Selected linguistic problems in indexing within the Canadian context

Lisa Rasmussen

Study of the problems inherent in indexing within a Canadian context. Takes into account the linguistic characteristics of Canadian English (the divided usage between British and American spelling and vocabulary; the literary warrant of words of Canadian origin) and of Canadian French (the frequency of vocabulary, morphological, and semantic anglicisms; the differences in vocabulary between standard and Canadian French) and the problems involved in bilingual indexing because of the trend in the English language towards nominalization.

Introduction

The effectiveness of most printed material depends, to a large extent, upon the adequacy of an index. A good index will lead the reader to the information sought in a quick and forthright manner, whereas a poor index may cause important facts to remain virtually inaccessible.

To complete their work, indexers rely heavily on the vocabulary of the author or the translator; indexers can make the concepts expressed within the text available to readers only if authors have expressed themselves clearly, or if translators have understood the authors’ language and faithfully reproduced their ideas.

Indexers use the natural language of the work or, if there is a chance that the natural language terminology may be unfamiliar to readers, provide a code to simplify the search. Readers will, of course, go first to the terminology most familiar to them, then to whatever possible alternative subject headings they can think of. The onus is on indexers to consider every possible legitimate alternative to the entry term, and to provide a lead-in vocabulary of synonyms which can make the jump from the reader’s vocabulary to that used in the text.

The purpose of this paper is to explore some of the linguistic problems indexers might encounter when indexing a text written in Canada, in either English or French, or translated from any language into English. It will look at some of the ways indexers can facilitate communication between author and reader through vocabulary control and, specifically, by selecting terms from the text to be used in compiling the index, by controlling synonyms to increase consistency, by defining ambiguous terms, and by providing cross-references to lead readers to the information sought.

These tasks are by no means exclusive to indexers of Canadian materials; au contraire, they constitute the basic principles used to compile an index in any language. However, when these principles are applied in index construction within a Canadian context, one must also take into consideration some of the linguistic characteristics of Canada’s two official languages: the divided usage between British and American English spelling and vocabulary; the literary warrant of words of Canadian origin outside Canada; the predominance of vocabulary, morphological, and semantic anglicisms in Canadian French; the current differences between standard French and Canadian French vocabulary; and the problems involved in bilingual indexing because of the trend in the English language towards nominalization.

British vs American English

Spelling

Since Americans broke with their mother country, the English language has evolved somewhat differently on both sides of the Atlantic as regards spelling, vocabulary, and pronunciation. Indexers, of course, are concerned primarily with spelling and vocabulary, and ordinarily conform to the form preferred by the author of the text, a standard which presents no problem for a British or an American publication. But what about Canadian English? If a text is written by an English-speaking Canadian, or translated from another language into English with a Canadian market in mind, the indexer’s task is complicated by a great deal of unsettled and divided usage, a situation brought about by some 200 years of influence from the United Kingdom on the one hand, and from the United States on the other. To appreciate the difficulties indexers face, we must understand the special relationship Canadian English enjoys with its two closest linguistic influences. It is invariably defined not by its own qualities, but by its deviation from what the Americans and the British deem the norm.

Canadian English has been influenced by the practice of both the Americans and the British, yet ‘in areas where American and British practices differ, Canadian usage is far from uniform’. It is not uncommon in Canada, for example, to find a text in which both British and American spellings are used. Traditionally, there has been a predilection for British forms in words such as *AXE, CATALOGUE,
CENTRE, CHEQUE, COLOUR, MEDIAEVAL, PLOW, SKILFUL and WOOLLEN (and words of similar pattern). In some cases, however, American spellings have asserted themselves to the virtual exclusion of the corresponding British forms, as in CONNECTION, CURB, JAIL, RECOGNIZE, TIRE and WAGON, for CONNEXION, KERB, GAOL, RECOGNISE, TYRE and WAGGON.  

There is, however, enough deviation from this generalization to warrant indexers using cross-references to lead the reader to the actual term used by the author in the text. For some words, such as TRAVELLER and ORGANISATION, a cross-reference is not indispensable because most Canadians are either aware of the alternative spelling (others are unaware that an alternative spelling exists, believing that only one is correct but never remembering which) or will happen upon the correct index entry while searching for it under the alternative spelling due to the alphabetical proximity of the two words.

To alleviate confusion over dual spellings, indexers can compile a list of the words to be indexed that have both a British and an American spelling and, using a dictionary, choose one form over the other to ensure uniformity and standardization. Cross-references must then be provided to lead the reader from the alternative to the preferred spelling. Another possibility is to look up the words in a Canadian dictionary, then choose the favoured spelling, which presupposes it is the one most widely used, and provide cross-references from the alternative form if there is any chance the reader might first look there.

Vocabulary

In vocabulary, as in spelling, Canadians are neither purely American or British, but in large measure a blend of both; and to this must be added many features that are typically Canadian. To an American, a Canadian sounds foreign because he prefers TAP, BRACES and PORRIDGE to FAUCET, SUSPENDERS and OATMEAL. The Englishman notices the American influence in the Canadian's choice of GAS, TRUCK, WINDSHIELD, MUFFLER, TRUNK and WRENCH over PETROL, LORRY, WINDSCREEN, SILENCER, BOOT and SPANNER. To the Canadian ear, SERVIETTE and NAPKIN both ring true, as do SHADES and BLINDS, and the choice of one term over the other is often an unconscious one, because both are widely used. Although SERVIETTE may be preferred in Britain and NAPKIN in the United States, in Canada user warrant is such that the two have become synonymous.

It is important to remember that these distinctions affect a relatively small number of words and that usage, for many, is divided in all three countries. Ultimately, user warrant will determine the indexer's choice. In instances where the author or translator has used a British or an American form unfamiliar to readers, indexers can direct readers to the information in the text by controlling the vocabulary so as to replace the foreign term with a familiar index term. When both British and American forms are widely used, indexers can choose one over the other and provide cross-references to the preferred term; if both forms are used interchangeably in the text, they can either index both and provide see also references or choose one and provide a see reference.

Words of Canadian origin

Recent years have seen the publication of dictionaries focusing on the way Canadian English differs from American and British English, and Canadian French from standard French; these include not only dictionaries of standard Canadian English and French but of regionalisms, slang, and colloquialisms. Importance is placed on words which originate in Canada, have evolved to acquire meanings in Canada different from their original meanings, or are spelled or pronounced unconventionally. Although many dictionary entries indicate the existence of multiple spellings, preference is usually given to the one most widely used in Canada. Some regionalisms, slang, and colloquialisms are found in dictionaries of standard Canadian English and French, while others are found only in specialized works such as the Dictionary of Canadianisms and Dictionnaire des petits ignorances de la langue française au Canada.

Words of Canadian origin can present a problem when a text is translated into standard English or French, especially if the word has no equivalent in the target language, or if another word is commonly used instead. According to Bertrand Russell, "no one can understand the word CHEESE unless he has a nonlinguistic acquaintance with cheese". If a text containing the word CHEESE is translated into a language for which there is no equivalent word, the translator has two options: provide a definition of the word in the target language or use the foreign word and footnote it, explaining the lack of an equivalent. If he defines CHEESE as "food made of the curd of milk, separated from the whey and pressed into a close mass", he may encounter a problem if the reader doesn't understand the meaning of CURD, for example, or if there is no equivalent in the target language for the word CURD. Cheese can be defined more largely as being simply a kind of food. Yet whenever an attempt is made to translate a term for which there is no equivalent in the target language, the indexer's work is further complicated because defining the concept invariably puts the emphasis on a more generic term, and this term is then indexed. If CHEESE is defined as a kind of food in a language in which there is no equivalent for the word CHEESE, then the only possible indexing term for this concept is FOOD. Because it is such a broad term, FOOD might not even be indexed. The resulting index, then, is not entirely true to the concepts expressed in the text.

It is conceivable that a given word of Canadian origin might not have an equivalent in another language, just as many of the words now embraced by French and English were at one time foreign to us. If there is a deficiency, terminology may be qualified or amplified by loanwords, neologisms, semantic shifts, or circumlocutions; the
Canadian word, or a variation of it, may be incorporated into the target language, or a relationship made between the word in question and a concept which is already familiar. Many of the words used by early explorers and settlers, for example, or for flora and fauna particular to Canada, were borrowed from the native people of North America, much as English has always borrowed words freely from the languages with which English-speaking people have come in contact.

Borrowed words that have since become part of our language present the same kinds of difficulties to indexers as the endless number of words that have crossed borders since the beginning of speech; this is not distinctly a Canadian problem, but one Canadian indexers share with their counterparts worldwide. Muskeg, from the Cree maskeek, has become part of our vocabulary and has also been adopted by French Canadians, who call it le muskeg, la fondrière de mousse, or la fondrière mousse. These three words are used interchangeably as the equivalent of muskeg, and any of them can be used as an index term, with cross-references made from the others. In France, however, one finds neither muskeg nor a word for muskeg, and the French would simply describe it as ‘un terrain marécageux et couvert de mousse’. If a text written in Canada translated from English into standard French includes the word muskeg, the indexer can either index muskeg, and might choose to so do if the translator has opted to keep the word in the text, or index a broader termMuskeg could be indexed under TERRAIN MARÉCAGEUX. In a Canadian book or journal destined for a British audience, either Muskeg or Ground, marshy could be indexed, with the cross-reference provided from one to the other.*

Anglicisms

Vocabulary and morphological Anglicisms

An indexer has the choice of using the author or translator’s natural language, or of substituting any part of the vocabulary he believes might be unfamiliar to the reader. In Canada, if a book or an article about cars is written in French or is translated into French, it is to be expected and even taken for granted that the author or translator will use standard French. In many subject areas, and the automobile industry is one in which this is the most noticeable, English and French words which are similar in appearance but have very different meanings. This is especially true in Canada, where a word shared by the two tongues is often used incorrectly in French because it is mistakenly thought to be synonymous with the English look-alike word. ACHIEVEMENT, which means END and not ACHIEVEMENT, is an example of a faux ami that might be translated incorrectly from English to French, or simply attributed with the English meaning. Similarly, the French noun les appointments means salary and not rendez-vous; une cave is a cellar, but not a cave; un éditeur publishes but does not necessarily edit, etc. Again the words like les freins, le capot, l’embrayage, le pare-brise, le pare-chocs, le ventilateur, le silencieux, le démarreur, la courroie, and le volant are sacrificed for les freins, le capot, la bouchon, la clé, le pare-craque, le pare-choc, le pare-soleil, the starter, the strappé, and le volant,* sometimes also called la roue. An indexer could, of course, simply compile an index from the standard French used in the text. But what good is an index if it does not allow readers to complete their search successfully? If users do not know the French word for steering wheel, they will be unable to locate the information on steering wheels in the text. If they look in the index under LE STEERING, borrowed from the English compound noun STEERING WHEEL, or under the French word LA ROUE, a literal translation of part of the English word STEERING WHEEL,* they will not find the information they seek unless there is a cross-reference from these two terms. The indexer must consider both literary warrant and user warrant when deciding which index term to use. In this case, Le volant has literary warrant; it is the standard French word for STEERING WHEEL. In Canada, however, Le steering and la roue have user warrant. The indexer might also consider compiling a classified rather than an alphabetical index in order to gather together like concepts under a heading such as VOITURES.

Morphological Anglicisms frequently occur because of a confusion between English and French words with the same etymological background which have preserved the same meaning and the same root, but have evolved with different but similar spellings. The overwhelming influence of English on Canadian French effects the use of Un pyromane rather than un pyromane, un voter instead of un votant, etc. The resulting deformation of a French word through inflexion of a suffix poses the same kind of problem to an indexer as the word STEERING WHEEL. If readers use the devise term VOTEUR rather than the word votant, how will they find the information on voters unless there is a reference directing them to votant?

Semantic Anglicisms

Like the translator, the indexer must beware of faux amis, English and French words which are similar in appearance but have very different meanings. This is especially true in Canada, where a word shared by the two tongues is often used incorrectly in French because it is mistakenly thought to be synonymous with the English look-alike word. ACHIEVEMENT, which means END and not ACHIEVEMENT, is an example of a faux ami that might be translated incorrectly from English to French, or simply attributed with the English meaning. Similarly, the French noun les appointments means salary and not rendez-vous; une cave is a cellar, but not a cave; un éditeur publishes but does not necessarily edit, etc. Again the

* Many standard dictionaries now include entries for words of Canadian origin and vocabulary particular to Canada. The Oxford Illustrated Dictionary, for example, defines muskeg as ‘level swamppy or boggy area in some regions of Canada’. Dictionaries can serve as a guide to the indexer by indicating the degree of literary warrant a Canadian word enjoys outside Canada.

* La roue is the French word for wheel, but is not used for steering wheel.
indexer is faced with a decision: indexing the natural language of the author or the translator might result in ambiguity because the concept expressed in the text is not obvious from the index term itself. The task of indexers is to choose terms for indexing that they would expect to find if they were looking for information. They can use a controlled vocabulary to bring together like concepts under one term, choosing an entry term that will not be ambiguous. Thus ACHÉVEMENT, ABOUTISSEMENT, and FIN might be grouped together under the term ABOUTISSEMENT, which will lead readers to the pages on which any of these three words is found.

**Demi-faux amis**, words that have the same meaning in French and English as well as another different meaning in either language, can also give rise to confusion. **Demi-faux amis** include words like L'ÉTIQUETTE, which means ÉTIQUE, but also PRICE TAG; L'EXPÉRIENCE, which means EXPERIENCE AS WELL AS EXPERIMENT; L'ADDITION, which can be A SUM OF NUMBERS AND THE BILL IN A RESTAURANT, etc. Again, indexers can avoid ambiguity by using vocabulary control, or by providing scope notes to indicate which meaning is intended in the text:

Addition (arithmétique)
Addition (notes des dépenses)

Scope notes are unnecessary when the correlation between the index term and the text is obvious. In a mathematics text book, ADDITION probably refers to a sum and not to the price of a meal. Likewise, in a book dealing with science, one can assume that EXPÉRIENCE alludes to an experiment, and not to an experience.

**Canadian vs standard French**

It is not unusual for the same French word to be used in France and Canada with entirely different meanings. DEPANNEUR, for example, the French word for a BREAK-DOWN MECHANIC or a TELEVISION REPAIRMAN, is used in Canada for CORNER STORE because it serves to DEPANNER, or TO BAIL OUT a person who cannot get to a grocery store. A KISS is UN BAISER in France and UN BEC in French Canada, whereas in France UN BEC is A BEAK. Incidentally, the French prefer the verb form EMBRASSER QUELQU'UN while in Canada one is more apt to hear DONNER UN BEC À QUELQU'UN.

Other ambiguities arise from dialectical varieties brought from France in the sixteenth and seventeenth centuries that have sometimes remained influential in Canada while losing ground in France. Words like CHAR, BREUVAGE, and TABAGIE are no longer used in standard French but are still part of a French Canadian's vocabulary, meaning, respectively, CAR, BEVERAGE, and TOBACCO SHOP. These inconsistencies are due partly to the influence of the English language, and partly to the geographical distance between France and Canada, which effects a divided evolution of the French language.

It is interesting that French Canadians, who readily embraced the English language for so long, have recently shown a reluctance to adopt new English words that have been accepted by the French, perhaps because they feel threatened by the anglicization of their language and their culture. The French have no qualms about SHOPPING,* while in Canada we insist on MAGASINER or FAIRE DU MAGASINAGE. This wariness to accept English words into the language, however, serves only to distance Canadian French from standard French and, for an indexer, increases the chances of ambiguity in vocabulary and, consequently, the need for vocabulary control.

The emergent terminology of the quickly growing information field is a prime example of the different ways neologisms make their way into the French and the French Canadian languages. Both groups borrow directly from English and also translate English terminology to create new words, but there seems to be some confusion due to the lack of standardization. HARD DISK translates as DISQUE RIGIDE and DISQUE DUR; TYPE FONT as POLICE CARACTÈRES in France and FAMILLE DE CARACTÈRES and FONT DE CARACTÈRES in Canada; DISKS are called both DISQUES and DISQUETTES in Canadian French, but usually only DISQUES in French; the usage among CD-ROM, COMPACT-DISQUE, and DISQUE COMPACT varies, but the French seem to prefer the abbreviated English form, while French Canadians tend to like DISQUE COMPACT. Because this is such a new domain, many of the buzz-words have not yet been established, and a good many of them have more than one correct form, all of which are used interchangeably. Neologisms create extra work for indexers, who must observe a changing vocabulary and provide the reader with relevant index terms.

**Syntactical problems**

Not exclusive to Canada, but prevalent because of the number of books and articles translated from French to English and vice versa, are the difficulties encountered when compiling an index for a work that has been translated, or when creating a bilingual index. One of the most troublesome areas in translation is caused by differences in language structure. In English there is a trend towards nominalization, while in French this is less true. It has often been stated, in fact, that any word can gain noun status in English, either by the addition of a suffix† or by conversion, the process whereby one part of speech is adapted

---

* The French also say, and just as frequently, faire des achats and faire les courses.
† A noun can be formed from a verb by adding a suffix such as -er or -or, potentially affixable to any verb in the English language (worker, writer, thriller); -ant (inhabitant, lubricant, disinfectant); -ee (payee, trainee, employee); -ation (fixation, exploration, foundation), etc. A de-adjectival noun is formed by adding a suffix such as -ness (happiness, selfishness, up-to-dateness); -ity (sanity, rapidity, elasticity), etc.
or converted to another without the addition of an affix.\footnote{In full conversion, one part of speech is derived from another (the nouns the desire, the attempt and the walk from the verbs to desire, to attempt and to walk); in partial conversion, one part of speech appears in a function typical of another (the nouns the wealthy, the well-dressed and the foolish from the adjectives wealthy, well-dressed and foolish). Although less common, conversion can also be made from closed-system words to nouns: we’re interested in the hows and the whys; and from phrases to nouns: during a recession, it’s the have-nots who suffer.} Given this flexibility, one might consider the English language an indexer’s dream come true, since nouns are postulated to be content indicators and therefore the natural preference of indexing languages.

Indexing involves doing a conceptual analysis to determine the contents of a book or an article, but indexing a translation might also mean using the existent index as a reference tool in creating a new one. While it is often relatively easy to translate common or concrete nouns, terms are not always equivalent in English and French because of the syntactic differences of the two languages, so a translator and an author might have selected different parts of speech to communicate the same concept. For these reasons a literal translation of the original index would be misleading, but it can be used as a guide to the content of the work. Doing a content analysis is especially important because it gives the indexer an overall idea of the concepts expressed in the work, and makes the task much easier than would relying solely on the original index or on the nouns in the translation to convey the author’s ideas to the reader.

The sentence Si vous acceptez cette mission, vous serez défrayé is translated as If you accept this assignment, your expenses will be covered. An indexer might pick out the word MISSION from the French sentence, and ASSIGNMENT and EXPENSES from the English sentence. Since DÉFRAYER QUELQU’UN means TO TAKE CARE OF SOMEONE’S EXPENSES, it would be redundant in French to repeat the noun EXPENSES, while in English, the noun is part of the idiom. In many cases, of course, more than one correct translation is possible. The English translation could just as easily be If you accept this assignment, you will be reimbursed, corresponding directly to the French Si vous acceptez cette mission, vous serez remboursé. Whether or not expenses are covered might be of immense importance to the reader, but the word DÉFRAYER would probably not be an index term, and reference to this paragraph might not appear in the index even though expenses are mentioned elsewhere in the text using another word, such as DÉPENSES or FRAIS. In this case, a noun was used in English to express the same concept relayed in French by a verb, although the French noun dépenses could be used as an index term. Even when a French equivalent for an English noun or noun phrase does exist, it will not necessarily be as widely used. English is a more supple language than French and, stylistically, lends itself well to this kind of construction, while a verb form is often more natural in French and is usually more correct stylistically. In this text, the indexer should note the page where DÉFRAYER occurs, using the noun DÉPENSES as a lead-in entry to the passage. Indexers regularly use this process of gathering together like terms and ideas when creating indexes for unilingual works. If the English language work has been indexed, it might prove helpful to the indexer because EXPENSES is probably already an index term.

**Conclusion**

Indexers try to facilitate communication between the author and the reader by performing a content analysis to select the relevant concepts in the text, then select the terms they think will lead the reader to the information needed. In Canada, indexing takes on an additional dimension because of the influence of powerful external linguistic forces, on French-speaking Canadians by France and English-speaking North America, and on English-speaking Canadians by Britain and the United States. Taking into consideration the potential problems of indexing in this milieu, this paper has examined some of the means by which indexers can surmount these difficulties so as to help readers find the information they are seeking.

**References**


**Department of great thoughts**

‘Where the title failed to convey the full contents of the article, the compiler has looked for suitable subjects to index from the general contents: thus Richard Phillips’ 1968 article “The last of the drovers” has been indexed not only under “drovers” but also under “Isaac, Dafydd”, the subject of the article. . . . Miss [. . .] has been responsible for typing most of the work, and her careful work and her wide knowledge of many of the subjects contained in the Index has enabled me to avoid many errors.’