as *plow* instead of *plough*, *to practice* instead of *to practise*, *program* instead of *programme*, *recognize* instead of *recognise*, *tire* instead of *tyre*, etc. However, as Canadian English is a blend of British and American English, it is not surprising to find, on the other hand, that there are several preferred orthographic variants that Canadian English shares with British English and which distinguish both varieties from American English, typified by such examples as *axe* instead of *ax*, *catalogue* instead of *catalog*, *centre* instead of *center*, *cheque* instead of *check*, *honour* instead of *honor*, and *travelled* instead of *traveled*, etc.

There are a few orthographic variations that distinguish French Canadian spelling from that of France. Canadian French prefers, for example, *abatis* to *abattis*, *aérobique* to *aérobic* and *canoé* to *canoë*.

3.4 French and English Lexicography in Canada

Despite the fact that the French and English languages have developed a significant number of distinctive features, and also despite the fact that the first *Official Languages Act* making Canada an officially bilingual country was passed over 35 years ago in 1969, there is still no up-to-date English-French, French-English general dictionary of Canadian origin on the market. Canadians are obliged, therefore, to use English-French dictionaries produced in Europe and geared to the European, and not the Canadian market. When Canadianisms are included in these dictionaries, their treatment is often inaccurate or at the least, incomplete. Moreover, the equivalents given often fail to take account of Canadian usage (Cormier, 2000: 214) and the examples chosen to illustrate the senses do not reflect anglophone and francophone culture in Canada (Roberts, 1990: 78). Even the most recent English-French dictionaries, such as the *Oxford-Hachette* (2001) and the *Collins-Robert* (2002) do not meet Canadian needs, since they generally omit typical Canadian words and senses (e.g. French *camelot* and English *paper carrier*, in the sense of "person who delivers newspapers").

Writing in 1970, Mark Orkin described language study in Canada as "a trackless waste into which none but a few brave pioneers had ventured" (5). Forming part of the study of language, bilingual lexicography in Canada has suffered a similar fate and has remained for many years a largely unexplored area. As was mentioned in Chapter 1, however, Canada has, on the other hand, developed great expertise in bilingual terminology and a large number of specialized bilingual glossaries and lexicons have been produced, most notably by terminologists working both in the federal government's Translation Bureau and in the Office québécois de la langue française. The fact remains, however, that only a very small number of bilingual French and English dictionaries covering the general, as opposed to the specialized, usage of Canadian French and English has ever been produced.

The production of a successful bilingual dictionary usually requires that there has been some preliminary work in unilingual lexicography in the two languages concerned. So before treating the evolution of bilingual lexicography in Canada, in section 3.4.2 below, I will first give a quick review of unilingual lexicography in Canada[14].

3.4.1. Unilingual Lexicography in Canada

While serious lexicographic activity surged ahead in the 18th century in Britain, the United States, and France, this was not the case in either of the Canadian language communities. At that time in Canada, the linguistic affinities of the educated classes were still tied most closely to the "mother country" and Canadians considered that existing dictionaries produced in Britain, the U.S. and France were not only adequate to their needs, but enshrined the proper linguistic standards to which they should aspire.

3.4.1.1 French Unilingual Lexicography in Canada

Although not a dictionary in the true sense, the first "lexicographic" work of note in Canadian French is *Façons de parler proverbiales, triviales, figurées, etc.*, written between 1743 and 1758 by Pierre-Philippe Potier, a Belgian Jesuit priest working as a missionary in New France. Potier recorded approximately a thousand words and expressions he encountered in

[14] A more detailed presentation of Canadian-produced English and French unilingual and bilingual dictionaries can be found in "Canadianisms and their treatment in dictionaries" (S. Hamilton, 1997).

conversation (mostly French, but also a few Indian words), sometimes giving examples of usage, but making no prescriptive statements whatsoever on the lexical items included. Potier's work provides invaluable information on the French spoken in Canada on the eve of the British conquest of French North America.

Although the Conquest in 1763 effectively ended all linguistic ties with France, the opinion prevailed, supported adamantly by the language purists, that Canadian French was a mere dialect of "proper" French and as such, not worthy of recording in a dictionary. The purists, "les clercs", maintained that good dictionaries of "gallofrançais" were all that was required and indeed all that should be consulted in order to uphold the standards of "correct" French.

It was not until 1810, therefore, that Canada produced its first lexicographic work recording Canadian French. Jacques Viger's *Néologie canadienne ou Dictionnaire des mots créés au Canada* included three types of words: words created in Canada, words spelt or pronounced differently in Canada, and foreign words. Viger's presentation is typical of the purist ideology that has characterized, even bedevilled, French Canadian lexicography. In common with many other subsequent French Canadian lexicographic works[15], Viger generally adopts a prescriptive rather than a descriptive approach to language, producing a dictionary that does not document Canadian usage in an objective manner, but that seeks to judge whether each lexical item treated is acceptable or unacceptable with reference to the standards of European French and, in most cases, to condemn the Canadian form.

The founding of the Société du Parler français au Canada in 1902 marked what Orkin describes as "the beginnings of a serious study of the French language in Canada" (1971: 14). The work of the Société culminated in the publication in 1930 of the *Glossaire du parler français au Canada*, the first comprehensive study of "les mots et les locutions en usage dans le parler de la Province de Québec" (ix). In that it does not include any prescriptive statements, the *Glossaire* displays a much more objective attitude to the recording of the particularities of

[15] The other principal French lexicographic works produced in Canada in the 19th century were Maguire's *Manuel des difficultés les plus communes de la langue française adapté au jeune âge et suivi d'un recueil de locutions vicieuses* (1841), Gingras' *Manuel des expressions vicieuses les plus fréquentes* (1867), Caron's *Petit vocabulaire à l'usage des Canadiens-français* (1880), Dunn's *Glossaire franco-canadien du français au Canada* (1880), and Manseau's *Dictionnaire des locutions vicieuses du Canada avec leur correction suivi d'un dictionnaire canadien* (1881).

Canadian French and as such is recognized as "une pièce maîtresse du patrimoine lexicographique québécois et un événement capital dans l'histoire du français au Québec" (Mercier, 2002: 395). In subsequent years, a series of French dictionaries produced in Canada straddled the prescriptive-descriptive divide to a greater or lesser degree and were either acclaimed or blamed by the critics—probably in equal measure—whichever side they chose!

However, by the 1960s, there was a clear acknowledgement of the need to document French Canadian usage without adverse comment and without attempting to correct it, and by the time of the publication of the *Dictionnaire du français plus* (a "general" Canadian French dictionary) in 1988 and the *Dictionnaire québécois d'aujourd'hui* in 1992, attitudes had advanced sufficiently that a geographic label was no longer attached to features of Canadian usage, but instead to the particularities of the French language outside Canada. Thus, in the *Dictionnaire québécois d'aujourd'hui*, *magasinage* and *séchoir*, for example, have no geographic label, whereas *shopping* and *sèche-cheveux* are labelled "France". French lexicography in Canada was finally able to recognize Canadian French as its language of reference, its own norm, and accept that, after almost 400 years on American soil, "le français d'Amérique . . . constitue un ensemble linguistique légitime" (*Dictionnaire québécois d'aujourd'hui*, 1992: xx).

Although it first appeared in 1988, the same year as the *Dictionnaire du français plus*, the *Multidictionnaire de la langue française*, a Canadian-produced unilingual French dictionary, now in its fourth edition, tends to be more prescriptive than either the *Plus* or the DQA. The *Multi* aims to present a picture of overall French usage, whilst giving special prominence to the "emplois lexicaux propres au français du Québec" (*Multidictionnaire de la langue française*, 2000: xv) and making it a policy to indicate to users "les formes fautives (anglicismes, calques, impropriétés, etc.)" (ibid.: xvi).

Research in Canada in French lexicography has accelerated and gained considerable ground over the last couple of decades and two important lexicographic projects are presently underway in Quebec. One, the Trésor de la langue française au Québec project produced the first

edition of its dictionary in 1998; the second, the *Dictionnaire québécois*, which is in preparation at the University of Sherbrooke, is promised for the end of 2006.

3.4.1.2 English Unilingual Lexicography in Canada

In marked contrast to the dynamism of lexicographical activity in Canadian French, an interest in lexicography was much slower to develop in English-speaking Canada. This is for the most part due to the fact that, because French has been a minority language in Canada for more than 200 years, the study of Canadian French has often stemmed more from the need to assert an ethnic, cultural, or political distinction than from a desire to conduct purely linguistic research. English-speaking Canadians, on the other hand, have always felt confident about their linguistic heritage, reassured by strong linguistic ties with the English-speaking world in Britain and the United States. This sentiment led not only to the conviction, as mentioned in 3.4.1, that British- and U.S.-produced dictionaries were quite adequate to the needs of Canadians, but also to a kind of indifference to the study of English usage in Canada. From the middle of the 19th century onward, only a handful of papers on Canadian English were written by amateur linguists, either with the aim of correcting what they saw as the growing corruption of the language by Americanisms or to record distinctive features of regional Canadian usage.

Although the establishment of the Canadian Linguistic Association in 1954 at last provided a focal point for research in Canadian English, progress was still slow: a project launched in 1930 to publish *The Linguistic Atlas of the United States and Canada* had, by 1965, produced only 75 field records for the whole of Canada (Avis, "Problems in the Study of Canadian English", 1978: 11). And while noting that same year that many Canadians, especially schoolteachers, had become dissatisfied with imported dictionaries because they realized that "dictionaries prepared for Americans or Britishers are not fully appropriate to Canadians" (ibid. 9-10), Avis was also forced to conclude that a project to produce a dictionary of Canadian English would not be "economically feasible" (ibid. 10) until Canada's population increased.

However, a start towards the preparation of truly Canadian dictionaries was made in the period from the late 1960s onward by the publication of a number of Canadianized versions of American English dictionaries (see Chapter 2, 2.2.1). To produce the *Dictionary of Canadian*

English series, for example, the Thorndike-Barnhart American dictionaries were Canadianized for spelling and pronunciation and a selection of lexical Canadianisms added. While the *Winston Dictionary of Canadian English* (1970) and the *Penguin Canadian Dictionary* (1990) were produced entirely in Canada, their compact size allowed only a partial coverage of Canadian English. It was only in 1998, therefore, with the publication of the *Canadian Oxford Dictionary* that a more extensive coverage of Canadian English usage was eventually provided. The *Canadian Oxford* is now accepted as the most thorough and up-to-date account of Canadian English recorded in dictionary form. A second revised edition was published in 2005.

3.4.2. Bilingual English and French Lexicography in Canada

The first significant event in the history of bilingual English and French lexicography in Canada was the publication in 1905 of Sylva Clapin's *Nouveau dictionnaire français-anglais et anglais-français*. Having already produced two unilingual dictionaries, the *Dictionnaire canadien français* in 1894 and the *New Dictionary of Americanisms* in 1902, the Quebec linguist and short-story writer took an earlier bilingual English and French dictionary compiled by Thomas Nugent[16] and adapted it (Clapin's title describes it as a "nouvelle édition, revue, corrigée et considérablement augmentée") to produce his own bilingual dictionary. Clapin's *Nouveau dictionnaire français-anglais et anglais-français* is an important landmark in the history of Canadian lexicography because, although it claims to cover Americanisms rather than Canadianisms, it acknowledges that North American English is different from British English. The dictionary includes, therefore, not only English words and expressions used in North America with "un sens différent de celui qui s'y attache en Angleterre" (Préface: vi), but also words such as *carpet-bagger* and *caucus*, which refer to specifically North American realities, and which may be "introuvables dans les dictionnaires anglais" (ibid. vii).

Over half a century elapsed before another Canadian French-English, English-French dictionary was produced. Jean-Paul Vinay et al.'s *Canadian Dictionary*, published in 1962, was the result of seven years' research at the Lexicographic Research Centre of the University of Montreal. In his Preface, Vinay claims that the dictionary contains not only "the rules and usage

[16] Nugent's *New Pocket Dictionary of the French and English Languages* was first published in London in 1767 and according to the 1895 *Dictionary of National Biography* was "frequently reprinted and redacted". A copy of Nugent's dictionary is catalogued as being in Thomas Jefferson's personal library.

which reflect international standards of French and English", but also "direct guidance on the terminology and style which are peculiar to the French and English of Canada" (viii). The publication of the *Canadian Dictionary* was, therefore, a significant event in the history of bilingual lexicography in Canada and aroused much comment. Although a f–w critics – the majority francophone - were unhappy with the inclusion of some examples of Canadian usage that they considered incorrect (for example, *fournaise* in the sense of "furnace"), most welcomed its appearance on the scene as "the first important and good study of Canada's two official languages and a fine tool for any punctilious translator" (Petrin, 1962). However, being a concise dictionary of approximately 20,000[17] entries, the *Canadian Dictionary* offered only limited coverage of the two languages and attempts to update and enlarge it in the 1970s and early 1980s were unsuccessful. The dictionary is now out of date and has long been out of print.

Since 1962, only two bilingual dictionaries of note have been produced in Canada and because both are dictionaries of Canadianisms only rather than of Canadian usage as a whole, they offer only partial coverage of the two official languages. In addition, because they are both unidirectional only (French-English), neither of these two more recent bilingual dictionaries provides a tool to serve all Canadians in their daily life.

Fully 20 years elapsed after the publication of the *Canadian Dictionary* before the first of these dictionaries appeared: Léandre Bergeron's *Québécois Dictionary* (1982), a French-English dictionary, is an adapted and abridged version of the same author's unilingual *Dictionnaire de la langue québécoise* (1980). In his introduction to the bilingual version, Bergeron expresses his hope that his dictionary of French Canadianisms will provide English Canadians (as well as other English speakers) with a "handbook to explore Quebec, to understand the Québécois" (xi) and to discover, through the differences between the two languages, the common ground, "the human aspects that we all share" (xi). Bergeron clearly states that the source material for his dictionary was his day-to-day conversations with "mechanics . . . shoppers . . . friends . . . and neighbours" (xi) and, in aiming to capture in the pages of his dictionary "the vitality of Québécois speech"

[17] Because the information given in dictionaries on the number of entries they contain is often misleading or inaccurate, the numbers of entries given in this study for any particular dictionary have been calculated by counting the number of entries on 10 pages, in order to arrive at an average number of entries per page, and by then multiplying that average number by the total number of pages in the dictionary. The figures are approximate because in one dictionary, an "entry" may have derivatives embedded within it, whereas in another, derived forms, such as compounds, for example, may be treated as separate entries.

(x), he produces a dictionary with an almost exclusive emphasis on oral – and often very informal – usage. Although a useful aid to the comprehension and translation into English of selected lexical items of Quebec vernacular speech, the limited scope of this small dictionary of approximately 5,000 entries means that it in no way meets the linguistic needs of Canadians seeking to understand and communicate effectively in both official languages in written as well as spoken form.

The second and later bilingual dictionary produced in Canada is Sinclair Robinson and Donald Smith's *Dictionary of Canadian French* published in 1990, a revised and expanded version of their *Practical Handbook of Quebec and Acadian French* (1984). This dictionary is organized thematically into 33 different subject fields and the information is presented in three columns, the first giving the Canadian French headword, the second the equivalent used in international French and the third the English equivalent. It is essentially a decoding dictionary, an aid to comprehension, rather than a tool to provide equivalents that would work in translation. The fact that the dictionary is unidirectional naturally reduces its usefulness and the great need for a bidirectional dictionary is felt in one reviewer's rather understated comment that "an alphabetized list for the anglophone with Quebec equivalents would be beneficial" (Ainsworth, 1992: 251).

3.5 Conclusion

This analysis has shown that there are distinct and recognizable features of the lexicon, the grammar, and the orthography of Canadian French and English that distinguish them from varieties of French and English spoken elsewhere in the world. Furthermore, it is clear from the review of lexicography that no up-to-date bilingual French and English dictionary produced in Canada is currently available that records these distinctive features. The Bilingual Canadian Dictionary Project seeks to redress this situation by producing a dictionary that records both these distinctive Canadian elements as well as those features that Canadian English and Canadian French share with varieties of the languages used elsewhere in the world. It is envisaged that the Bilingual Canadian Dictionary will, therefore, be truly "représentatif de la communauté linguistique et culturelle" (Clas, 2002: 9) in Canada.

4. THE TREATMENT OF REGIONALISM IN DICTIONARIES

The first part of this study has provided answers to two questions posed in the title by describing *what* the Bilingual Canadian Dictionary Project is and *how* the Dictionary is produced. By showing that there is no up-to-date bilingual French and English dictionary of Canadian origin currently available that records the distinctive features of Canadian French and Canadian English, Chapter 3 has provided part of the answer to the question of *why* a bilingual Canadian dictionary is needed. There are, however, as has been noted, recently revised editions on the market of bidirectional English-French dictionaries produced in Europe. In order, therefore, to complete the answer to the third question, *why?*, this study examines the handling of regionalism in general and of Canadian usage in particular in two of these dictionaries: the sixth edition of the *Collins-Robert French Dictionary* published in 2002 (ISBN 0-00-710526-6) and the third edition of the *Oxford-Hachette French Dictionary* (ISBN 0-19-860363-0) published in 2001. The two dictionaries, both of which are the result of Franco-British collaborative projects, are designated here by the abbreviations CR and OXHA. The way in which the BCD handles regionalism in general and Canadian usage in particular will then be examined. The *Collins-Robert* and the *Oxford-Hachette* have been chosen for this comparative study because they are long-established and top-selling bilingual English-French dictionaries with a high international reputation. CR and OXHA are of comparable size to the BCD and both dictionaries claim in their front matter to provide extensive coverage of North American usage. CR and OXHA are the principal bilingual dictionaries used by lexicographers in the BCD Project, both being recommended as reliable sources of information (Roberts, 1993: 12). It was decided that for the size and scope of this study, two bilingual English-French dictionaries would provide sufficient material for comparative purposes.

Firstly, therefore, this chapter identifies the features that play a significant role in the successful handling of regionalism in dictionaries. A framework of analysis based on these features is then used to determine the policy on the treatment of regionalism firstly, of the two European dictionaries (CR and OXHA) according to their front matter, and secondly, of the BCD

according to its policy documents[18]. It should be noted that as the BCD Project is an undertaking of "scientific lexicography" (Hausmann, 1989: xvi), based on a "constant interaction between theoretical aspects ... and practical applications" (Roberts, 1993: 3), its lexicographic policy is extremely well documented, not only in the *Methodology for Research Assistants* (a document of some 45,000 words) and the "Grant Proposal", but in a very large number of articles and graduate dissertations written by members of the BCD team. Indeed, the *Methodology* has been recommended for publication as a "new textbook for lexicography" containing a "set of explicit instructions for dictionary compilers" ("Grant Proposal Evaluation Report": 4, 1). By contrast, comparatively little information is given in the front matter of the current editions of CR and OXHA regarding their lexicographic policy on the treatment of regionalism. The publishers and editors of the two dictionaries may well possess internal documentation regarding their lexicographic policy on the treatment of regionalism, but it is not publicly available. In order, therefore, to assemble as much data as possible, the front matter of all preceding editions of both dictionaries have been consulted. Because of the shortage of information in the front matter and because it is recognized in any case that claims made in dictionary prefaces may not always be substantiated in the entries themselves, a selection of entries from CR and OXHA containing examples of Canadian usage is studied in Chapter 5 and compared with entries in the BCD.

4.1 Significant Features of the Treatment of Regionalism in Dictionaries

Metalexicographers such as Al-Kasimi (1977) and Landau (2001) have identified a number of features that play a significant role in the successful handling of regionalism in dictionaries. Amongst the most important of these are the identification and selection of regionalisms, the use of geographic labels and the handling of equivalents. This section will examine each of these features in turn.

4.1.1. Identification and Selection of Regionalisms

As indicated in Chapter 3, regionalism is linguistic usage that is restricted according to the geographic origin of the language user. When recording examples of restricted usage in a

[18] These policy documents will no doubt form the basis for the preface to the published version of the BCD.

dictionary, Al-Kasimi (1977: 86) specifies two major issues that the lexicographer must address: firstly, the identification and selection of the examples of restricted usage to be included and secondly, the differentiation on the dictionary page of examples of restricted usage from what he terms "standard usage"[19]. In order to collect data on which examples of restricted usage to include in a dictionary and how to mark them, Barnhart (1967) proposed the use of questionnaires, and others have suggested the establishment of a "usage panel", a method used to provide information for the compilation of the *American Heritage Dictionary* (1969). Both methods are open to criticism if either the sample of respondents or examples of usage considered is not sufficiently large or randomized. Today, with the rapid growth and increasing ease of use of information technology, modern lexicography turns almost exclusively to corpora of electronic of texts to provide objective evidence of the choices and combinations of choices that language users make.

When identifying and selecting regionalisms for inclusion in a dictionary, lexicographers are assembling what Sinclair terms a "coherent subset of a language" (Sinclair, "Corpora for lexicography", 2003: 167), that is, "a subset chosen by the application of external criteria to the totality of language behaviour" (ibid. 178). In the case of regionalisms, the external criteria applied could be the presence and frequency of occurrence of any given lexical item in a corpus of texts from a specific geographic region, compared with the absence or presence and comparable frequency of the same item in a corpus of texts produced in a different geographic region. By comparing the content of two or more regional corpora in this way and by evaluating the results with reference to the nomenclatures of existing regional dictionaries, a reliable subset of regional usage can be established.

This subset would normally include both types of regionalism identified in Chapter 3 (section 3.1): firstly, those that may be termed "linguistic regionalisms", that is, words and expressions used in one geographic region to denote realities that are *not* unique to that region; and secondly, those that may be termed "cultural regionalisms", that is, words and expressions used in one geographic region to denote realities that *are* unique to that region. The results of

[19] The differentiation of regional usage on the dictionary page will be dealt with below in the section of this chapter (4.1.2) that deals with geographic labelling.

three separate surveys of the requirements of dictionary users (Barnhart 1962, Béjoint 1981 and Hartmann 1983) showed that, out of a range of up to seven different reasons for which a dictionary is consulted, the highest-ranking activity was using a dictionary to discover the meaning of a lexical item (Svensén, 1993: 14). Regionalisms can certainly be difficult to understand for those from outside a given region and they present special problems in the field of second-language acquisition where "the more idiosyncratic a lexical item is to the specific culture of a language community, the more difficulty it causes for the foreign learner" (Snell-Hornby, 1987: 166). Hartmann's survey of bilingual dictionary use for reading foreign-language text showed that, on the question of which types of words are most often looked up by users, culture-specific words achieved the second-highest ranking (62%) (ibid.). The fact that consulting a dictionary to discover the meaning of a lexical item and consulting a dictionary to look up culture-specific items are the two highest-ranking activities in these surveys confirms the need to include both linguistic and cultural regionalisms in a dictionary.

4.1.2. Geographic Labelling of Regionalisms

It is customary in lexicography to mark with a label lexical items whose usage is restricted to a particular geographic area or areas, since the user may interpret the absence of a regional indicator as a positive instruction to use the lexical item indiscriminately in cultural contexts of whatever kind (Burkhanov, 2003: 108). Hence, a lexicographer's failure to provide indications when they are needed may lead to incorrect usage, possible confusion and even embarrassment. For example, a Canadian asking an Englishman where he could buy *suspenders* ("a pair of straps of usu. elastic material that may be worn over the shoulder and fastened to the waistband of a pair of pants ...", *Canadian Oxford Dictionary*) would be surprised perhaps to be directed to a lingerie store ("*suspender* . . . *Brit.* . . . an elastic strap attached to a belt or corset . . . for holding up women's stockings", *Collins English Dictionary*). The word *suspenders* is marked by a geographic label in both the *Canadian Oxford* and the *Collins English*, although, in these two dictionaries, as in most dictionaries, the vast majority of lexical items remain unmarked by any geographic label.

Geographic labels in lexicography (known in French as "les marques topolectales") are markers of usage intended to specify restrictions on the use of a lexical item according to the

place in which it is used. The *Collins*, while stating that it offers a "wide-ranging and in-depth survey of the English language as it is used in Britain and around the world, . . . from Australia to the United States" (*Collins English Dictionary*, Foreword: ix) and that it uses national and regional labels to identify "words or senses restricted to or associated with a particular country or region" ("Guide to the Use of the Dictionary": xxi), does not define the geographic status of the unmarked lexical items in its pages. One may well infer that, since the *Collins* is a dictionary produced in Britain, its content will reflect predominantly British usage, but without the clear statement of lexicographic policy in the front matter of the dictionary, as recommended by Al-Kasimi (1977: 109) as long ago as 1977, the user cannot be certain whether these unmarked items reflect exclusively British usage or usage common to most (if not all) regional varieties of English.

In fact, the presence or absence of a geographic label in lexicography always highlights a contrast between, on the one hand, usage that is restricted to one specified regional variety of the language and, on the other hand, usage that may be termed "unrestricted", being generalized across several regional varieties. Decisions on whether to attach a label to the example of restricted usage, to the example of unrestricted usage, or to both constitute an important aspect of the policy and planning stages of any dictionary, but these decisions become even more crucial for dictionaries that aim to give more complete coverage of the language of a particular geographic region, such as, for example, of the French language used in Canada.

What types of lexical items do dictionaries label? Most unilingual dictionaries label items that are restricted in usage to a region outside of their own or to smaller areas within the territory in which they are produced. Thus, the *Collins,* a British dictionary, primarily labels words or senses used in the United States or Australia, for example, as well as those used only in a particular region of Great Britain. However, exceptionally, some British words, such as *suspenders*, mentioned above, are also labelled. This is because, as the editors explain in the "Guide to the Use of the Dictionary", the label *Brit.* fulfils a very specific function in their dictionary, being used "mainly to distinguish a particular word or sense from its North American equivalent or to identify a term or concept that does not exist in North American English" (*Collins English Dictionary*: xxi).

The issue of whether to mark lexical items used only in the country where the dictionary is produced is particularly delicate in the case of regional dictionaries, i.e. dictionaries created in regions where the language variety used is not that which is most widespread. Dictionaries published in Canada, for instance, have traditionally labelled Canadian words in French and English, in order to ensure that users recognize these less well-known regional varieties. However, the publication of two French Canadian dictionaries, the *Dictionnaire du français plus* in 1988 and the *Dictionnaire québécois d'aujourd'hui* (DQA) in 1992, marked a significant change in the policy regarding the geographic labelling of Canadian French in unilingual Canadian dictionaries. Since both these dictionaries aim to reflect Canadian usage, their editors argued that logically they do not need to attach a geographic label to Canadian features, but rather to those features that are *not* part of Canadian usage. Having noted that dictionaries produced in France do not affix a geographic label to examples of usage of the French language that are generalized across the whole of France, but only to "les mots dont l'usage est restreint à une région du territoire français ou à une partie d'un autre territoire de la francophonie" (*Dictionnaire québécois d'aujourd'hui*, 1992: xx), the editors of the *Dictionnaire québécois d'aujourd'hui* decided that their dictionary, having a Quebec point of view, would not attach a geographic label to lexical items "qui constituent les particularismes de la langue française au Québec" (ibid.).

This shift in policy regarding the geographic labelling of Canadian French sparked much debate in Canadian lexicography, a debate that continues to divide lexicographers and metalexicographers into two main camps. The first group, favouring the policy of the *Dictionnaire québécois d'aujourd'hui* and the *Dictionnaire du français plus*, argues that if a lexical item is used freely in a particular area, such as in French-speaking Canada, or indeed, in France, then it does not require a geographic label. It is the opinion of the editor-in-chief of the *Dictionnaire du français plus*, Claude Poirier (professor in the Département de langues, linguistique et traduction at Laval University and director of the Trésor de la langue française au Québec[20] project), that if Canadianisms are identified in dictionaries by a geographic label it is as if they have a warning tag attached to them saying "attention! ce n'est pas tout à fait français"

[20] This project is primarily concerned with the production of the *Dictionnaire historique du français québécois* the first edition of which was published in 1998.

(Poirier, 1988: 91). Others of the same opinion are Pierre Auger (linguist, terminologist and present director of the Département de langues, linguistique et traduction at Laval University) and Normand Beauchemin (professor from 1969 to 1993 in the Faculté des lettres et sciences humaines at the University of Sherbrooke and author of several works on linguistics), both of whom, in their foreword to the *Dictionnaire du français plus*, criticize the geographic labelling of regionalisms as if they were "des papillons . . . des curiosités explicitant le pittoresque ... des protubérances du tissu langagier « normal »" (*Dictionnaire du français plus*, 1988: ix). Therefore, neither the *Dictionnaire québécois d'aujourd'hui* nor the *Dictionnaire du français plus* labels Canadian usage, but both affix the geographic label "France" or "surtout en France" to lexical items that are used exclusively in France, very often with the specific intention of highlighting a contrast between European and Canadian usage. At the headword *tu*, for instance, the *Dictionnaire québécois d'aujourd'hui* describes the form "tu viens ?" as being "plus courant en France" than "viens-tu ?".

The second group of lexicographers and metalexicographers considers that a geographic label is of enormous value to the user, agreeing with De Villers that such a label "ne constitue pas une marque d'infamie" (1998: 159). Not surprisingly, given its much earlier publication date of 1962, the *Dictionnaire canadien*, whose stated aim was to offer "guidance on the terminology and style which are peculiar to the French and English of Canada" (*Dictionnaire canadien*, Preface: viii), employed a more traditional approach to geographic labelling, attaching the marks © and [Fr.] respectively to "les mots ou expressions d'usage exclusivement canadien ou français (européen)" (*Dictionnaire canadien*, Introduction: xxix). The *Multidictionnaire de la langue française* (2003), a Canadian-produced French dictionary of which De Villers is editor, while acknowledging the difference between the linguistic regionalism and the cultural regionalism, chooses to mark only the former (*babillard* and *banc de neige*, for example), leaving unmarked "les désignations de réalités qui nous appartiennent, mais qui sont ainsi dénommées dans l'ensemble de la francophonie ... ex. : acériculture, acériculteur, acéricultrice" (*Multidictionnaire de la langue française*, Introduction: xv). Jean-Claude Corbeil[21] supports this

[21] Jean-Claude Corbeil is a former professor of linguistics at the University of Montreal, a former director of the Office québecois de la langue française and former secretary general of the Conseil international de recherche et d'étude en linguistique fondamentale et appliquée. He is currently the editorial director for reference of the publishing company Québec Amérique and editor-in-chief of the company's *Visuel* range of illustrated French dictionaries.

approach (Corbeil, 1994), considering the geographic label to be essential in a French Canadian dictionary in order to properly identify "des faits culturels propres à la France" (e.g. *arrondissement*, *département*, *préfet*), to identify different senses attributed in France and Canada to the same word (e.g. *bleuet*, *déjeuner*, *scotch*) and to identify lexical items that refer to the same concept (e.g. *bas/chaussette*, *efface/gomme*, *mitaine/moufle*).

Geographic labels may be divided into three types, each type denoting a different degree of restriction. The first and perhaps the commonest type identifies a word or sense restricted to or associated with a particular country: *Austral.*, *Brit.* and *U.S.*[22], for example, commonly identify English words and senses used in Australia, Great Britain and the Unites States. A second type of geographic label denotes a greater degree of restriction, being affixed to a lexical item that is used in a specific region of a single country, for example, *en Suisse romande*, *dans le Nord de la France*, etc. A third type of label, at the opposite end of the range, designates a larger geographic area covering more than one single country, for example, *Am* or *NAm* to denote North America. A geographic marker in lexicography may be an abbreviation (e.g. *N.Z.* for New Zealand), a single word (e.g. *Belgique*), a phrase (e.g. *northeast English dialect*), a symbol, such as © (used by the *Dictionnaire canadien* to denote a Canadianism) or a combination of these (e.g. *Brit dialect*). Geographic labels can be very general, such as *dial.* for "dialectal (*Canadian Oxford Dictionary*) or "région." for "régional", or very specific, for example, "en Acadie" (DQA) or *Central Scot. urban dialect* (*Collins English Dictionary*). Whichever labels are used, in order to be of maximum benefit to the user, "the signification of each usage label employed in the bilingual dictionary must be precisely defined in the front matter" (Al-Kasimi, 1977: 88). Zgusta goes so far as to assert that the correct selection of usage labels and their proper use is "one of the most important tasks of the lexicographer" (1971: 333).

Having selected the appropriate labels and defined and stated their use in his dictionary, the lexicographer must also decide where to place them within the dictionary entry. The position of usage labels has varied over the years according to prevailing tendencies and the preference of individual lexicographers. However, James Iannucci observed in the late 1960s (1967: 204) that

[22] All examples of geographic labels are taken in English from the *Collins English Dictionary* (2000) and in French from *Le Nouveau Petit Robert* (2003) unless otherwise specified.

"explanatory matter in dictionaries conventionally refers to what precedes" and Zgusta (1971: 333) further clarified this question with special reference to bilingual lexicography by stating that, as a general rule, whether attached to the source-language or the target-language item, a usage label "pertains to what precedes" (ibid.).

4.1.3. Equivalents

Having tackled the problems associated with the identification and selection of linguistic and cultural regionalisms in the source language and with the geographic labelling of these regionalisms, for both the French-English side of the dictionary and the English-French side, the bilingual lexicographer must also address the particular problems posed by the need to establish equivalence between source- and target-language items. As Zgusta explains, "the way the lexicographer handles the equivalence and how he presents it varies largely with the different intentions and other properties of the planned dictionary" (1971: 315). The treatment of equivalents will differ depending on whether, for example, the dictionary is intended primarily for decoding or encoding, for first- or second-language users, for beginners or for advanced language users. Similarly, a decision to include regionalisms in the source-language nomenclature of a dictionary will have an impact upon target-language equivalents. Before presenting an examination of the special issues affecting the treatment of equivalence in a dictionary that includes regionalisms, an overview from a semantic perspective will be given of the main principles of equivalence, as established by linguists in the second half of the 20th century.

Zgusta defines an equivalent as "a lexical unit of the target language which has the same lexical meaning as the respective lexical unit of the source language" (1971: 312), while acknowledging that absolute equivalence between source- and target-language items cannot always be achieved because different languages organize concepts in different ways. In acknowledging this lack of correspondence between languages, this "anisomorphism", as he calls it, Zgusta (1971: 294) is developing a theory of translation suggested as far back as 1813 by Friedrich Schleiermacher in his essay "On the different methods of translating"[23]. Roman Jakobson and Eugene Nida, writing in the late 1950s and early 60s, were also prominent amongst

[23] Some of Schleiermacher's concepts were further developed in the 1930s by the American linguists Sapir and Whorf.

translation theorists who recognized that human beings dissect the natural world and organize it into concepts largely according to the codified patterns of their own language.

Jakobson made it clear that, although "there is ordinarily no full equivalence between code-units" (Jakobson, 1966: 233), translation is still possible because "all cognitive experience and its classification" (ibid. 234) can be conveyed in any existing language by using different kinds of translational process. Linguists have identified varying degrees of "translatability" of lexical items and characterized them in a variety of different ways. Mary Snell-Hornby (1987: 166), for example, describes translatability, or what she terms "semantic overlap" in interlingual relationships, as a "gradual transition" from complete equivalence at one end of the scale, through "equivalence with discrimination" to "partial over-lapping" and, finally, at the other extreme, to definition.

For the practical purposes of bilingual lexicography, Zgusta (1971: 319) identifies only two types of equivalent, the translational (or insertible) equivalent and the explanatory (or descriptive) equivalent. Pursuing Zgusta's argument, Al-Kasimi defines the translational equivalent as a lexical item that can be "immediately inserted into a sentence in the target language" (ibid. 60), whereas the explanatory equivalent is one that "cannot always be inserted into a sentence in the target language" (ibid.). The English noun *boy*, for example, can usually be rendered in French by the translational equivalent "garçon", as in this free combination in the entry for *boy* in the BCD: he was a good boy il était un bon garçon. However, "état de garçon", offered as an explanatory equivalent of the English noun *boyhood*, is a paraphrase that cannot be readily inserted into all the target-language constructions a user may need to produce. The phrase "during his boyhood", for example, could not be translated idiomatically into French using the explanatory equivalent "état de garçon" and is probably better rendered by a phrase such as "quand il était garçon" or "pendant sa jeunesse". Although an explanatory equivalent may provide more information to the dictionary user by, as its name suggests, *explaining* the meaning of the source-language item (here, for example, the explanatory equivalent could be paraphrased as "state of being a boy"), it has the disadvantage of not always being insertible into the target-language construction. An explanatory equivalent may, of course, be "une traduction consacrée" (Roberts, 1997: 38), a translation that is "standardized by acceptance and use in the language"

(Al-Kasimi, 1977: 60). This is often the case with fixed expressions, the specific vocabulary of which may vary from culture to culture, while conveying the same meaning. For example, the French fixed expression *on a l'âge de ses artères* cannot be rendered accurately in English using any combination of "age" and "arteries", the normal translational equivalents of *âge* and *artères*. But valid equivalence is satisfactorily achieved using the explanatory equivalent "you are only as old as you feel", which accurately conveys the meaning of the French expression.

By contrast, an explanation or explanatory gloss – which is not an equivalent at all (Zgusta, 1971: 319) – is required when no translational or explanatory equivalent adequately conveys the exact or full sense of the source-language item. This is often the case with cultural or geographic realities, where a source-language item may be so specific to a particular culture that the inclusion of an explanatory gloss is of immense value to target-language users in clarifying an element of meaning that may not be evident from the equivalent alone. For example, the translational equivalent of *métropolitaine* is given in the BCD as "Metropolitan Boulevard", but the inclusion of an explanation, "part of the Trans-Canada Highway running through Montreal", provides supplementary information allowing the user who is not familiar with that Canadian reality to pinpoint exactly the meaning of the word. The difference between an explanatory equivalent and an explanation is, as Zgusta (1971: 319) puts it, "not absolutely sharp", in the same way that it is sometimes difficult to differentiate clearly between translational and explanatory equivalents, but the equivalents in a bilingual dictionary of culture-specific items may be best defined as "rather a matter of encyclopedic information than of interlingual relationships" (Snell-Hornby, 1987: 166).

Although there are various ways of handling equivalence and although the treatment of equivalence will differ depending on whether a particular dictionary is intended primarily for decoding or encoding, for first- or second-language users, for beginners or for advanced language users, metalinguists are agreed that no bilingual dictionary can possibly guarantee that it will always offer all suitable equivalents for every possible application.

The handling of equivalence presents special problems in a dictionary that includes regionalisms in the nomenclature. Bilingual lexicographers try always to provide equivalents that

match the source-language item as closely as possible in terms of meaning, of course, but also in terms of register (formal, informal, slang, etc.), currency (archaic, obsolete, old-fashioned, etc.) and other aspects of usage (ironic, pejorative, figurative, etc.). In some special cases, like that of an officially bilingual country such as Canada, it may even be possible to match a source-language regionalism with a target-language equivalent that has not only the same meaning and is of the same register, but is also used in the same geographic region. The English Canadianism *dump of snow,* for example, signifying a heavy snowfall (*Canadian Oxford Dictionary*), is perfectly matched by the French Canadianism "bordée de neige" as far as all three aspects of meaning, register and regional distribution are concerned.

Although, as has been noted above, cultural regionalisms usually require an explanatory equivalent in order to be clear to those outside the language community concerned, in some cases a cultural regionalism in the source language can be perfectly matched by another in the target language. The French Canadianism *carte-soleil,* for example, is perfectly rendered by the English Canadianism "health card" (*Canadian Oxford Dictionary*) which serves in all contexts as a fully insertible equivalent, though a brief explanation of the allusion to the "soleil" pictured on the Quebec provincial health card would add a bonus of information for dictionary users not familiar with this aspect of Canadian culture.

It is interesting to note that even those rare cases where a perfect match does exist between a source-language regionalism and a regionalism in the target language can create further problems for the lexicographer who is trying to present an accurate representation of regional usage. If all the senses of a polysemous headword can be accurately rendered in the target language by the same single equivalent, bilingual dictionaries often choose to avoid repetition and thus save space by conflating all the semantic divisions into one. The drawback of this method is that if one of the senses is regional, its existence and the existence of its regional equivalent may go unremarked. All the meanings of the English noun *confederation*, for example, can be rendered in French by the equivalent "confédération", but if this is all the information that is given in the dictionary entry, then the existence of the two semantic Canadianisms, one in English and one in French, is hidden.

However, in most cases, it is not possible to match one regionalism in the source language with another in the target language and so an equivalent has to be sought in the standard variety of the target language. In these cases, even more than in the case of most cross-language equivalents, the match of source- and target-language items is unlikely to be perfect for a variety of reasons connected with the anisomorphism of language (as discussed above) that manifests itself in three main areas. Firstly, in the semantic area, where the meanings covered by each item in its respective language (and culture) may not perfectly coincide with one another. In fact, it is often difficult to determine the precise meaning of a regionalism. The French Canadian noun *bébelle*, for instance, has many different meanings ("toy", "gadget", "knick-knack", "thing") depending on the context in which it is used. The meaning of this kind of catch-all word is often difficult to pin down with precision, even sometimes for native speakers. Moreover, unilingual dictionaries of regionalisms do not always agree on the precise meaning of the lexical items they include and their explanations and examples often present differing individual interpretations. In a case like *bébelle*, no one English equivalent can provide a perfect semantic match for the source-language item and it can only be rendered by a range of different equivalents that are either more generic or more specific and that have to be differentiated in the bilingual dictionary by sense indications, as can be seen in this excerpt from the BCD entry for this word:

1 (*jouet*) toy . . .

2 (*chose au sens vague*) thing . . .

3 (*chose plus ou moins utile, de peu de valeur*) knick-knack; (*objet clinquant*) trinket . . .

Secondly, in the stylistic area, even if the meaning of the regionalism can be clearly identified and a semantic equivalent found in the target language, lexicographers may have problems finding a target-language equivalent whose register matches that of the source-language item. This is a special problem that has to be tackled when providing equivalents for French Canadianisms, a high proportion of which are informal in register. It often impossible to find an equivalent that is of the same register. The French Canadian fixed expression *en bedaine*, for example, is informal in register, but can only be rendered in English by equivalents that are standard in register, such as "bare-chested", "shirtless", "stripped to the waist" or "without a shirt on". Sometimes a match of register can be achieved for some contexts but not for others. In the

case of the informal French Canadian adjective *bâdrant*, for example, "pesky" provides a good informal equivalent to collocate with certain nouns (e.g. *une mouche bâdrante* "a pesky fly"), but in order to provide an equivalent that will collocate with other nouns, it is necessary to have recourse to an equivalent from the standard register such as "irritating" or "annoying" (e.g. *un bruit bâdrant* "an irritating noise"). Neither is it only the matching of register that poses a problem for the bilingual lexicographer: similar difficulties are encountered when attempting to match lexical items in different languages with regard to currency and usage.

As has been shown, it is rare to be able to match a source-language regionalism perfectly in register, currency, usage, etc. with an equivalent from the same geographic region and often difficult to do so with an equivalent from the standard variety of the language. If an equivalent with both the same semantic coverage and the same register as the source-language regionalism can be found, it is often in a third area, that of geographic distribution, that a further "imbalance" may arise when the source-language item has a narrower or wider geographic distribution than its target-language equivalent. The French noun *autoroute*, for example, is used in all francophone countries, but its English equivalents, "highway", "expressway" and "motorway", are used in different parts of the English-speaking world, the first two being current throughout North America and the third only in Britain.

4.2 Significant Features of the Treatment of Regionalism in Dictionaries According to Dictionary Prefaces and Policy Documents

The framework of analysis established above will now be used to determine the policy on the treatment of regionalism, firstly, of the two European dictionaries (CR and OXHA) according to their front matter and secondly, of the BCD according to its policy documents.

4.2.1. Identification and Selection of Regionalisms in CR, OXHA and BCD

4.2.1.1 Identification and Selection of Regionalisms in CR and OXHA

When establishing their nomenclature, the editors of all general bilingual dictionaries are obliged to make a selection from the millions of lexical items in both languages. The size of the

intended dictionary obviously influences this selection, but, ideally, it is made in accordance with the needs of the targeted user group and the purpose for which the dictionary is intended (see Chapter 1, section 1.2). Top-selling modern French and English bilingual dictionaries like CR and OXHA seek to target a wide audience in order to maximize their sales in as many English- and French-speaking countries of the world as possible, but most particularly in Europe and North America. To this end, these dictionaries include a number of examples of regionalism in both French and English.

4.2.1.1.1 In CR

It is perhaps because of his "ancêtres acadiens" that Paul Robert, in the opening paragraph of his preface to the first edition of CR (1978), gives such prominence to the fact that this new dictionary is "une œuvre faite en commun dans les deux vieux pays [France and Great Britain] pour aider à la communication entre les anglophones et les francophones de l'ancien et du nouveau monde" (1978: vi) and that it does not restrict itself to covering only European English and French but that it includes also "les emplois courants de l'anglais d'Amérique, ainsi que les termes les plus répandus du français du Canada" (ibid.). What is somewhat strange is that the importance of North American varieties of English and French is not highlighted in this way in the English preface to the same edition which, although it informs the reader of the dictionary's aim to "promote accurate and easy communication between the speakers of two great world languages" (1978: vii), has a far more general focus than its French counterpart, merely offering the dictionary to "people of each nationality [French and English] who are interested in the language, literature and culture of the other" (ibid.).

The preface to the second edition of CR informs the user that more space has been devoted to American English "dans les deux parties de l'ouvrage" (1987: ix) and that, of the new items added, a substantial number are "American expressions" (ibid.). The preface to the third edition of CR states that the dictionary presents "une image fidèle du français, de l'anglais et de l'américain tels qu'ils sont parlés dans le monde à l'aube du vingt-et-unième siècle" (1993: xiii). The contribution of Guy-Jean Forgue[24] on the "American language" is acknowledged in both the

[24]A former professor of Yale University and of the Sorbonne, Guy-Jean Forgue is a consultant lexicographer and co-author of *La langue des Américains* (Paris: Aubier Montaigne, 1972).

second and third editions and that of Guy Rondeau[25] on French Canadian usage in the first three editions (1978: xxvii).

A section of the introduction to the first and second editions of CR entitled "The scope and nature of the language treated" (1978: ix and 1987: xi) states that the spread of language treated in the dictionary "is not confined to British English or metropolitan French: American English and Canadian French are given due attention" (ibid.). In the third edition (1993), this statement has been modified to provide the information that "North American English is described in great detail and its specific usages are pinpointed on both sides of the dictionary" (1993: xv) and that "the most common Canadian, Belgian and Swiss French terms are also covered" (ibid.).

The first edition of CR includes a section entitled "Using the Dictionary" in which the following example is given: honor ... *n US* = honour. However, it is not made clear that this example is meant to inform the user that in this dictionary American spelling variants that are not alphabetically adjacent are cross-referred to the British form.

The Introduction to the fifth and sixth editions of CR informs the user that the dictionary contains entries "giving details about institutions, culturally significant events, traditions and customs in French- and English-speaking countries" (1998: xv) and CR also uses a system of "Cultural Notes" that are described as "extra information on culturally significant events, institutions, traditions and customs that cannot be given in an ordinary translation or gloss" (1998: xxiv). This information is presented in the form of boxed notes following the relevant entry. At the English headword *kissagram*, for example, the italicized explanatory equivalent *baiser télégraphié* is followed by a "cultural note" presented in a box and consisting of a five-line explanation or definition, also in italic type. However, the front matter of CR provides no

[25] Guy Rondeau is a former director of the Department of Linguistics and Modern Languages at the University of Ottawa and was also formerly a professor in the Department of Languages, Linguistics and Translation at Laval University, secretary general of the Association internationale de terminologie, director of the Groupe interdisciplinaire de recherche scientifique et appliquée en terminologie, a member of the Commission de terminologie of the Office québécois de la langue française and chairman of the Canadian consultative committee for ISO/TC 37. He has written many works on French linguistics, applied linguistics, computational linguistics and terminology, most notably his *Introduction à la terminologie* (Montreal: Centre éducatif et culturel, 1981).

information on how these cultural regionalisms are identified and selected for inclusion in the dictionary.

4.2.1.1.2 In OXHA

Since OXHA is a much more recent arrival on the dictionary scene, with only three editions (1994, 1997, 2001) published to date, its front matter provides even less information than CR on the dictionary's lexicographic policy on regionalism. All three editions of OXHA state that the dictionary has "a wide coverage of North American as well as British English" (Introduction: ix) and that it was prepared with the help of "specialists in the French- and English-speaking worlds" (1994: viii). The introduction to the current (third) edition states that the French nomenclature of the dictionary "prend en compte le français des pays francophones autres que la France, qu'il s'agisse d'entrées à part entière ou d'acceptions" (2001: ix). In a section entitled "The structure of English-French entries" (1994: xlix), the first edition of OXHA gives two examples of entries containing a British/North American variant spelling (*meagrely/meagerly* and *mouldy/moldy*), but includes no information at all on how these variant forms are identified or selected for inclusion.

As far as the identification and selection of cultural regionalisms in OXHA are concerned, the dictionary's Introduction states that it incorporates "un grand nombre de notes culturelles qui abordent des points spécifiques de la civilisation et des institutions de la France, du Royaume-Uni et des États-Unis" (2001: x). But although a list is included at the end of the dictionary of the headwords for which cultural notes are given (e.g. *Groundhog Day*, *soap operas*, *Westminster*, *Élysée*, *Garde républicaine*, *pompiers*, etc.), no information is provided on how this material was identified or selected for inclusion.

4.2.1.1.3 Conclusion

Naturally enough, in their aim to appeal to as wide an international readership as possible, the editors of CR and OXHA attempt in their front matter to give the impression that their dictionaries include a large number of examples of regional usage, with particular emphasis on North American English. However, such expressions as "the most common terms", "les emplois courants", "les termes les plus répandus", "une image fidèle", etc., used in the context of

regionalisms, are subjective opinions that do not provide any solid information, but only raise questions in the reader's mind regarding their precise meaning. Which lexical items and senses are the "most common", which are "les plus répandus", are questions that can only be answered confidently by analysis of a large corpus of texts, a corpus that can be "trusted to represent the way language is used by those people whose usage one is interested in describing" (Landau, 2001: 331). It is surprising that, although OXHA was the first corpus-based bilingual French and English dictionary and despite the fact that CR has been prepared using a "databank of real written and spoken language" (CR 1993:xiii) since its third edition, neither dictionary provides any explicit information in its front matter on whether or how their corpora are used to identify regional words and senses for inclusion.

4.2.1.2 Identification and Selection of Regionalisms in the BCD Project

In the BCD Project, by contrast, setting and stating the parameters for the identification and selection of regionalisms was a priority. The identification of Canadianisms was naturally a prime concern, but a number of non-Canadian regionalisms are also included in the BCD, as explained below.

4.2.1.2.1 Canadianisms

Having first defined a Canadianism (see Chapter 3, section 3.2), BCD lexicographers worked to establish a preliminary list of all potential French and English Canadianisms by comparing unilingual dictionaries and glossaries produced in Canada with those from France, the U.S. and Britain, by comparing the evidence of Canadian and non-Canadian texts in the BCD's comparable corpora (see Chapter 2, section 2.2) and by consultation with experts (Roberts, 1999: 5).

As the BCD is a dictionary that will be consulted by both francophones and anglophones for the purpose of decoding Canadian texts in French and English, it is essential that it should include a large number of Canadianisms in both languages as source-language words or expressions (Roberts, "Methods of bilingual dictionary making: the Canadian experience", 1992: 97). However, limitations of space and time mean, of course, that "the BCD has to make choices, as do all dictionaries" (Roberts, 1999: 6). No general bilingual Canadian dictionary could

reasonably include all possible Canadianisms as well as those features of French and English that are common to Canada and to other French- and English-speaking countries.

Frequency of occurrence in TEXTUM (the BCD comparable corpora) is almost invariably the main criterion used to determine the inclusion in or omission from the Dictionary of any lexical item. In this respect, both linguistic and cultural Canadianisms are subject to the same procedure, being selected for inclusion principally according to the frequency of their occurrence in the BCD corpora. However, not all Canadianisms identified, even those that are fairly frequently used, will appear in the BCD. For example, when assessing an anglicism in Canadian French for inclusion in the nomenclature of the BCD, its frequency of usage in the Project's Canadian corpora is always one of the main criteria used (ibid. 5). However, other aspects of an anglicism may need to be considered. For example, anglicisms in Canadian French that are identical or almost identical in form and meaning to the English word ("anglicismes intacts" or "quasi-intacts" such as *truck* or *muffler*) are "far less likely to figure in the nomenclature" (ibid.: 6) of the BCD than those like *sloche* (an "anglicisme francisé" from "slush") or the interjection *bienvenue!* (a semantic anglicism that has borrowed its meaning from the English expression "you're welcome") that would be more difficult for an anglophone to decipher. Anglicisms in Canadian French that have been linguistically productive, giving rise to new lexical items, are also more likely to be included in the nomenclature of the BCD (ibid.). Thus, the noun *chum* (labelled "anglicisme" in the DQA[26]) is included in the BCD because it has led to the creation of not only such collocations as *chum de bureau* and the compounds *chum de fille* and *chum de gars*, but also to an adjectival form, as in *elle est très chum avec lui.*

The selection of Canadianisms to be included in the French nomenclature is complicated by two additional factors. The first of these is that many French Canadianisms are informal in register, as was shown by Grenon-Nyenhuis (1997: 71) who, in a sample of Canadianisms, identified around 40% of the French lexical items as having some degree of informality ("familier", "très familier", "vulgaire"), compared to only around 5% of the English lexical items. Since the BCD, like any general bilingual dictionary, must make it a priority to present as complete a coverage as possible of standard-register items, it is necessary to limit the number of

[26] Anglicisms are not marked in the BCD (Roberts and Bossé-Andrieu, 2002: 248).

informal lexical items included in the nomenclature. Moreover, as the BCD corpora are composed predominantly of texts reflecting written usage, non-standard elements of the spoken language, such as colloquialisms and slang, are not extensively represented in them. But since the BCD aims to be a descriptive dictionary that presents as objective a record as possible of the French and English languages as used in Canada at the end of the 20th and the start of the 21st century, a dictionary which "ne porte pas de jugement de valeur, mais décrit les usages" (Roberts and Bossé-Andrieu, 2002: 236), non-standard items, which form an integral part of language, cannot be entirely neglected. The BCD includes, therefore, informal Canadianisms such as, for example, *beurrée* in the sense of "grosse somme d'argent" and *sacrer* in the sense of "lancer". Informal Canadianisms in English are also included in the BCD, for example, the adjective *chippy*, which is used to describe either a person who is "short-tempered or irritable" or a hockey game that is "characterized by rough or dirty play". However, as was mentioned above, far fewer such elements exist in English than in French.

The second factor that complicates the selection of Canadianisms to be included in the French nomenclature is the regulatory influence of l'Office québécois de la langue française (OQLF), which makes recommendations concerning anglicisms and other "emplois critiqués". BCD lexicographers take account of these recommendations when assessing a French Canadianism for inclusion in the Dictionary as a source-language item, examining each case individually, but always respecting the Project's own lexicographic criteria, as stated above. For example, while the OQLF acknowledges that "la création de nouveaux noms féminins par l'ajout du *e* au nom masculin en *-eur* a permis de combler certaines lacunes lexicales" (Office québécois de la langue française. "Noms féminins en -eure") and proposes that some of these feminine forms should be retained (e.g. *auteure*, *docteure*, *entrepreneure*, *gouverneure*, *pasteure*, *professeure*, etc.), this organization does not recommend the retention of those new feminine forms which "concurrencent indûment des formes régulières déjà existantes" (ibid.). Thus, in some cases where a feminine form ending in *-euse* or *-trice* already exists, the OQLF describes as "noms en *-eure* non retenus" such forms as *administrateure*, *chauffeure*, *chercheure*, *directeure*, *entraîneure*, *recteure*, etc. The entries in the BCD make no reference to the recommendations of the OQLF and variant forms are included in the Dictionary on the grounds of their frequency in the BCD corpora. In the case of *chercheure* and *chercheuse*, for example,

examination of concordances from the *Presse canadienne française* sub-corpus of the BCD corpus shows that both forms have a very similar number of occurrences in Canadian texts (*chercheure* 44 occurrences and *chercheuse* 53 occurrences). Both forms are, therefore, included as source-language items in the BCD, their order being determined by their frequency in the corpora, thus: **chercheur** *nm*, *f* **chercheuse**, *f* **chercheure** *(CD)*.

A large number of cultural Canadianisms appear in the BCD, including *blue box* and *sovereignty association* in English and *fête du Canada* and *souveraineté-association* in French, as well as the many others quoted throughout this study. These lexical items, representing cultural realities unique to Canada, are identified and selected, as was mentioned above, according to the same procedures as those used for the linguistic Canadianisms, that is, principally according to their frequency in the Project's corpora of texts.

4.2.1.2.2 Canadianisms Confined to Specific Areas of Canada

Since the BCD aims to present general Canadian usage, regionalisms that have narrow limits of distribution within Canada are not generally included (Cormier and Roberts, 2000: 216), for example, "un mot dont l'usage est exclusif au Lac-Saint-Jean ou à un village acadien n'est pas retenu" (ibid.). While a limited number of linguistic and cultural regionalisms particular to the different regions of Canada are included in the BCD, they are only those that have more general usage and are "connues dans tout le Canada" (Roberts and Bossé-Andrieu, 2002: 239). Decisions on whether the usage of a Canadian regionalism is widespread enough to warrant its inclusion as a source-language item are made based upon consultation of the BCD comparable corpora and the recommendation of specialist advisors in Canadian English and French. Based on the procedure just outlined, narrower linguistic regionalisms such as *bluff* in the sense of "a grove or clump of trees" (Prairies) (*Canadian Oxford Dictionary*), *screech* in the sense of "potent dark rum of Newfoundland" (*Canadian Oxford Dictionary*), *aboiteau* ("[surtout en Acadie] digue étanche munie de clapets", DQA) and *bombe* in the sense of "kettle" (labelled as "région." in DQA, that is, "mot ou emploi particulier au français parlé dans une ou plusieurs régions"(DQA: xxxv) are all included in the BCD.

The same procedure is used to assess whether lexical items specific to the culture of Canada's regions are to be included in the dictionary. One example of a cultural Canadianism that is included in the BCD is *député* in the sense of "membre d'une assemblée législative provinciale", for which the equivalent varies depending on the province: *(en Ontario)* member of the Provincial Parliament; *(au Québec)* member of the National Assembly; *(à Terre-Neuve)* member of the House of Assembly; *(autres provinces)* member of the Legislative Assembly.

4.2.1.2.3 North American English

As was discussed in Chapter 3, section 3.2.2 above, Canadian English usage is not comprised solely of Canadianisms. As Canadian English shares a large number of features with U.S. English, linguistic and cultural regionalisms having a wide currency in North American English (e.g. *airplane*, *automobile*, *tire* (in the sense of "pneu"), *Thanksgiving*, etc.) are included as source-language items in the BCD and identified by the label *NA*.

4.2.1.2.4 Non-Canadian Regionalisms

In recognition of the fact that Canadians of both linguistic groups are exposed to and need to understand a lot of material emanating from Europe and the United States (Roberts, 1997: 66), some source-language items are included in the BCD that are common in, but particular to the usage of France, Great Britain and the United States. These non-Canadian lexical items, which correspond either to "des usages linguistiques inusités au Canada" (Roberts and Bossé-Andrieu, 2002: 239) or to "des réalités inexistantes au Canada" (ibid.), are identified and selected for inclusion from the nomenclatures of current unilingual dictionaries and from the BCD comparable corpus (ibid.).

Examples of linguistic regionalisms common in France, Great Britain and the U.S. and that figure in the BCD include:

sèche-cheveux (FR) meaning "appareil électrique manuel . . . qui sert à sécher les cheveux mouillés" (DQA);

banger (GB) in the informal senses of "sausage", "old decrepit car" and "firework" (*Collins English Dictionary*); and

cracker, in the informal senses of (GB) "thing or person of notable qualities" and (US) "poor white" (*Collins English Dictionary*).

Some important non-Canadian cultural regionalisms may be included in the BCD when they are well represented in the non-Canadian sections of the BCD comparable corpus and substantiated in current editions of unilingual dictionaries, and especially when the "signifiant" designates a purely Canadian, as well as a non-Canadian cultural reality. This is the case with the compound *Assemblée nationale*, for example, of which the equivalents are differentiated in the BCD according to whether the source-language item denotes the cultural entity found in Canada or in France. Here, the inclusion of the explanatory glosses "parlement du Québec" and "parlement de France" help to avoid confusion and provide the user with additional useful information.

Having discussed the non-Canadian regionalisms that are included in the BCD, the exclusions should also be noted. Given the special Canadian focus of the Dictionary and the fact that Canadians are its principal targeted user group, the following linguistic and cultural regionalisms are not included in the BCD: in French, words and senses that are restricted to francophone countries other than Canada and France or to regions of France; and in English, words and senses that are restricted to Commonwealth countries other than Canada and Great Britain or to regions of the United States or Great Britain. These exclusions apply to lexical variants as well as to morphological and spelling variants (Roberts and Bossé-Andrieu, 2002: 237).

4.2.1.2.5 Conclusion

As was seen above, metalexicographers have proposed two main methods of identifying source-language items for inclusion in a bilingual dictionary, firstly, the use of a user survey to identify the types of lexical items to be included, and secondly, the use of a large corpus of electronic texts to identify the specific items to be included.

The prime objective of the BCD is to serve the special needs of sophisticated Canadian language users, more specifically, "translators, bilingual editors, bilinguals among the general

public and advanced second-language learners" (Roberts, 1994: 52). The results of the user survey conducted by the Project show that the targeted user-group placed the inclusion of "words with a specific Canadian sense" and " words designating Canadian realities" (ibid. 53) very high on the list of lexical items they wished to find in a bilingual Canadian dictionary.

From the account of the creation of the BCD corpora in Chapter 2 it is clear that the Project has fully complied with the recommendation of metalexicographers that any dictionary produced today must be not merely "based on" a corpus of authentic texts, but "faithful to corpus evidence" (Sinclair, "Corpora for lexicography", 2003: 167). Moreover, at more than 310 million words, the BCD comparable corpus fully meets the latest recommendations of acknowledged world specialists in corpus lexicography such as Graeme Kennedy who specify the need for a very large corpus because "one cannot get an accurate picture of the variability of language use unless one has a very large corpus" (Landau, 2001: 317). Because TEXTUM is a large comparable corpus, it is possible to identify Canadianisms reliably in both English and French by crosschecking concordances from the different regional sub-corpora, for example, the *English Canadian Press* against the U.S. *Wall Street Journal* for English and the *Presse canadienne française* against *Le Monde* and *Ouest-France* for French[27].

The above analysis shows that the BCD Project has formulated and implemented a coherent policy on the identification and selection of regionalisms for inclusion in the Dictionary, a policy that is entirely consistent not only with the stated aims of the Project, but also with the recommendations of metalexicographers on the subject.

4.2.2. Geographic Labelling of Regionalisms

The importance of geographic labelling in the bilingual dictionary, though recognized by metalexicographers (as discussed above), is not always acknowledged by dictionary makers in their lexicographic policies and practices. Even in those dictionaries where a list of geographic labels is provided, they are not always applied systematically, leading to confusion for the user who is often unable to ascertain with certainty the geographic area of usage of any given lexical item. The simplest and most diplomatic way of avoiding this confusion is probably to mark with

[27] See also footnote 6, page 10.

a geographic label both source- and target-language items that fall outside what has been termed the "common core of the language" (Hamilton, 1997: 126). Bilingual English and French dictionaries produced in Europe generally follow the same practice in regional labelling as that used by the *Collins English Dictionary* and other European-produced unilingual English and French dictionaries, attaching geographic labels to words and senses that vary from the norms of standard usage in Britain and France.

4.2.2.1 Geographic Labelling of Regionalisms in CR

In the front matter of the first and second editions of CR, the editors provide a brief though clear explanation of what they understand by standard usage and deviations from standard usage, stating that in their dictionary, "words and expressions that are unmarked for style or register in source or target language are to be taken as standard language appropriate to any normal context or situation. Wherever this is not the case the nature of the restriction is indicated: formal, literary, U.S., military slang, humorous, pejorative, and so on" (1978: ix and 1987: xi). CR lists 11 geographic labels, *Austral*, *Belg*, *Brit*, *Can*, *Écos/Scot*, *Helv*, *Ir*, *N Angl/N Engl* and *US*, plus the general label *dial* to denote a dialectal form. A few examples will serve to illustrate how geographic labels are used in CR, firstly on source-language items on the French-English side:

> **bobet, -ette** *adj* (*Helv = sot*) stupid, foolish.
>
> **niaiseux, -euse** (*Can*) **1** *adj* stupid, idiotic. **2** *nm,f* idiot.
>
> **nonante** *adj (Belg, Helv)* ninety;

and then on the English-French side:

> **butty** *n* (*Brit dial*) sandwich *m*
>
> **pavement** *n* (**a**) (*Brit*) trottoir *m*
>
> **sheila*** *n (Austral*) nana**f*
>
> **sidewalk** *n* (*US*) trottoir *m*.

British and U.S. usage are differentiated in CR not only for source-language items, but also for target-language items, for example:

> (*Brit*) **the tube** (= *the Underground*) le Métro; ... (*US = television*) **the tube*** la télé*
>
> **chemin de fer** railway (*Brit*), railroad (*US*)

However, a sophisticated dictionary user familiar with North American English may lament certain inaccuracies in CR's labelling of usage common to the U.S. and Canada. For example, CR's label *US*, defined as "anglais des U.S.A." (1987: xiii), is attached to the lexical items *jumper cable* and *kiss-off* both of which are, in fact, North American, being used in both the United States and Canada.

It is significant to note that because CR uses the French language of France as its standard, no geographic label is attached to examples of linguistic usage or to cultural entities that are exclusive to France. Thus, whereas the cultural regionalism *National Health Service*, for example, is labelled *Brit* in CR, the dictionary's entries for the cultural regionalism *École normale*, for example, or for *CAP* (*certificat d'aptitude professionnelle*) make no mention whatsoever of the fact that these cultural regionalisms designate entities unique to France.

4.2.2.2 Geographic Labelling of Regionalisms in OXHA

The editors of OXHA provide next to no information in the front matter of succeeding editions of their dictionary on their lexicographic policy regarding geographic labelling. OXHA lists seven geographic labels, *Austral*, *Belg*, *Can*, *GB*, *Helv*, *Ir* and *US*, plus the general label *dial* denoting a dialectal form, along with a brief one- or two-word explanation of the scope of each, but provides no details of how they are used in the dictionary. For example, the OXHA label *Can* is defined as "Canadian French/canadianisme", without any explanation of whether Canadian English usage is also included in the dictionary and if it is, whether it is labelled. What is obvious, as has been noted earlier, is the priority that OXHA, like CR, gives to the differentiation of British and American English. This is, of course, entirely natural, given the fact that both dictionaries are produced in Britain for British and American markets. Hence, the introduction to the third edition of OXHA notes that "les usages britanniques ou les américanismes sont spécifiés comme tels" ("exclusively British or North American usage is marked") (2001: ix). However, the same confusion in the differentiation of North American and U.S. usage prevails in OXHA as in CR, as can be seen from the fact that the first edition of OXHA, in a section entitled "The structure of English-French entries" (1994: xlix), actually defines its geographic label *US* as "North American usage", quoting the following abbreviation as an example:

MDT *n US* > **Mountain Daylight Time**.

Since Mountain Time is the time in the zone that includes the Canadian province of Alberta, as well as U.S. states in or near the Rocky Mountains, and Mexico, the use of the label *US* in this case is extremely misleading.

Like CR, OXHA does not attach any geographic label to examples of linguistic usage or to cultural entities specific to France because the French language of France is the standard that the editors have chosen to use in their dictionary. This means that the entries in OXHA for the cultural regionalisms *École normale* and *certificat d'aptitude professionnelle* (like those in CR discussed in the preceding section 4.2.2.1) make no mention whatsoever of the fact that these referents are unique to France, whereas the cultural regionalism *National Health Service*, for example, is labelled *GB*.

4.2.2.3 Geographic Labelling of Regionalisms in the BCD

As was discussed in Chapter 3 (3.4.1.1), the editors of two Canadian-produced unilingual French dictionaries, the *Dictionnaire du français plus* and the DQA, chose to depart from what had been the normal practice in French Canadian lexicography, by not attaching a geographic label to Canadian features, but rather to those features that are *not* part of Canadian usage. While this policy may make perfect sense for unilingual French dictionaries whose target audience is French Canadians, it does not seem practical in the context of a bilingual dictionary whose users include not only French speakers but English speakers also. Anglophone Canadians do not have an instinctive awareness of what constitutes French Canadian usage and what does not. Therefore, for practical reasons, the BCD labels the distinctive features of all the language varieties it includes: Canadian English and French *(CD)*, British English *(GB)*, North American English *(NA)*, English of the United States *(US)* and French from France *(FR)*. Presented below are some examples from the BCD:

tuque *nf* 1 *(bonnet tricoté) (CD)* tuque *(CD)*. . .

cuppa *n (informal esp. GB)* 1 *(cup of tea)* tasse *f* de thé . . .

autoroute *nf* highway *(NA)*, expressway *(NA)* . . .

congressman *n*, *pl* **congressmen**, *f* **congresswoman**, *pl* **congresswomen** *(US Polit)* 1 *(member of the national legislative body)* membre du Congrès *mf* . . .

hockey sur glace *nm (FR)* hockey *n (CD)*, ice hockey *n*.

The geographic label *CD* is used in the BCD to identify both linguistic regionalisms as well as cultural regionalisms, that is, lexical units denoting realities that exist only in Canadian culture. Thus, for example, because the Royal Canadian Mounted Police is a specifically Canadian institution, the *CD* label is attached in the BCD both to *Mountie* and to *RCMP*[28]. Furthermore, the BCD systematically follows a policy of marking with the *CD* label all examples of exclusively Canadian usage in both source and target languages, so that the user can identify them with absolute confidence and avoid the possibility of confusing them with elements of non-Canadian or general usage (Roberts, "Lexicographie bilingue du français et de l'anglais au Canada": 14). The celebrated Canadian meat pie, *tourtière*, for example, appears in the BCD as a headword on both the French-English and English-French sides, with "tourtière" given as the first equivalent in both entries and with the *CD* label attached to both source- and target-language occurrences. What remain unmarked in the BCD are those elements of French and English that are common to both the Canadian variety and to the other major varieties of each language.

4.2.3. Equivalents

This section will begin by describing briefly the information presented in the front matter of the two European-produced dictionaries regarding their treatment of general equivalence. This information will serve as a comparison with the next section, which examines the inclusion of regionalisms as equivalents in the two European-produced dictionaries. The section will conclude with a description of the information presented in the policy documents of the BCD regarding its handling of equivalents, both general and regional.

4.2.3.1 Equivalents in CR and OXHA

In none of the prefaces or introductions to a series of six editions do the editors of CR choose to give any detailed information on their overall lexicographic policy regarding the handling of equivalence in their dictionary, other than to mention in a very general way that they

[28] Both nouns are used informally to denote an officer of the Royal Canadian Mounted Police. *Mountie*, however, is used in all varieties of English, whereas *RCMP* seems to be used in this sense only in Canada. The Canadian Oxford Dictionary defines this sense as "an RCMP officer" and labels it "Canadian informal". Its usage is illustrated by the following example in the English Canadian Press section of the BCD corpus: "an RCMP who sets out to track down his father's killer". This sense is not attested at all in the U.S. sections of the BCD corpus.

have "taken care that each side of the Dictionary is equally helpful for translation from and into the foreign language" (1987: xi).

The editors of OXHA, on the other hand, state in the prefaces and introductions to succeeding editions of their dictionary that their policy on equivalents is to identify what they call "one good translation" (2001:ix) of a lexical item, to check its use in their electronic corpora and then, on many occasions, to give this one equivalent as "the only general translation for one sense of the headword" (ibid.). It is explained in addition that two or more translations of a headword are only given "in those rare cases where they are consistently interchangeable" (ibid.) and, furthermore, that when one equivalent is considered "adequate" (ibid.), the addition of "less safe alternatives" (ibid.) is avoided. Although it would have been helpful to have been offered some information on what the editors of OXHA mean exactly in this context by the adjectives "good" and "adequate", which are open to a variety of interpretations, it is nonetheless reassuring to see a statement of lexicographic policy laid out clearly in this way in the front matter of a bilingual dictionary, as recommended by Al-Kasimi (1977: 109) as long ago as 1977.

4.2.3.1.1 Inclusion of Regionalisms as Equivalents in CR

Given the lack of information provided by CR on their handling of equivalence in general, it is not surprising that its front matter is relatively silent as far as the inclusion of regionalisms as equivalents is concerned. However, succeeding editions of the guide to the use of the dictionary do include two examples of regionalisms being included as target-language equivalents. Both are examples of lexical differences between British and U.S. English:

appartement *nm* (a) flat (*Brit*), apartment (*US*) (1978: x)

câble de démarrage (*Aut*) jump lead (*Brit*), jumper cable (*US*) (1993: xx)

and the two geographic labels *Brit* and *US* appear to be used systematically in the body of the dictionary to highlight differences between British and U.S. usage, either lexical differences in entries such as:

caddie *nm* trolley (*Brit*), cart (*US*) ...

and

trottoir *nm* (a) (*accotement*) pavement (*Brit*), sidewalk (*US*) ...

or orthographic differences in such entries as:

chasse-neige *nm* snowplough (*Brit*), snowplow (*US*) ...

Unfortunately, when CR steps off the well-trodden path of British-U.S. variation, its coverage is not so secure. The French Canadianism *cabane à sucre*, for example, is included as a compound in the sixth edition (2002) of CR, quite correctly labelled *Can*. However, the equivalent provided, "sap house", although confidently labelled *Can* also, is not in fact a Canadianism at all, nor is it even used in Canada, as is attested by the fact that it appears neither in the *English Canadian Press* section of the BCD corpus nor in the *Canadian Oxford Dictionary*. Indeed, it is not even included on the English-French side of CR. The correct Canadian equivalent, "sugar shack", which is included in the *Canadian Oxford Dictionary* and is well represented in the *English Canadian Press* section of the BCD corpus, does not appear in CR.

As far as equivalents for cultural regionalisms are concerned, CR's "Using the Dictionary" section explains that the symbol ≈ is used "when the source language headword or phrase has no equivalent in the target language and is therefore untranslatable. In such cases, the nearest cultural equivalent is given." (2002: xxii) Presented below are some examples from both sides of the dictionary of CR's use of this symbol of cultural equivalence:

baccalauréat *nm* ≈ (*en France*) . . . A-levels (*Brit*), ≈ high school diploma (*US*) . . . (*au Canada* = licence) ≈ BA

bêtise de Cambrai *nf* ≈ mint humbug (*Brit*), piece of hard mint candy (*US*)

National Health Service *n* (*Brit*) ≈ Sécurité *f* sociale.

As can be seen from these examples, CR implements its stated policy of illustrating differences between British and U.S. usage, but because the language and culture of France constitute the standard that the editors of CR have chosen to use in their dictionary, they do not include any target-language material referring to features of other French-speaking cultures. Thus in CR's entry (2002) for the compound *teacher training college*, both target-language equivalents given reflect only French culture:

teacher training college *n* ≈ Institut *m* universitaire de formation des maîtres, IUFM *m*.

In cases where there is no cultural equivalent in the target language, CR provides (2002)

an explanatory gloss in italics, for example:

Yorkshire pudding *n (Brit Culin) pâte a crêpe cuite qui accompagne un rôti de bœuf*

4.2.3.1.2 Inclusion of Regionalisms as Equivalents in OXHA

Like their fellow lexicographers at Collins-Robert, the editors of OXHA do not enunciate a general lexicographic policy regarding the inclusion of regionalisms as target-language equivalents. But like CR, OXHA does include regionalisms as target-language equivalents in the case of British-U.S. orthographic and lexical variation. The editors state in the introductions to the first and third editions of OXHA that, where appropriate, "dans la partie français-anglais, nous proposons des traductions en anglo-americain chaque fois qu'une variante graphique ou lexicale s'impose" (2001: ix).

Like CR, OXHA also includes a guide to the use of the dictionary entitled "The structure of French-English entries" (1994: xxxii) in which two examples are given of entries where regional variants are offered as target-language equivalents. The first, on the English-French side of the dictionary, is described as a "North American translation":

bagage à main hand luggage, carry-on baggage *US*.

This type of entry probably confuses more than helps the dictionary user who remains uncertain not only about whether "carry-on baggage" is a compound used throughout North America or only in the United States, but also about whether the other equivalent offered, "hand luggage", is also a regionalism and if so, of which region. The lack of a geographic label attached to the equivalent "hand luggage" is open to a variety of interpretations. The second example, at the headword *formidablement*, is described as a "traduction avec sa variante nord-américaine":

il a grossi formidablement he's got *GB* or gotten *US* awfully fat.

4.2.3.2 Equivalents in the BCD

Given its special Canadian orientation, the BCD generally includes as target-language equivalents only those lexical items that are used in North America. This means that lexical items that are particular, for example, to British English or to continental French and not used in Canada or North America as a whole are not offered as equivalents. Thus, for example, at the headword *camion*, only the equivalent "truck" is given and the British linguistic regionalism

"lorry" is not included. Similarly, on the English-French side of the Dictionary, the compound *high school* is translated as "école secondaire", and "lycée", the cultural regionalism used predominantly in France[29], is not given as an equivalent. A comparison of the numbers of occurrences of the term *lycée* in the *Presse canadienne française* sub-corpus and in the sub-corpus of texts from *Le Monde* shows that the term is far and away more frequent in the texts produced in France, occurring around 1,000 times in a corpus of 17 million words, as against only 200 occurrences in the French Canadian sub-corpus of more than 77 million words. Analysis of 100 concordances generated (of the total 200) for *lycée(s)* in the *Presse canadienne française* sub-corpus shows that, with fewer than 10 exceptions, the word occurs in Canadian texts only when the educational establishment referred to is located in a country other than Canada, usually in France (the names of the famous *lycées* of Paris recur: Henri IV, Louis-Le-Grand, Saint-Louis and Voltaire), but also, in countries with educational systems as widely varied as the United States ("un lycée de Fairfax, près de Los Angeles"), Israel ("au lycée Rashidya de Jérusalem"), Italy ("dans un lycée du centre de la Sicile") and Japan ("au lycée de Tazawako [préfecture d'Akita]"). In other words, the term *lycée* is not used in Canadian texts to refer to Canadian educational establishments, apart from a few well-known exceptions such as the Lycée Claudel in Ottawa.

Although the special Canadian focus of the BCD guarantees the inclusion of many Canadian regionalisms as source-language items, this does not necessarily mean that they all appear on the other side of the Dictionary as target-language equivalents. This is primarily because the BCD, like all dictionaries, has to make a selection due to limitations, in the case of the printed dictionary, of space, from all the material that could potentially be included in every entry. Canadianisms are, therefore, included as target-language equivalents in the BCD according to the following three main criteria:

(i) if the Canadianism appears frequently in the BCD Canadian corpora: for example, given the fact that the Canadianism *épinette* is 50 times more frequent in the BCD French Canadian sub-

[29] The fact that, in the *Dictionnaire québécois d'aujourd'hui*, the label "France" is attached to the entry *lycée*, means that this term is known in Canada but that its use is "plutôt passif" (*Dictionnaire québécois d'aujourd'hui*, 1993: xxi), that is to say that the word is known and understood in Canada (through television, radio, cinema and the print media), but that a Canadian speaker would use it rarely in his own language production, "sauf lorsqu'il veut créer un effet rhétorique" (ibid.).

corpus than the standard French word *épicéa*, it is naturally included as one of the equivalents for the English noun "spruce".

(ii) if the Canadianism is linguistically productive: for example, the verb *magasiner* has led not only to the creation of the derivatives *magasineur* and *magasineux*, but also to the addition of the meaning "shopping" to the noun *magasinage*. Hence, *magasiner* is considered worthy of inclusion as an equivalent for the verb "to shop".

(iii) if no other lexical item exists in the target language that corresponds exactly to the concept expressed by the source-language item. The Canadianism, *citron*, for example, a calque of the English word "lemon" with the meaning of "objet acheté qui a un ou des vices de fabrication, spécialement une automobile" (DQA) is included in the BCD as an equivalent for the English headword "lemon" in this sense; and in the French-English side of the Dictionary, the Canadianisms to *francize* and *francization* are given as the only equivalents for the headwords "franciser" and "francisation".

The anglicism is a special type of Canadianism that occurs with relatively high frequency in Canadian French usage as a result, as was discussed in Chapter 3, of the particular circumstances of the historical evolution of the French language in Canada. The BCD Project has established special guidelines for the inclusion and exclusion of anglicisms, not only as headwords (see section 4.2.1.2.1) but also as equivalents. While anglicisms are normally included as source-language items in the French-English section of the BCD according, for the most part, to their frequency in the BCD corpora, anglicisms will be included as equivalents for English headwords only if there is a special justification for doing so. This BCD policy is based on the premise that, while the Dictionary has a duty to record the use of anglicisms in Canadian French, it is not its role to propagate them by systematically proposing them as equivalents. All anglicisms included in the BCD as source-language items do not, therefore, necessarily appear on the other side of the Dictionary as target-language equivalents. When assessing the suitability of an anglicism for inclusion as an equivalent on the English-French side of the Dictionary, BCD lexicographers take account of the most recent recommendations of l'Office québécois de la langue française and examine each case individually according to the same three criteria described above in section 4.2.3.2 for the evaluation of Canadianisms in general. The anglicism *surtemps* (a calque of English "overtime"), for example, given its frequent occurrence in the

BCD's corpora of French Canadian texts, is included in the Dictionary both as a headword and as an equivalent for the headword "overtime". On the other hand, an anglicism is not included in the BCD as an equivalent if it is "concurrent d'un mot français . . . dont la fréquence d'utilisation est égale" (Roberts, "Lexicographie bilingue du français et de l'anglais au Canada": 14). For example, the anglicism "audience" is not given as an equivalent for the English headword *audience* (in its first sense of "physical body of spectators or listeners"), because the French equivalents "public", "auditoire", "assistance" and "spectateurs" are all represented in the BCD's French Canadian sub-corpus with greater frequency.

As its policy statements make clear, the BCD is a work of scholarly endeavour and as such, has been compiled in accordance with many of the recommendations of metalexicographers and linguists such as Zgusta and Al-Kasimi, whose theories were highlighted in section 4.1 above. As far as equivalence is concerned, BCD lexicographers always aim whenever possible to supply a full translational (insertible) equivalent that matches the source-language item as closely as possible in meaning, register and geographic distribution. To achieve this for lexical regionalisms often presents difficulties, as was discussed above, but for cultural regionalisms these difficulties are often even greater. In these cases where no translational equivalent exists for a cultural regionalism, the BCD provides an explanatory equivalent. Because the degrees of "translatability" for culture-specific entities are just as varied as they are for purely linguistic items, ranging from near-perfect equivalence at one extreme to a straightforward definition in the source language at the other, the BCD provides equivalents of different kinds.

At one end of the spectrum lies the case of the culture-specific item for which a regionally distinct translational equivalent exists that matches the source-language item as perfectly as is possible in sense, register and geographic distribution. The French Canadianism *tourtière*, for example, is perfectly translated into English by the borrowing "tourtière", so both appear as source- and target-language items in both sides of the BCD. The French/English Canadianisms *dépanneur/depanneur* and *cégep/CEGEP* provide further examples of this type of perfect regional equivalence, which in many, if not most cases, is achieved by the linguistic process of borrowing.

Further along the translatability scale, it is sometimes possible to coin a translational equivalent that can be readily inserted in all contexts, for example:

> **outport** *n* **1** (*isolated fishing village, esp. on the Newfoundland coast*) *(CD)* petit port de pêche isolé *m*, petit village de pêcheurs *m*.

In other cases, even when there is an equivalent term in the target language, the BCD may include a brief explanation or gloss (located after the equivalent) to make quite clear to target-language users an element of meaning that may not be evident from the equivalent alone. Here are two examples of this type, the first from the French-English side of the BCD and the second from the English-French side:

> **polyvalente** *nf (CD)* (*école secondaire*) high school, secondary school *(offering both technical and academic training)*.
>
> **residential school** *n (CD) (boarding school maintained by the federal government for Inuit and First Nations children)* pensionnat *m (du gouvernement fédéral pour les enfants autochtones et inuit)*.

However, when no "designatum" exists or can be coined in the target language to denote these exclusively regional concepts, the BCD provides an explanatory translation. For instance, there is no real equivalent in English for the noun *bleuet* in the sense of "habitant du Saguenay-Lac-Saint-Jean", so the BCD provides an explanatory equivalent, "inhabitant of the Saguenay-Lac-Saint-Jean region". This kind of explanatory equivalent is always given in italics in the BCD as an attention-grabbing device to alert the user by distinguishing it from the normal typeface of a translational equivalent. BCD policy also ensures that an explanatory equivalent is always accompanied by an illustrative example to show the user how the material offered can be adapted for use in his own particular context. In the case of *bleuet*, the explanatory equivalent is followed by this illustrative example:

> **il s'appelle Tremblay – ça doit être un bleuet** his name's Tremblay – he must be from the Saguenay-Lac-Saint-Jean region.

4.3 Conclusion

The aim of this chapter has been to complete the answer to the question of *why* a bilingual Canadian dictionary is needed by examining three different aspects of the handling of regionalism in general and of Canadian usage in particular, firstly, according to the front matter of the two top-selling English-French dictionaries (CR and OXHA) and secondly, according to the policy documents of the BCD.

Regarding the identification and selection of regionalisms as source- and target-language items, CR and OXHA both state in their front matter that it is their policy to include examples of regional usage in English and French, but it is apparent from the sample entries given in the front matter that the emphasis in this selection is primarily on lexical and orthographical variation between British and North American English. From this it may be deduced that certain criteria regarding the identification and selection of regionalisms are being met, but that the editors of CR and OXHA decided not to include these details of their lexicographic policy in the front matter of their dictionaries. Even when the contribution of regional material is specifically acknowledged in the front matter, as with the French Canadian material provided in early editions of CR, no details are given of where these lexical items were found or how they were selected for inclusion, although this information is no doubt available in internal documentation.

As the product of a scholarly project that is still in progress, the BCD does not, as yet, have a published preface or "guide to the use of the dictionary", but there is ample detail in the Project's policy documents of the methodology used to identify and select both source- and target-language regionalisms, as described in section 4.2.1.2 above.

It is clear from the front matter of both CR and OXHA that, as Franco-British collaborative projects, both dictionaries use British English and the French of France as their standard. However, as far as geographic labelling is concerned, it is never explicitly stated in either dictionary that lexical items that do not have any geographic label attached are deemed to form part of "standard language" (CR 1978: ix and 1987: xi) in both English and French and as such are, therefore, considered common to a number of different regions. Moreover, neither CR

nor OXHA offers an explanation of why examples of exclusively British usage and exclusively North American usage are labelled when examples of usage specific to France are not.

It may be concluded, therefore, that although CR and OXHA are both excellent dictionaries in their own right, because they are designed principally for English-French dictionary users in Europe and the United States, they do not meet the needs of English- and French-speaking Canadians. By including a large number of Canadianisms in both source- and target languages and by clearly labelling examples of Canadian usage as well as other regional usage, the BCD, on the other hand, will provide the tool that Canadians need in order to understand and communicate effectively in both official languages, in their own country as well as abroad.

5. THE TREATMENT OF REGIONALISM IN DICTIONARIES ACCORDING TO BILINGUAL DICTIONARY ENTRIES

Chapter 4 first established the criteria against which the treatment of regional variation in bilingual dictionaries is assessed in this study and then began the process of evaluating the handling of regionalism in CR, OXHA and the BCD against these criteria. This first stage of the evaluation was carried out using the evidence of the front matter of the two European-produced English-French dictionaries and the policy documents of the BCD. In the second stage of this evaluation, Chapter 5 will now examine the treatment of regionalism in the three dictionaries according to the evidence presented by a selection of specific lexical items.

5.1 The Regionalisms to be Studied

The ten lexical items chosen for study represent the major regional varieties of English and French, namely those used in North America (NA), the United States (US), Great Britain (GB), Canada (CD) and France (FR). Because this study is concerned with an examination of the treatment of regionalism in bilingual English and French dictionaries, the selection is made up of five English and five French items. Given that the BCD, the primary focus of this study, is a Canadian dictionary, five of the ten lexical items chosen are Canadianisms, of which two are English and three French, since Canadianisms in French outnumber those in English (Grenon-Nyenhuis, 1997: 57).[30]

Three types of regionalism were presented earlier in this study: lexical regionalism, grammatical regionalism and orthographic regionalism. Of the three types examined in Chapter 3 in an exclusively Canadian context, the most important numerically and the most widely known are the lexical regionalisms (Cormier, 2000: 217). This study will restrict itself, therefore, to an analysis of only lexical regionalisms. Although regional varieties of French and English share the vast bulk of their vocabulary with French and English spoken in other parts of the world, lexical

[30] It should be noted that this study does not examine the more restricted levels of regionalism, such as those used in a more limited geographic area (e.g. Newfoundland, Texas, Brittany, Guadeloupe, etc.) or in a particular city (e.g. London, New York, Montreal, Paris, etc.), or by a particular group of people (e.g. Creoles, Northern Irish, etc.), as these finer variations are hardly ever included in bilingual dictionaries.

regionalisms form the majority of the features that distinguish regional usage. The presence or absence of a number of lexical regionalisms in the three dictionaries will, therefore, be a good primary indicator of how well regional usage is covered. Of the ten items selected for study, seven are linguistic regionalisms and three are cultural regionalisms, as defined in Chapter 3.

Although a limited sample of lexical items is presented in detail here, the selection is based on a preliminary study of a much larger sample eventually narrowed down to a shortlist of 15 English and 8 French regional lexical items[31].

5.1.1. Choosing the Lexical Items

Having established the geographic areas from which the regionalisms would be selected and the type of regionalism to be studied, the status of candidate regionalisms as valid examples of regional usage worthy of further study was confirmed on the basis of their frequency in corpora and their presence in unilingual dictionaries.

5.1.1.1 Frequency in Corpora

The BCD corpus, containing authentic examples of texts representing Canadian, U.S. and French usage and containing over 300 million words (see Chapter 2, section 2.2.4), was used to validate the status of NA, US, CD and FR regionalisms. The British National Corpus, a corpus of texts containing 100 million words of written and spoken English from British sources (<http://www.natcorp.ox.ac.uk/>), was used to validate the GB candidate regionalisms. All ten lexical items are well represented in the appropriate corpora, as the following tables show.

[31] The following lexical items were studied and all are included in the BCD: NA - *auditor* (university student), *faucet*, *hood* (of a car), *coulee* (ravine), *trunk* (of a car); US – *cotillion* ("formal ball"), *cotton wool* ("raw cotton"); GB: *banger* (sausage, old car), *bonnet* (of a car), *boot* (of a car); CD English - *cheese cutter* (a child's ice skate), *hydro* (electricity), *hydro line*, *snye* (channel of water), *tuque*; CD French – *beigne* (pâtisserie), *breuvage* ("boisson non alcoolisée"), *cabane à sucre* ("bâtiment . . . où l'on prépare les produits de l'érable"), *huard*, *prendre le plancher* (to take the floor), *tabagie* ("établissement commercial où l'on vend surtout du tabac"); FR - *beigne* ("coup, gifle"), *tabac* ("établissement commercial où l'on vend surtout du tabac").

Table 4

Presence in Corpora of the Sample English Lexical Items

			Number of occurrences		
	Region	**Definition (OXCAN)**	**English Canadian Press (ECP)**	**Wall Street Journal (WSJ)**	**British National Corpus (BNC)**
attorney	US	lawyer	280	17	229
muffler	NA	device attached to a motor vehicle's exhaust system to reduce noise	150	14	0
riding	CD	district whose voters elect a representative member to a legislative body	8,284	0	1
riding association	CD	unit of organization of a political party at the level of the riding	514	0	0
silencer	GB	device attached to a motor vehicle's exhaust system to reduce noise	0	0	33

Notes:

attorney: In order to focus the analysis on the word *attorney* only, the parameters for the corpus searches were set to exclude occurrences of *city attorney*, *county attorney*, *Crown attorney*, *district attorney*, *state attorney*, *U.S. attorney*, *attorney general* and *power of attorney*. The number of concordances generated in this way in the ECP for *attorney* alone was 280 of which a sample of 155 was randomly selected for further examination. From an examination of these 155 concordances (i.e. not the full text), it was clear that most of the press reports were concerned with legal activity taking place in the United States. This was easily deduced by the fact, for example, that the word *attorney* was preceded by or used with the name of a U.S. state (e.g. "a Florida attorney", "attorneys for the Tennessee Department of Health", etc.) or with the name of a person widely known to have been involved in legal proceedings that took place in the United

States (e.g. "Farrow's attorney", referring to the 1993 Woody Allen-Mia Farrow child custody hearing and "attorneys for Whitney Houston", etc.) or occurred in a phrase clearly indicating a U.S. context (e.g. ""Our attorneys tell us we're an American newspaper and we print what we choose," said Stan Evans, assistant managing editor of the Buffalo News", "if you're injured while wintering in Florida . . . ", "the FBI said cult attorneys would not . . ." , etc.). Each of the press articles in the corpus is tagged, in the case of Canadian publications, with the name of the publication in which it first appeared and, in the case of non-Canadian publications, with the name of the town, city, state or country in which the report first appeared (many of the articles being syndicated, of course, in a large number of newspapers and magazines worldwide). Analysis of the full text of 10% of the 155 concordances generated for *attorney* revealed that they all appeared originally either in a U.S. publication or concerned an event that took place in the U.S.

When searches in the BNC for *attorney* were similarly restricted, the number of concordances generated was 229, of which a sample of 100 was randomly selected for further examination. From an examination of these 100 concordances (i.e. not the full text), it was clear that in only around six cases was the word used in the U.S. sense of "lawyer ... representing a client in a court of law", as opposed to being used in the sense that is the most common in Canada and Britain of "person . . . appointed to act for another in business or legal matters". This was easily deduced from the same indicators as those mentioned above, that is, the inclusion, for example, of the name of a U.S. state or city (e.g. "the Washington attorney was a close friend".)

muffler: The vast majority of the 150 occurrences in ECP are in the automobile sense given above. There are only one or two examples among the 150 occurrences of the "scarf" sense. There are no occurrences of *muffler* in the sense of "device attached to a motor vehicle's exhaust system to reduce noise" in the BNC.

riding: It is difficult to analyze in detail a large number of concordances such as the 8,000-plus occurrences of *riding* in the ECP, but it was possible to restrict the search to some extent by using the query strings "the riding", "a riding", "his riding", "her riding" , "ridings", etc., in order to avoid some of the –ing forms of the verb *to ride*. Of the 160 occurrences of *riding*

produced in this way from restricted queries in the BNC, only one was found to be in the Canadian sense of "electoral district": "Jean Chrétien, a French Canadian who for years represented Quebec ridings as a minister in Pierre Trudeau's era"

silencer: In the automobile sense, this is a GB regionalism as is attested by the fact that 71% (33 out of 46) of the total number of occurrences in the BNC are in this sense, whereas it is completely absent in this sense from the Canadian and U.S. subcorpora.

Table 5

Presence in BCD Corpora of the Sample French Lexical Items

	Region	Definition	Subcorpora from Canada	Subcorpora from France
accommoder	CD	rendre service à (DQA)	640	0
basket	FR	chaussure de sport (DQA)	21	34
fête du Canada	CD	Fête légale observée le 1er juillet (Meney)[32]	104	0
premier étage (See Notes)	CD	rez-de-chaussée (DQA)	159	115
premier étage (See Notes)	FR	the floor above the ground floor (OXCAN)	159	115

Notes:

accommoder: It is difficult to analyze in detail all of the large number of concordances such as the 640 occurrences of *accommoder* in the Canadian subcorpora. However, analysis of the total number of 157 occurrences of *accommoder* in the *Le Monde* sub-corpus reveals that there is not one single occurrence of the verb in the sense of "rendre service à". It seems fair to deduce from this analysis that *accommoder* is not commonly used in France in this sense.

basket: The apparently small number of occurrences of this lexical item (21 in the subcorpora from Canada and 34 in the subcorpora from France) is a result of setting the parameters for the corpus search to exclude occurrences of *basketball* and *basket-ball*. The difference in the number of occurrences of this regionalism in the two regionally distinct corpora becomes even more

[32] This definition of *fête du Canada* is from Meney's *Dictionnaire québécois français*, 1999.

marked when the number of occurrences is expressed as a percentage: 50% of the occurrences of this word in the subcorpora from France are in the sense of "chaussure de sport" (the others being in the sense of "basketball"), whereas only 10% of the occurrences in the Canadian subcorpora are in the sense of "chaussure de sport". In other words, *basket*, in the sense of "chaussure de sport", occurs five times more frequently in the subcorpora from France than in the Canadian subcorpora. Further analysis of the full text from which the 21 concordances in the Canadian subcorpora are generated reveals that although these articles appeared in Canadian publications, the majority of them deal with activities taking place in France: a good number of the articles, for example, are reviews of books and films produced in France and originally appearing, most probably, in newspapers and magazines published in France.

premier étage: In the vast majority of corpus examples, it is impossible to know whether this lexical item is being used in the sense of "ground floor" or "floor above the ground floor".

5.1.1.2 Presence in Unilingual Dictionaries

To confirm their status as commonly occurring lexical items worthy of study, the regionalisms were searched in the most recent college-size (i.e. not concise) editions of appropriate general unilingual dictionaries. The English items were searched in a Canadian dictionary (*Canadian Oxford Dictionary*, OXCAN), a U.S. dictionary (*Random House Webster's Unabridged Dictionary*, RHWUN) and a British dictionary (*Collins English Dictionary*, CED). The French items were searched in a Canadian dictionary (*Dictionnaire québécois d'aujourd'hui*, DQA, 1993) and in a dictionary produced in France (*Le Nouveau Petit Robert*, NPR, 2003). The results of these searches are presented in the following tables.

Table 6

Presence in Unilingual Dictionaries of the Sample English Lexical Items

	Geographic region	OXCAN	RHWUN	CED
attorney	US	YES	YES	YES
muffler	US	YES	YES	YES
riding	CD	YES	NO	YES
riding association	CD	YES	NO	NO
silencer	GB	YES	YES	YES

Table 7

Presence in Unilingual Dictionaries of the Sample French Lexical Items

	Geographic region	DQA	NPR
accommoder	CD	YES	NO
basket	FR	YES	YES
fête du Canada	CD	NO	NO
premier étage	CD	YES	YES
premier étage	FR	NO	YES

As can be seen in the preceding tables the sample lexical items do not figure in all the dictionaries. Whereas all five of the English sample lexical items are included in OXCAN, the two Canadianisms are not included in the RHWUN and one of the Canadianisms does not appear in the CED. Similarly, not all of the sample French regionalisms appear in both dictionaries used. One Canadianism and one example of a regionalism from France are omitted from the DQA and two of the French Canadianisms are omitted from NPR.

5.1.1.3 Final Choice

After the preliminary stages presented above, the final choice of lexical items was made, taking into account a number of different parameters: different regional varieties; linguistic vs. cultural regionalisms; simple vs. complex items, etc.

The pair *muffler* and *silencer* were selected as an example of how different regional varieties of a language (in this case, American and British English) may choose completely

different lexemes to express the same concept (i.e. "a device attached to a motor vehicle's exhaust system to reduce noise"). English Canadian usage is represented by the two Canadianisms *riding* and its derived compound *riding association*. French Canadian usage is represented by the linguistic Canadianism *accommoder* and the cultural Canadianism *fête du Canada*.

The collocation *premier étage* represents the opposite phenomenon to the *muffler/silencer* pair presented above, as it illustrates the case where one lexical item has become polysemous. The lexical item *premier étage* has taken on an additional meaning in its transition from France to Canada, a meaning probably acquired due to the influence of the North American English "first floor" in the sense of "the floor on the ground level". It is interesting to note, however, that the close co-habitation of French and English in Canada does not always produce the same linguistic effect. The anglicism *basket*, meaning "chaussure de sport", borrowed into French from English in the 1950s, did not cross the Atlantic to French-speaking Canada, where *espadrille* is by far the most widely used term for this kind of sports shoe.

5.2 Significant Features of the Treatment of Regionalism as Exemplified in Dictionary Entries

5.2.1. Identification and Selection of Regionalism as Exemplified in Dictionary Entries

5.2.1.1 In CR and OXHA

The following tables present the results of searches for the ten lexical items in CR and OXHA and, for comparison, their presence in the BCD.

Table 8

Presence in Bilingual Dictionaries of the Sample English Lexical Items

	Geographic region	CR	OXHA	BCD
attorney	US	YES	YES	YES
muffler	US	YES	YES	YES
riding	CD	NO	NO	YES
riding association	CD	NO	NO	YES
silencer	GB	YES	YES	YES

Table 9

Presence in Bilingual Dictionaries of the Sample French Lexical Items

	Geographic region	CR	OXHA	BCD
accommoder	CD	NO	NO	YES
basket	FR	YES	YES	YES
fête du Canada	CD	NO	NO	YES
premier étage[33]	CD	YES	NO	YES
premier étage	FR	YES	YES	YES

The data in the above Tables 8 and 9 shows that both CR and OXHA, in accordance with statements made in their front matter (see Chapter 4), do include US, GB and FR regionalisms. On the other hand, assertions in the front matter of the two dictionaries that they also include regionalisms from other parts of the English- and French-speaking worlds are not fully corroborated by this sample. Although CR does include the French Canadianism *premier étage*, neither CR nor OXHA includes the very common English Canadianisms, *riding* and *riding association*.

[33] OXHA gives the noun collocation *le premier étage*, whereas CR gives the adverbial collocation *au premier étage*.

5.2.1.2 In the BCD

In accordance with its policies (as discussed above), the BCD selects lexical items for inclusion in its nomenclature based on the frequency of their occurrence in the Project's corpora and on their presence (or absence) in appropriate unilingual dictionaries. All ten lexical items presented here fully meet these criteria and, therefore, are all included in the BCD.

5.2.2. Geographic Labelling as Exemplified in Dictionary Entries

5.2.2.1 In Unilingual Dictionaries

As was discussed above, the sample lexical items were searched using the most recent editions of six highly respected unilingual dictionaries of English and French. The two tables below present the details of the geographic labelling of the sample items in these dictionaries. The first table, Table 10 below, shows which of the sample English lexical items have geographic labels attached in each of the three dictionaries. OXCAN states in its "Guide to the Use of This Dictionary" (2004: xvii) that "if the use of a word is restricted in any way, this is indicated by any of various labels" and thus all five regionalisms have the appropriate geographic label attached. As is entirely normal for a dictionary produced in the U.S., the RHWUN has elected to use U.S. English as its standard and does not, therefore, attach any geographic label to *attorney* and *muffler*, since both "are considered to be in general use throughout the U.S." (RHWUN 1992: xix). The lexical items that remain unmarked for geographic region in the RHWUN, as in the CED, are those that form part of the regional variety of English which these dictionaries have elected to use as their standard. Hence, the RHWUN, an American dictionary, attaches an appropriate geographic label to the British regionalism *silencer*, whereas the CED, being a British-produced dictionary whose standard language is British English, does not. Although the CED does not attach a geographic label to *silencer*, it does provide additional information for the user in the form of a cross-reference to *muffler*, the North American synonym in this sense ("device attached to a motor vehicle's exhaust system to reduce noise"). By attaching appropriate geographic labels, the CED accurately and clearly differentiates the exclusively U.S. item (*attorney*) and the exclusively Canadian item (*riding*) from the example of North American usage (*muffler*).

Table 10

Geographic Labelling in Unilingual Dictionaries of the Sample English Lexical Items

	Geographic region	OXCAN label attached	RHWUN label attached	CED label attached
attorney	US	US	NO	U.S.
muffler	NA	N Amer	NO	U.S. and Canadian
riding	CD	in Canada	lexical item not included	in Canada
riding association	CD	Cdn	lexical item not included	lexical item not included
silencer	GB	Brit.	Chiefly Brit.	no label but a cross-reference to *muffler*

As far as the French lexical items are concerned, Table 11 below shows that, in accordance with its stated policy (see chapter 3, section 3.4.1.1), the DQA does not attach geographic labels to those elements of the French language that it considers to be used exclusively or predominantly in Canada, but that it is "les particularismes lexicaux européens répertoriés, essentiellement ceux de la France, qui sont accompagnés d'une marque" (DQA 1993: xi). Thus, in the DQA, the Canadian senses of *accommoder* and *premier étage* are not labelled for geographic region, whereas *basket*, being an example of regional usage from France, does receive an appropriate label. The sample lexical item *fête du Canada* does not appear in the DQA[34]. Being a dictionary that uses the French of France as its standard, NPR does not attach any geographic label to the lexical items *basket* and *premier étage* (in the sense of "floor above the ground floor"), since they are both regionalisms used chiefly in France. NPR does, however, attach the label "au Canada" to the Canadian sense of *premier étage* meaning "ground floor".

[34] In its entry for *fête du Canada*, the *Dictionnaire québécois français*, a Canadian special-purpose dictionary, which, like the DQA, has as its standard language the French of Quebec and more broadly of Canada as a whole, does not attach a geographic label to the cultural Canadianism *fête du Canada*.

Table 11

Geographic Labelling in Unilingual Dictionaries of the Sample French Lexical Items

	Geographic region	DQA	NPR
accommoder	CD	no label	this sense not included
basket	FR	France	no label
fête du Canada	CD	lexical item not included	lexical item not included
premier étage	CD	no label	au Canada
premier étage	FR	lexical item not included	no label

5.2.2.2 In CR and OXHA

The three sample English entries *attorney*, *muffler* and *silencer* (see Appendix) that are included in CR and OXHA were analyzed to see how closely their use of geographic labels accords with statements made in the front matter of these two dictionaries. Analysis of CR's entries for these three items shows that, in accordance with the statements made in its front matter, geographic labels are attached to those lexical items whose usage is restricted to a specific geographic area. Analysis of OXHA's entries shows that, although (as was noted in Chapter 4) OXHA makes no comment in its front matter on its use of geographic labelling, the appropriate geographic labels are attached to the three sample lexical items. All three of the sample English items included in CR and OXHA are identified by the appropriate geographic label: *attorney* and *muffler* are labelled *US* in CR and OXHA, and *silencer* is labelled *Brit* in CR and *GB* in OXHA.

As was noted in Chapter 4, because both CR and OXHA use the French language of France as their standard, they do not as a general rule attach any geographic label to examples of linguistic usage or to cultural entities that are exclusive to France. This is borne out in the entry for *basket*, which, despite the fact that it is a regionalism from France, has no geographic label attached in either CR or OXHA. However, in the case of the polysemous linguistic regionalism *premier étage*, CR opts for a more detailed presentation, identifying the finer regional differences of the source-language item with the contrasting glosses *(en France)* and *(au Canada)*. OXHA does not label the source-language item *premier étage* as having any specifically regional meaning. In this aspect, therefore, CR's presentation of the sample lexical item is far more helpful than that of OXHA, for it provides the dictionary user with additional

information on regional usage, thereby facilitating his task, be it one of comprehension or of translation.

As was discussed in Chapter 4, it is entirely natural, given the fact that both are dictionaries produced in Britain for the British and North American markets, that CR and OXHA should give priority to the differentiation of British and North American English. Both dictionaries state in their front matter that "exclusively British or North American usage is marked" (OXHA, 2001: ix). The entry for *premier étage* in CR and OXHA serves not only as an example of how regional and standard usage may be differentiated, but also of how different geographic varieties may be presented. The difference between the use of the source-language item *premier étage* in France and in Canada is highlighted by CR but not by OXHA. The difference between the use of the target-language items "first floor" and "second floor" in Britain and North America is highlighted by both CR and OXHA. However, as was noted in Chapter 4, neither CR nor OXHA accurately differentiates North American usage from United States usage. Both CR and OXHA include the label *US*, CR defining it as "américain, États-Unis/American, United States" (CR, 2002) and OXHA as "anglais américain/American" (OXHA, 2001), but its use is confused and confusing because it is attached to examples of both United States and North American regional usage. This confusion can be seen in the entries for *premier étage* where both CR and OXHA attach the *US* label to the equivalent "second floor", which is, in fact, North American, being used in both the United States and Canada. This confusion would be avoided by the use of three distinct labels identifying the three major regional varieties of North American English, namely, North American, United States and Canadian.

5.2.2.3 In the BCD

The ten sample entries from the BCD (see Appendix) illustrate the way in which each aspect of the Dictionary's stated policies on geographic labelling (examined in Chapter 4) is implemented. The sample entries have been chosen to illustrate usage of all five of the BCD's geographic labels and provide evidence of the fact that, as stated in BCD policy documents, a geographic label is attached to all source-language regionalisms that are included in the Dictionary.

In addition to stating that geographic labels are attached to all source-language regionalisms in the BCD, the Dictionary's policy also makes it clear that all regionalisms in the target language are similarly labelled. The implementation of this policy statement is illustrated in the entries for *basket*, *fête du Canada* and *premier étage* where a label is used to identify the geographic regions of the target-language items "running shoe" *(CD)*, "runner" *(esp. CD)*, "sneaker" *(esp. NA)*, "trainer" *(GB)*, "Canada Day" *(CD)*, and "first floor", "ground floor" and "second floor" (labelled *(NA)* and *(GB)*, depending on meaning).

BCD policy states, moreover, that a geographic label is attached not only to linguistic regionalisms, but also to cultural regionalisms. This statement is corroborated in the entry for *fête du Canada*, where both the source-language item and the target-language item ("Canada Day") are cultural regionalisms. As these items denote a reality existing only in Canadian culture, they are both identified by the *CD* label.

As well as setting out the parameters for the inclusion of a geographic label, the BCD's policy documents also state that what remains unmarked by any geographic label in the Dictionary are those elements of French and English that are common to both the Canadian variety and to the other major varieties of each language. This aspect of BCD policy is illustrated by the fact that, amongst the ten sample entries, there are certain senses of the source-language items as well as several target-language items which, being lexical items common to North American, Canadian and British varieties of English and to Canadian French and French from France, are not marked by any geographic label. In the sample entries, those senses unmarked for geographic region are *accommoder* in the sense divisions numbered 4, 5, 6, and 7, *basket* in the first sense division (with the meaning of "basket-ball"), *muffler* in sense divisions 2 ("scarf") and 3 ("device used to reduce noise"); and *silencer* in the first sense division (with the meaning of "device used to reduce noise"). In the entries for *accommoder*, *attorney*, *riding* and *silencer*, all the target-language items are unmarked for geographic region; in the entry for *muffler*, all the target-language items except one ("foulard") are unmarked for geographic region, and in the entry *basket*, the target-language item in sense 1 ("basketball") is unmarked for geographic region.

5.2.3. Treatment of Equivalents as Exemplified in Dictionary Entries

Having examined how the policies of the three bilingual dictionaries are implemented in the ten sample dictionary entries with regard to the identification and selection of regionalism and the geographic labelling of regionalism, the following section of this study will examine to what extent the sample entries confirm the dictionaries' stated policies on the inclusion of regionalisms in the target language.

5.2.3.1 Regionalisms as Equivalents in CR and OXHA Entries

As was discussed in Chapter 4, the editors of CR do not explicitly state any policy on the inclusion of regionalisms as target-language equivalents, but give two examples in the guide to the use of the dictionary highlighting lexical differences in English between British and U.S. usage. It is entirely consistent with this emphasis that, in the sample entries studied, regionalisms are only included as target-language equivalents in CR when they reflect British-U.S. lexical differences. Thus, the equivalents CR offers for the French lexical item *baskets* include two British regionalisms ("trainers" and "basketball boots") and two U.S. regionalisms ("tennis shoes" and "high-tops"). Likewise, for the French collocation *au premier étage* in the sense used in France of "floor above the ground floor", CR offers as equivalents the British regionalism "on the first floor" and the U.S. regionalism "on the second floor" and in the sense used in Canada of "au rez-de-chaussée" offers as equivalents the British regionalism "on the ground floor" and the U.S. regionalism "on the first floor".

Like CR, OXHA includes regionalisms as target-language equivalents in the case of British-U.S. orthographic and lexical variation, but by contrast with CR, the editors of OXHA state in the introductions to the first and third editions of the dictionary (see Chapter 4) that this practice forms part of the lexicographic policy of the dictionary. The implementation of this policy statement is illustrated in the sample entry for the French lexical item *le premier étage*, for which OXHA offers as equivalents the British regionalism "the first floor" and the U.S. regionalism "the second floor".

5.2.3.2 Treatment of Equivalents in BCD Entries

As was discussed in Chapter 4, BCD policy documents state that the equivalents included on both sides of the Dictionary are generally only those lexical items used in North America. These items fall into two categories: those identified by the geographic labels *NA*, *US* and *CD* and those that remain unmarked by any geographic label. For the marked items, the implementation of this policy is illustrated in the sample entries by the inclusion as equivalents of the Canadian regionalisms "Canada Day", "running shoe", "runner", "association de circonscription" and "foulard", and by the North American regionalisms "sneaker", "first floor" (in the sense of "ground floor") and "second floor" (in the sense of "floor above the ground floor"). All the other equivalents in the sample entries are unmarked for geographic region since they are elements of English and French that are common to both the Canadian variety and to the other major varieties of each language.

Having stated that only lexical items used in North America will normally be included as equivalents in the Dictionary, the policy of the BCD is sufficiently flexible to allow the inclusion as equivalents of regionalisms from a geographic region other than North America when it is considered to be especially helpful to the user. Thus, the British-only lexical item "trainer", for example, is offered as an equivalent for *basket*, since it does occur in a few texts of British origin that are reproduced in articles taken from certain Canadian publications in the BCD Canadian corpora. In the case of the lexical item *premier étage*, the British regionalisms "ground floor" (in the sense of "rez-de-chaussée") and "first floor" (in the sense of "floor above the ground floor") are included alongside the North American equivalents in order to provide as much helpful information as possible for the user and to avoid possible confusion.

5.2.3.2.1 Canadianisms as Equivalents in BCD Entries

Although the special Canadian focus of the BCD guarantees the inclusion of many Canadian regionalisms as source-language items, this does not necessarily mean that they all appear as target-language equivalents on the other side of the Dictionary. As was seen in Chapter 4 (4.2.1.2.1) in the discussion of the BCD's policy on the identification and selection of Canadianisms, a large number of French Canadianisms do not appear in the BCD either as source- or as target-language items because they are informal in register and their use is

criticized by l'Office québécois de la langue française and other language experts. The BCD, like all dictionaries, has to make a selection from all the material that could potentially be included in every entry and, because it is general dictionary, it makes it a priority to present as complete a coverage as possible of standard-register items.

As far as Canadianisms are concerned, BCD policy documents state that they are included as target-language equivalents in the Dictionary according to three main criteria: frequency in the BCD Canadian corpora, linguistic productivity and when no other lexical item exists in the target language that corresponds exactly to the concept expressed by the source-language item. An analysis according to these three criteria of the five Canadianisms included as equivalents in the sample entries confirms the implementation of the BCD's stated policy, as follows:

(i) ***if the Canadianism is more frequent in the BCD Canadian corpora***: the Canadianism *foulard* occurs twice as frequently in the BCD French Canadian sub-corpus as the standard French word *écharpe*, which justifies its position as the first of four equivalents offered for the English noun *muffler*.

(ii) ***if the Canadianism is linguistically productive***: *running shoe,* having been adopted as the commonest way of denoting in Canadian English "any of various shoes having an upper made of nylon, canvas, etc. and a rubber or synthetic sole" (OXCAN), was shortened to *runner*[35] by the normal linguistic process in which many of the most frequently used lexical items, especially in the less formal registers, are reduced to a simpler and shorter form. In this way, an extra meaning has been added to the noun *runner*. Because of their linguistic productivity, therefore, these two Canadianisms are considered worthy of inclusion as equivalents for the French noun *basket* and are given before the other equivalents.

(iii) ***if the Canadianism is the only lexical item existing in the target language that expresses exactly the concept denoted by the source-language item***: the Canadianism "Canada Day" is given in the BCD as the sole equivalent for the French lexical item *fête du Canada* because it is the only lexical item that exists in the English language to denote this concept of "the annual holiday commemorating the creation of the Dominion of Canada" (OXCAN).

---

[35] "Runner" is labelled "*esp. CD*" in the BCD because although it is most widely used in Canadian English, it is also used in other varieties of English, such as Irish.

5.3 Conclusion

In order to answer the question of *why* a bilingual Canadian dictionary is needed, Chapter 4 studied three aspects of the lexicographic treatment of regionalism in general and of Canadian usage in particular, according to the front matter of two European-produced bilingual English-French dictionaries, the *Collins-Robert* and the *Oxford-Hachette*. The aim of Chapter 5 has been to complete the answer to this question by studying the same three aspects, but this time according to the evidence provided by an analysis of ten sample lexical items in these two dictionaries.

Regarding the identification and selection of regionalisms as source-language items, CR and OXHA both have identical coverage of the five sample English lexical items, including in their nomenclatures the U.S. regionalisms *attorney* and *muffler*, and the British regionalism *silencer*, but omitting the Canadianisms *riding* and *riding association*. On the French side, both dictionaries provide identical coverage of four of the regionalisms as source-language items, including in their nomenclatures the regionalisms from France *basket* and *premier étage*, but omitting the Canadianisms *accommoder* and *fête du Canada*. It is only in the case of the Canadian regional usage of *premier étage* that the coverage of the two dictionaries varies in that CR includes this item whereas OXHA does not.

As far as geographic labelling is concerned, analysis of the sample entries in CR and OXHA reveals that both dictionaries provide appropriate geographic labelling of source-language and target-language items from the United States and Great Britain. However, because neither dictionary labels examples of French usage that are exclusive to France nor clearly differentiates North American, U.S. and Canadian usage, the user sometimes lacks the detailed pointers he may need to enable him to make the most informed choices.

The evidence of the ten sample entries in English and French has shown that both dictionaries do include regionalisms as equivalents, but that this coverage is limited to examples of regional usage from the United States and Great Britain. Examples of specifically Canadian usage are not included as equivalents in either dictionary.

This analysis of the ten sample entries has shown that CR and OXHA generally serve the user well in providing a certain amount of information on regional usage by including some regionalisms as source-language items, by attaching geographic labels and glosses to certain regionalisms and by the inclusion of some regionalisms as equivalents. However, the comparative analysis of the ten sample entries in the BCD has shown that, as is to be expected from a dictionary specializing in the regional usage of English and French, the BCD's treatment of regionalism in all three of the aspects examined is far more detailed than that of the two European-produced dictionaries. The BCD has demonstrated that it is the more finely honed linguistic tool, a tool that enables the user to pinpoint regional usage with a high degree of precision.

CONCLUSION

The primary objective of this research project was to present an overall evaluation of the Bilingual Canadian Dictionary Project and of the extent to which the Project has achieved its principal aims. I have been privileged to have access to a large collection of BCD documentation, published and unpublished, as well as to the Project's corpora, database and lexicographic tools. My study of this material and my personal knowledge of these systems have enabled me to present here a comprehensive description of the Dictionary, its evolution, structure and production methods.

I have been fortunate also to have been part of a large team of many dedicated lexicographers, including both professors and students of translation studies, and professional translators and terminologists. We have all benefited enormously from the knowledge and experience gained as members of the Project and we now swell the numbers of trained lexicographers working in Canada. The BCD Project has indeed fully achieved its principal aims of developing Canadian expertise in bilingual lexicography, establishing an electronic corpus of Canadian texts in English and French and of producing a multiple-purpose dictionary database. The publication of the Dictionary will be the culmination of this substantial collaborative effort.

The use of the electronic corpus has been of prime importance in the production of the BCD and there is no doubt that its use has considerably enhanced the quality of the Dictionary's entries, enabling the reliable identification of regional senses and the suggestion and confirmation of equivalents. The construction from scratch of the BCD's own corpora of English and French texts was a mammoth task and one that required a huge investment of time and energy in the initial stages of the Project. This investment has, however, paid very handsome dividends as the Dictionary can now confidently boast that its lexicographers have been able to make authoritative choices regarding both source- and target-language lexical items based on their analysis of examples of authentic language. The conviction of BCD Project leaders as early as 1988 that an electronic corpus of authentic texts was an essential requirement for the production of a reliable bilingual dictionary has been amply corroborated over the intervening years, not only by lexicographical theoreticians, but also by the fact that it is rare today for a new

dictionary to appear on the market that does not lay claim to having used a corpus of texts to a greater or lesser degree.

Chapter 3 of this study reviewed the significant events in the evolution of lexicography in Canada and concluded that no up-to-date bilingual French and English dictionary produced in Canada is currently available. The BCD Project occupies an important position, therefore, as the first project since the 1970s designed to produce a bilingual dictionary recording the distinctive features of the English and French languages in Canada, as well as those features that Canadian English and Canadian French share with varieties of the languages used elsewhere in the world. The 2001 Canadian census reported that almost 6.8 million people stated French as their mother tongue and 17.5 million people English. These figures represent a significant market for a bilingual dictionary that puts the presentation of Canadian English and French as its top priority. The comparison in this study with two European-produced dictionaries has revealed that, by its thoroughgoing presentation of Canadian English and French, the BCD will also be of great use to those outside Canada who seek to position Canadian varieties of English and French in their wider international context alongside other regional forms.

The practical exercise of comparing a selection of dictionary entries in the BCD with the same entries in two European-produced French-English bilingual dictionaries has shown that the two European-produced dictionaries do include some examples of regional usage, but that incomplete or inaccurate geographic labelling often renders their entries unhelpful and confusing to users. Neither CR nor OXHA offers, for instance, an explanation of why examples of exclusively British usage and exclusively North American usage have geographic labels attached, when examples of usage specific to France do not. Similarly, it is clear from analysis of the front matter and of the sample entries in CR and OXHA that the emphasis in their selection of regionalisms is based primarily on lexical and orthographical variation between British and North American English, although no information is provided on how these regionalisms are identified or selected.[36]

---

[36] My research inclines me towards the hypothesis that, of the relatively few examples of Canadian usage included in CR and OXHA, Canadian French is better represented than Canadian English. Unfortunately, because of the lack of information in the dictionaries' front matter on the identification and selection of regionalisms, it is difficult to be precise. This would be an interesting avenue to pursue in further research.

The comparison of the BCD with the two European-produced bilingual dictionaries has clearly shown that, although CR and OXHA are both excellent dictionaries, they are not, despite claims in their front matter that may suggest the contrary, dictionaries specializing in regionalism. These are dictionaries designed principally for English-French dictionary users in Europe and, to some extent, the United States and, as such, often fail to take account of Canadian usage in both source and target languages and, therefore, do not meet the needs of English- and French-speaking Canadians. By including a large number of Canadianisms in both source- and target languages and by its systematic labelling of examples of Canadian usage as well as of other regional usage in both source and target languages, the BCD, on the other hand, offers the user far greater accuracy in pinpointing the regional lexical items and equivalents for which he is searching. The BCD provides the reliable tool that Canadians need in order to understand and communicate effectively in both official languages, in their own country as well as abroad.

APPENDIX: DICTIONARY ENTRIES

I. Entries from the Bilingual Canadian Dictionary

English entries:

attorney *n (esp. US) (Droit)***1** *(lawyer)* avocat *m, f avocate.*

***the suit has been turned over to our attorneys** la cause a été remise à nos avocats; **calls to the movie studios and the singer's attorney were not returned** les appels que nous avons faits aux studios de cinéma et à l'avocat du chanteur ne nous ont pas été retournés.

muffler *n* **1** *(device used to reduce the noise of a motor) (NA)* silencieux *m*, pot d'échappement *m.*

***cars with no mufflers** des voitures sans silencieux.

2 *(scarf)* foulard *m (CD)*, écharpe *f*, cache-nez *m invar*, cache-col *m.*

***a girl wearing a muffler around her neck** une fille portant un foulard autour du cou.

3 *(device used to reduce noise)* silencieux *m*, assourdisseur de bruit *m.*

riding *n* **1** *(constituency) (CD)* circonscription électorale *f*, circonscription *f*, comté *m.*

***a southern Alberta riding** une circonscription etc. du sud de l'Alberta.

riding association *n (CD)* association de circonscription *nf (CD).*

silencer *n* **1** *(device used to reduce noise)* silencieux *m*, assourdisseur *m*, insonorisateur *m*; *(of a firearm)* silencieux *m.*

***roof vents have been equipped with silencers to reduce noise** les bouches d'aération du toit ont été équipées d'assourdisseurs; **a gun equipped with a silencer** un fusil muni d'un silencieux.

2 *(device used to reduce the noise of a motor) (GB)* silencieux *m*, pot d'échappement *m.*

French entries:

accommoder *vt* 1 *(rendre service à qn, répondre aux besoins de qn, convenir) (CD)* to accommodate sb, to suit sb.

***la réunion a été reportée pour accommoder un des membres du comité** the meeting was postponed to accommodate one of the committee members;

il m'a accommodé en me prêtant sa voiture he helped me out by lending me his car;

pour accommoder les congressistes, on a prévu des ordinateurs dans chaque salle computers have been set up in every conference room for delegates' convenience.

2 *(contenir)(CD)* to accommodate, to hold, to fit.

***la cafétéria peut accommoder 1 200 personnes** the cafeteria can accommodate up to 1,200 people, the cafeteria can hold up to 1,200 people, the cafeteria seats 1,200.

3 *(héberger)(CD)* to accommodate, to put up *(informal)*.

***plusieurs familles ont offert d'accommoder les sinistrés** many families have offered to accommodate the disaster victims, many families have offered to house the disaster victims, many families have offered to put up the disaster victims.

4 *(préparer des aliments, des plats) (Cuis)* to prepare (**à**, in, with), to season, to cook, to dress, to do . . .

5 *(adapter)* to adapt, to fit, to adjust, to accommodate . . .

6 *(arranger, régler)[ses affaires]* to arrange; *(réconciler) [ennemis]* to reconcile, to bring together; *[personne]* to accommodate, to make comfortable . . .

7 *(ridiculiser qn)* to slander, to criticize, to ridicule, to give sb a good dressing-down, to tear sb to pieces (behind his back); *(maltraiter, frapper)* to give harsh treatment to, to beat . . .

basket *nm* **1** *(Sport)* basketball.

***je joue au basket depuis l'âge de neuf ans** I have been playing basketball since the age of nine.

2 *(chaussure de sport)* **basket** *nf (FR)* running shoe *(CD)*, runner *(esp. CD)*, sneaker *(esp. NA)*, trainer *(GB)*.

***il enfile ses baskets** he puts on his running shoes etc.

fête du Canada:

Canada *nm* . . . (**cmp2**) . . . **fête du Canada** *nf (le 1^er^ juillet) (CD)* Canada Day *n (CD).*

premier étage:

étage *nm* . . . ****premier étage** *(rez-de-chaussée) (CD)* first floor *(NA)*, ground floor *(GB)*;
(*étage au-dessus du rez-de-chaussée*) *(FR)* second floor *(NA)*, first floor *(GB).*

II. Entries from the *Collins-Robert French Dictionary*

English entries:

attorney *n* **a** *(Comm, Jur = representative)* . . .

b (*US:* also **attorney-at-law**) avocat(e) *m(f)* . . .

muffler *n* **a** (= *scarf*) cache-nez *m inv*, cache-col *m inv*

b (*US Aut*) silencieux *m*

silencer *n* (*on gun, Brit: on car*) silencieux *m.*

French entries:

basket *nm* (*Sport*) basketball

nf (= chaussure) **baskets** (*gén*) sneakers, trainers (*Brit*), = tennis shoes (*US*); (*pour joueur*) basketball boots (*Brit*), high-tops (*US*)

◆ **être à l'aise dans ses baskets*** to be at ease with o.s.; ➔ **lâcher**

étage *nm* . . . ◆ **au premier étage** (*en France*) on the first floor (*Brit*), on the second floor (*US*); (*au Canada*) on the ground floor (*Brit*), on the first floor (*US*).

III. Entries from the *Oxford-Hachette French Dictionary*

English entries:

attorney *noun* **1** *US* (*lawyer*) avocat *m* . . .

muffler *n* **1** *Fashn* cache-nez *m inv*;

2 *US Aut* silencieux *m*.

silencer *n Mil, GB Aut* silencieux *m*.

French entries:

basket *nm* **1** (*sport*) basketball;

2 (*chaussure de sport*) trainer.

Idiomes

lâcher les baskets à qn[!] to give sb a break[!];

être bien *or* **à l'aise dans ses baskets[!]** to be very together[!].

étage *nm* . . . **le premier étage** the first floor *GB*, the second floor *US*.

BIBLIOGRAPHY

1. Theoretical Works

Ainsworth, J. Rev. of *NTC's Dictionary of Canadian French. Modern Language Journal* 76.2 (1992): 251.

Al-Kasimi, Ali M. *Linguistics and Bilingual Dictionaries*. Leiden: E.J. Brill, 1977.

Avis, Walter S. "Canadian English in its North American Context." *Walter S. Avis: Essays and Articles*. Comp. T. Vincent et al. Kingston: Royal Military College of Canada, 1978. 35-49.

- - - . "Problems in the Study of Canadian English." *Walter S. Avis: Essays and Articles*. Comp T. Vincent et al. Kingston: Royal Military College of Canada, 1978. 3-12.

Barnhart, Clarence L. "Problems in editing commercial monolingual dictionaries." *Problems in Lexicography*. Ed. F. Householder and S. Saporta. Bloomington: Indiana University Press, 1967. (17-22).

Béjoint, Henri. "Scientific and Technical Words in General Dictionaries." *International Journal of Lexicography*. 1 (1988): 354-368.

Bell, Roger T. *Translation and Translating: Theory and Practice*. London: Longman, 1991.

Bouchard, Chantal. *La langue et le nombril. Histoire d'une obsession québécoise*. Montréal: Fides, 1998.

Boulanger, Jean-Claude. "Le pacte normatif du français québécois: réflexions sur les marques lexicographiques diatopiques." *Les marques lexicographiques en contexte québécois : Actes de la table ronde tenue à Montréal les 3 et 4 novembre 1994*. Ed. L. Mercier and C. Verreault. Montréal: Office de la langue française, 1998. 171-187.

Bowker, L. *Computer-Aided Translation Technology: A Practical Introduction*. Ottawa: University of Ottawa Press, 2002.

- - - . "Specialized lexicography and specialized dictionaries." *A Practical Guide to Lexicography*. Ed. P. G. J. van Sterkenburg. Amsterdam: Benjamins, 2003. 154-164.

Bowker, L., and J. Pearson. *Working with Specialized Language: A Practical Guide to Using Corpora*. London: Routledge, 2002.

Brazil, David. "The development of English as a world language." *Collins English Dictionary*. Glasgow: HarperCollins, 2000. xxvi-xxxiii.

Burkhanov, I. "Pragmatic specifications: usage indications, labels, examples; dictionaries of style, dictionaries of collocations." *A Practical Guide to Lexicography*. P. G. J. van Sterkenburg. Amsterdam: Benjamins, 2003. 102-113.

Clas, André. "Un dictionnaire homoglosse et hétéroglosse : le dictionnaire canadien bilingue." Allocution présentée à l'École des hautes études commerciales de Montréal, le 27 mars 2002.

Clas, André, and Roda P. Roberts. "Le dictionnaire bilingue : une mosaïque culturelle?". *Les écarts culturels dans les dictionnaires bilingues : Actes des 3e journées d'études sur la lexicographie bilingue, organisées par l'INALCO, Institut national des langues et civilisations orientales, Paris, les 12 et 13 octobre 2000*. Ed. T. Szende. Paris: Honoré Champion, 2003. 239-247.

Corbeil, Jean-Claude. "Les marques d'usage comme technique de description des aspects connotatifs du lexique." *Les marques lexicographiques en contexte québécois : Actes de la table ronde tenue à Montréal les 3 et 4 novembre 1994*. Ed. L. Mercier and C. Verreault. Montréal: Office de la langue française, 1998.

Darbelnet, Jean. *Le français en contact avec l'anglais en Amérique du nord*. Sainte-Foy: Les Presses de l'Université Laval, 1976.

Cormier, Monique C., and Roda P. Roberts. "Lexicographie comparée du français et de l'anglais au Canada : *le Dictionnaire canadien bilingue*." *Approches contrastives en lexicographie bilingue*. Ed. T. Szende. Paris: Honoré Champion, 2000. 213-222.

De Villers, Marie-Éva. "Les marques lexicographiques : des points de repère essentiels pour l'usage des mots." *Les marques lexicographiques en contexte québécois : Actes de la table ronde tenue à Montréal les 3 et 4 novembre 1994*. Ed. L. Mercier and C. Verreault. Montréal: Office de la langue française, 1998. 157-169.

- - - . "Les innovations de la 4e edition du *Multidictionnaire*." Chroniques linguistiques, HEC Montréal. 2000. August 2006 <http://neumann.hec.ca/~x067/chroniqueslinguistiques/multiquatre.html>.

Halliday, M.A.K. *Language as Social Semiotic: the social interpretation of language and meaning*. London: Arnold, 1978.

Hamilton, S. "Canadianisms and their treatment in dictionaries." Unpublished thesis. University of Ottawa, 1997.

Harris, B. "Bi-Text: A New Concept in Translation Theory." *Language Monthly*. 54 (1988): 8-10.

Hausmann, F. J. et al., ed. *International Encyclopedia of Lexicography. Vol. I-III*. Berlin: W. de Gruyter, 1989-1991.

Jackson, H. *Lexicography. An Introduction*. London: Routledge, 2002.

Jakobson, R. "On Linguistic Aspects of Translation." *On Translation*. Ed. R. Brower. Oxford: Oxford University Press, 1966. 232-239.

Josselin-Leray, Amélie. "Place et rôle des terminologies dans les dictionnaires généraux unilingues et bilingues. Étude d'un domaine de spécialité : volcanologie." Unpublished PhD Thesis. Université Lyon II, France, 2005.

Laffling, J. *Towards High-Precision Machine Translation Based on Contrastive Textology*. Berlin: Foris Publications, 1991.

Landau, Sidney I. *Dictionaries: The Art and Craft of Lexicography*. Cambridge: Cambridge University Press, 2001.

L'Homme, M.C. *Terminologie : principes et techniques*. Montréal: Les Presses de l'Université de Montréal, 2004.

Mercier, Louis. *La société du parler français au Canada et la mise en valeur du patrimoine linguistique québécois (1902-1962)*. Quebec: Les Presses de de l'Université Laval, 2002.

Nida, Eugene. "Principles of Correspondence." *The Translation Studies Reader*. Ed. L. Venuti. London: Routledge, 2000. 126-140.

Nida, Eugene, and Charles Taber. *The Theory and Practice of Translation*. Leiden: Brill, 1982.

Office québécois de la langue française. "Noms féminins en -eure." 2 June 2005 <http://w3.oqlf.gouv.qc.ca/bdl/gabarit_bdl.asp?t1=1&id=3336&t2.x=f%E9minin>

- - - . *Répertoire des avis linguistiques et terminologiques,* 3e édition revue et augmentée (171-186). Quebec: Publications du Québec, 1990.

Orkin, Mark M. *Speaking Canadian English*. Toronto: General Publishing Company Limited, 1970.

- - - . Mark M. *Speaking Canadian French*. Toronto: General Publishing Company Limited, 1971.

Petrin, L. "Two New Dictionaries Prove Canadian Language Flourishes." *The Globe and Mail* 14 Apr. 1962: 17.

Poirier, Claude. "Problèmes et méthodes d'un dictionnaire général du français québecois". *Revue québécoise de linguistique théorique et appliquée* 7.1 (1988): 13-54.

Poirier, Claude. "Les québécismes dans le dictionnaire: marqués ou non marqués." *Québec français* 79 (1990): 90-91.

Roberts, Roda P. *Bilingual Canadian Dictionary: Methodology for Research Assistants*. Version 8.1. University of Ottawa, 1997.

- - - . "Bilingual dictionaries prepared in terms of translators' needs." *Translation in the Global Village. Proceedings of the 3rd Conference of the Canadian Translators and Interpreters Council, May 4-8, 1994, Banff, Alberta.* Ed. Nésida Loyer. Ottawa: Canadian Translators and Interpreters Council, 1994. 113-119.

- - - . "Definition, Identification and Selection of French Canadianisms in the Bilingual Canadian Dictionary." *Circuit* 63 (1999): 4-6.

- - - . "Grant Proposal for the BCD Project." University of Ottawa, 1993.

- - - . "Methods of bilingual dictionary making: the Canadian experience." *Symposium on Lexicography V. Proceedings of the Fifth International Symposium on Lexicography, May 3-5, 1990, University of Copenhagen.* Ed. Karl Hyldgaard-Jensen and Arne Zettersten. Tübingen: Max Niemeyer Verlag, 1992. 65-71.

- - - . "Lexicographie bilingue du français et de l'anglais au Canada." *Terminogramme : bulletin de la Direction de la terminologie* 80 (1996): 11-16.

- - - . "Organization of information in a bilingual dictionary entry." *Proceedings of the Fourth Euralex International Congress-Euralex '90.* Ed. Bibliograf and Commission of the European Communities. Barcelona: Bibliograf, 1992. 219-231.

- - - . "Parallel text analysis and bilingual lexicography." 1996. <http://www.dico.uottawa.ca/articles-en.htm>.

- - - . "Translation and the bilingual dictionary." *Meta* XXXV.1 (1990): 74-81.

Roberts, Roda P., and J. Bossé-Andrieu. "Le Dictionnaire canadien bilingue et les normes linguistiques." *La représentation de la norme dans les pratiques terminologiques et lexicographiques : Actes du colloque tenu les 14 et 15 mai 2001 à l'Université de Sherbrooke dans le cadre du 69e Congrès de l'ACFAS.* Ed. P. Bouchard and M. Cormier. Montréal: Office québécois de la langue française, 2002. 233-242.

- - - . "Présentation de la diversité lexicale : variations dans les pratiques dictionnairiques." *La variation dans la langue standard : Actes du colloque tenu les 13 et 14 mai 2002 à l'Université Laval dans le cadre du 70e Congrès de l'ACFAS.* Ed. P. Bouchard. Montréal: Office québécois de la langue française, 2004. 141-166.

Roberts, Roda P., L. Langlois, and D. Megginson. "SGMLizing the Bilingual Canadian Dictionary: Reasons, Processes, and Problems." *Proceedings of the Joint International Conference of the Association for Computers and the Humanities and the Association for Literary and Linguistic Computing, June 3-7, 1997, Queen's University, Kingston, Ontario.*

Roberts, Roda P., and C. Montgomery. "The use of corpora in bilingual lexicography." *Proceedings of the Seventh Euralex International Congress-Euralex '96.* Ed. M. Gellerstam et al. Göteborg: Göteborg University, 1996. 457-464.

Schleiermacher, F. "On the different methods of translating." *Theories of Translation.* Ed. R. Schulte and J. Biguenet. Chicago and London: University of Chicago Press, 1992. 3654.

Sinclair, J. "Corpora for lexicography." *A Practical Guide to Lexicography.* Ed. P. G. J. van Sterkenburg. Amsterdam: Benjamins, 2003. 167-178.

- - - . "Corpus processing." *A Practical Guide to Lexicography.* Ed. P. G. J. van Sterkenburg. Amsterdam: Benjamins, 2003. 179-193.

Snell-Hornby, M. "Towards a Learner's Bilingual Dictionary." *The Dictionary and the Language Learner.* Ed. A. Cowie. Tübingen: M. Niemeyer, 1987. 159-170.

Svensén, Bo. *Practical Lexicography: Principles and Methods of Dictionary-Making.* Oxford: Oxford University Press, 1993.

Varantola, Krista. "Translators and Disposable Corpora." *Corpora in Translator Education.* Ed. F. Zanettin et al. Manchester: St Jerome, 2003. 55-70.

Wiegand, H. E. "On the structure and contents of a general theory of lexicography." Lexeter '83 *Proceedings.* Ed. R.R.K. Hartmann. Tübingen: Max Niemeyer Verlag, 1984.

Yule, George. *The Study of Language.* Cambridge: Cambridge University Press, 2006.

Zanettin, Frederico. "DIY Corpora: The WWW and the Translator." *Training the Language Services Provider for the New Millennium.* Ed. M. Belinda et al. Porto: Faculdade de

Letras, Universidade do Porto, 2002. 239-248.

http://www.federicozanettin.net/DIYcorpora.htm

Zgusta, L. *Manual of Lexicography*. The Hague: Mouton, 1971.

2. Dictionaries

Canadian Oxford Dictionary. Ed. K. Barber. Toronto: Oxford University Press, 2005.

Collins Cobuild English Dictionary. London: HarperCollins, 1995.

Collins English Dictionary. Ed. D. Treffry. Glasgow: HarperCollins, 2000.

Collins-Robert French Dictionary. New York: HarperCollins, 1993, 1995, 1998, 2002.

Collins-Robert French-English, English-French Dictionary. London: Collins, 1978, 1987.

Dictionary of Canadian French. Ed. Sinclair Robinson and Donald Smith. Toronto: Stoddart,1990.

Dictionary of Canadianisms on Historical Principles. Ed. W. Avis et al. Toronto: Gage, 1967.

Dictionary of Lexicography. Ed. R.R.K. Hartmann and G. James. London: Routledge, 1998.

Dictionary of National Biography. Vol. XLI. London: Smith, Elder, & Co., 1895.

Dictionary of Newfoundland English. Ed. G.M. Story et al. Toronto: University of Toronto Press, 1982.

Dictionary of Prince Edward Island English. Ed. T.K. Pratt. Toronto: University of Toronto Press, 1988.

Dictionnaire des canadianismes. G. Dulong. Montréal: Larousse, 1989.

Dictionnaire canadien/The Canadian Dictionary. Ed. J.P. Vinay et al. Toronto: McClelland & Stewart, 1962.

Dictionnaire canadien-français. Sylva Clapin. Montréal: Beauchemin, 1894. Sainte-Foy: Les Presses de l'Université Laval, 1974.

Dictionnaire du français plus. Ed. Claude Poirier. Quebec: CEC, 1988.

Dictionnaire historique du français québécois. Ed. Claude Poirier et al. Sainte-Foy: Les Presses de l'Université Laval, 1998.

Dictionnaire de la langue québécoise. Ed. L. Bergeron. Montréal: vlb éditeur, 1980.

Dictionnaire des locutions vicieuses du Canada avec leur correction suivi d'un dictionnaire canadien. Ed. Joseph Amable Manseau. Québec: J.A. Langlais, 1881.

Dictionnaire nord-américain de la langue française. Ed. L.-A. Bélisle. Montréal: Beauchemin, 1986.

Dictionnaire québécois d'aujourd'hui. Ed. J.-C. Boulanger. Montréal: Dicorobert, 1992.

Dictionnaire québécois d'aujourd'hui. Ed. J.-C. Boulanger. Montréal: Dicorobert, 1993.

Dictionnaire québécois français. Ed. Lionel Meney. Montréal: Guérin, 1999.

Funk & Wagnall's Canadian College Dictionary. Toronto: Fitzhenry and Whiteside, 1989.

Gage Canadian Dictionary. Ed. W. Avis et al. Toronto: Gage, 1983.

Glossaire franco-canadien du français au Canada. Ed. Oscar Dunn. Québec: 1880. Sainte-Foy: Les Presses de l'Université Laval, 1976.

Glossaire du parler français au Canada. Société du Parler français au Canada. Québec: 1930. Sainte-Foy: Les Presses de l'Université Laval, 1968.

Lexis: dictionnaire de la langue française. Ed. Jean Dubois et al. Paris: Larousse, 1987.

Manuel des difficultés les plus communes de la langue française adapté au jeune âge et suivi d'un recueil de locutions vicieuses. Ed. Thomas Maguire. Québec: Fréchette et Cie, 1841.

Manuel des expressions vicieuses les plus fréquentes. Ed. J. F. Gingras. Outaouais: Imprimerie du "Canada", 1867.

Multidictionnaire de la langue française. Montréal: Québec Amérique, 1997.

Néologie canadienne ou Dictionnaire des mots créés au Canada. Ed. Jacques Viger ms. 1810. Ottawa: Les Presses de l'Université d'Ottawa, 1998.

New Dictionary of Americanisms, Being a Glossary of Words Supposed to be Peculiar to the United States and the Dominion of Canada. Sylva Clapin. New York: 1902.

Nouveau dictionnaire français-anglais et anglais-français. Sylva Clapin. Montréal: Beauchemin, 1905.

Le nouveau petit Robert : dictionnaire alphabétique et analogique de la langue française. Ed. Josette Rey-Debove and Alain Rey. Paris: Dictionnaires Le Robert, 2003.

Nugent's improved pocket-dictionary of the French and English languages: in two parts, French and English, English and French; containing all the words in general use and the usual technical terms, the accentuation and pronunciation adapted to the French and English idiom. Ed. Thomas Nugent. Montréal: D. & J. Sadlier, 1872.

Oxford-Hachette French Dictionary. Oxford: Oxford University Press, 1994 1997, 2001.

Penguin Canadian Dictionary. Ed. T. Paikeday. Mississauga: Penguin, 1990.

Petit vocabulaire à l'usage des Canadiens-français. Napoléon Caron. Trois-Rivières: Journal des Trois-Rivières, 1880.

Practical Handbook of Quebec and Acadian French. Sinclair Robinson and Donald Smith. Toronto: Anansi, 1984.

Québécois Dictionary. Ed. L. Bergeron. Toronto: Lorimer, 1982. TEXTUM, the BCD Bilingual Comparable Corpus

Random House Webster's College Dictionary. Ed. R.B. Costello et al. New York: Random House, 1992.

Richesses et Particularités de la langue écrite du Québec. Ed. A. Clas, E. Seutin et al. Montréal: Université de Montréal, 1980.

Le Robert & Collins dictionnaire français-anglais, anglais-français senior. Ed. Beryl T. Atkins et al. Paris/London: Le Robert/Collins, 1993.

Webster's Ninth New Collegiate Dictionary. Ed. Frederick F. Mish. Markham, ON: Thomas Allen & Son Ltd, 1987.