NAME OF AUTHOR/NOM DE L'AUTEUR: Murray T. Wilton

TITLE OF THESIS/TITRE DE LA THÈSE: Bilingual Lexicography: Theoretical Foundations and Practical Methodology With Special Reference to French and English

UNIVERSITY/UNIVERSITÉ: Simon Fraser University

DEGREE FOR WHICH THESIS WAS PRESENTED/GRÂDE POUR LEQUEL CETTE THÈSE FUT PRÉSENTÉE: Ph. D.

YEAR THIS DEGREE CONFERRED/ANNÉE D'OBTENTION DE CE GRÂDE: 1978

NAME OF SUPERVISOR/NOM DU DIRECTEUR DE THÈSE: Dr. Pierre Guiraud

Permission is hereby granted to the NATIONAL LIBRARY OF CANADA to microfilm this thesis and to lend or sell copies of the film.

The author reserves other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without the author's written permission.

DATED/DATÉ: 31 August 1978 SIGNED/SIGNÉ:

PERMANENT ADDRESS/RÉSIDENCE FINALE: 767 VICTORIA AVE

VICTORIA, B.C. V8S 4N2
NOTICE

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

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

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

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

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

THIS DISSERTATION HAS BEEN MICROFILMED EXACTLY AS RECEIVED

Bibliotheque nationale du Canada
Direction du catalogage
Division des theses canadiennes

AVIS

La qualite de cette microfiche depend grandement de la qualite de la these soumise au microfilmage. Nous avons tout fait pour assurer une qualite superieure de reproduction.

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

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

Les documents qui font deja l'objet d'un droit d'auteur (articles de revue, examens publies, etc.) ne sont pas microfilmes.

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

LA THÈSE A ÉTÉ MICROFILMÉE TELLE QUE NOUS L'AVONS REÇUE
BILINGUAL LEXICOGRAPHY:
THEORETICAL FOUNDATIONS AND PRACTICAL METHODOLOGY, WITH
SPECIAL REFERENCE TO CANADIAN FRENCH AND ENGLISH

by

Murray Thomas Wilton
B.A., University of New Zealand, 1961
M.A., University of Auckland, 1964

A THESIS SUBMITTED IN PARTIAL FULLFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
in the Department
of
Languages, Literatures and Linguistics

© Murray Thomas Wilton 1978

SIMON FRASER UNIVERSITY
August 1978

All rights reserved. This thesis may not be reproduced in whole or in part, by photocopy or other means, without permission of the author.
Name: Murray Thomas Wilton

Degree: Doctor of Philosophy

Title of Dissertation: Bilingual Lexicography: Theoretical Foundations and Practical Methodology, with Special Reference to Canadian French and English.

Examining Committee:

Chairman: John Knowles

Pierre Guiraud (Senior Supervisor)

Charles P. Bouton

Barrie E. Bartlett

Marguerite Saint-Jacques

Jean-Paul Vinay Matthew H. Scargill

External Examiners Professors
University of Victoria, Victoria, B.C.

Date Approved: August 4, 1978
I hereby grant to Simon Fraser University the right to lend my thesis or dissertation (the title of which is shown below) to users of the Simon Fraser University Library, and to make partial or single copies only for such users or in response to a request from the library of any other university, or other educational institution, on its own behalf or for one of its users. I further agree that permission for multiple copying of this thesis for scholarly purposes may be granted by me or the Dean of Graduate Studies. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Title of Thesis/Dissertation:
Bilingual Lexicography: Theoretical Foundations and Practical Methodology With Special References to French and English

Author: ____________________________

(signature)

Murray T. Wilton

(name)

August 4, 1978

(date)
BILINGUAL LEXICOGRAPHY:
THEORETICAL FOUNDATIONS AND PRACTICAL METHODOLOGY, WITH SPECIAL REFERENCE TO CANADIAN FRENCH AND ENGLISH

Abstract

Lexicography in the Indo-European languages has been practised for centuries. In fact, by the time Samuel Johnson published his Dictionary of the English Language in 1755 lexicography had already inherited a methodology developed by the earliest dictionary-makers. The first dictionaries were bilingual and evolved as a reaction to social and practical phenomena -- the need to comprehend texts written largely in Latin and, subsequently, the demand for communication between different language groups. These kinds of considerations have continued to exert an influence on lexicography, to such an extent that modern bilingual dictionaries still draw much of their inspiration from traditional methodology instead of attempting to appeal to scientific procedure predicated on theoretical linguistic research.

During the past two decades a significant amount of theoretical research has been done in monolingual lexicography in an attempt to bridge the gap between theory and practice in that discipline. But bilingual lexicography remains a vaguely delineated subject on the fringe of linguistics: it is regarded by some as a branch of translation (itself an ill-defined field); others consider it an offshoot of monolingual lexicography, even though the reverse is true.

In this thesis an attempt is made to demonstrate that bilingual lexicography is an important and independent discipline within linguistics, with its own theoretical basis set in a general theory of linguistics and a distinctive methodology which extends into technological applications.
The first part of the thesis is devoted to a descriptive and theoretical analysis of the field and associated disciplines. A diachronic view of the developmental factors which have influenced French and English lexicography is followed by a typological classification of bilingual dictionaries according to criteria which will form the basis for the discussion on linguistic methodology in Part II.

Various aspects of modern linguistic theory are surveyed in order to determine theoretical approaches which are of significance for bilingual lexicography. It will be determined that, in spite of obvious relationships, translation and bilingual lexicography constitute two independent studies with different goals and separate theoretical issues. An examination of semantic theory as it concerns lexicography will be used as the basis for a model of lexical structure in bilingual dictionaries, appealing to methods of distributional and componential analysis and suggesting the manner in which these can be fruitfully applied to solving the problem of lexical equivalence. Various issues in transformational generative grammar are analysed in order to shed light on a number of contentious issues in lexicography and the findings of psycholinguistic research into bilingualism are shown to have significant ramifications for bilingual lexicographers.

The goal of the first part of the thesis is therefore to demonstrate that bilingual lexicography has its own unique theoretical foundation: the notion that the equivalence relation, the very source of its endeavour, is an abstract entity and that the dictionary represents a model of "lexical competence" where the units of manipulation are artificially removed from contextual settings and constitute elements of "potential equivalence". It is the aim of the second part of the study to indicate how, in practical terms, the apparently unattainable ideal equivalence is at least approximated through the application of a special kind of methodology and the observance of a number of basic
prerequisites. A detailed analysis of structural features is undertaken in order to reveal the special conditions which apply to dictionaries compiled for non-native users. Then specific problems of a differential nature are examined and model solutions proposed. An important chapter on regional variation is designed to show how this particular feature is dealt with in the Canadian context, although much of what is said would be applicable to other similar linguistic situations. Part II ends with a description of the application of computer technology to the editing of a bilingual dictionary; not only does automatic processing facilitate the accomplishment of repetitive and demanding tasks, but also it adds a measure of control over otherwise elusive materials and enhances accuracy and consistency.
For my wife and children who endured so much in order that this work could be undertaken and brought to fruition.
ACKNOWLEDGEMENTS

This research was largely inspired by contacts with teachers and lexicographers over a number of years. Like any undertaking, it owes much to a number of individuals. In particular I would like to express my gratitude to the members of my Supervisory Committee for their constant guidance and encouragement and especially for their patient understanding during periods of relative stagnation.

I would also like to thank the University of Victoria and its Canadian Bilingual Dictionary Project whose research materials were freely accessible to me and who allowed me generous use of their computing facilities.

Finally, I wish to recognize the financial assistance of McClelland & Stewart Ltd. of Toronto in the typing and duplication of the manuscript.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>APPROVAL</th>
<th>(ii)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>(iii)</td>
</tr>
<tr>
<td>DEDICATION</td>
<td>(vi)</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>(vii)</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>(xvi)</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>(xviii)</td>
</tr>
<tr>
<td>ABBREVIATIONS USED IN THE TEXT</td>
<td>(xix)</td>
</tr>
</tbody>
</table>

## PART I

DESRIPTIVE AND THEORETICAL ANALYSIS

Chapter

1. INTRODUCTION ........................................ 1
   1.0 The Nature of the Problem .................... 1
   2.0 Terminology .................................... 9

2. HISTORICAL DEVELOPMENT .......................... 30
   1.1 Arrangement of Entries ....................... 31
   2.0 Canonical Form ................................ 34
   3.0 Grammar Identification ...................... 39
   4.0 Pronunciation ................................ 42
   5.0 Ordering of Senses ............................ 48
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.0 Meaning Discrimination</td>
<td>51</td>
</tr>
<tr>
<td>7.0 Contextual Examples</td>
<td>55</td>
</tr>
<tr>
<td>8.0 Cross-references</td>
<td>56</td>
</tr>
<tr>
<td>9.0 Prefatorial Information</td>
<td>58</td>
</tr>
<tr>
<td>10.0 Conclusion</td>
<td>62</td>
</tr>
<tr>
<td>Notes</td>
<td>65</td>
</tr>
<tr>
<td>References</td>
<td>68</td>
</tr>
</tbody>
</table>

3. DESCRIPTIVE ANALYSIS | 69
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0 Typological Classification</td>
<td>69</td>
</tr>
<tr>
<td>2.0 Aims and Objectives</td>
<td>71</td>
</tr>
<tr>
<td>3.0 Choice of Languages</td>
<td>77</td>
</tr>
<tr>
<td>4.0 Directional criteria</td>
<td>80</td>
</tr>
<tr>
<td>5.0 Relational Criteria</td>
<td>85</td>
</tr>
<tr>
<td>6.0 Classification by Range</td>
<td>87</td>
</tr>
<tr>
<td>7.0 Lexical Concentration</td>
<td>91</td>
</tr>
<tr>
<td>8.0 Structural Criteria</td>
<td>95</td>
</tr>
<tr>
<td>9.0 Discriminatory Features</td>
<td>102</td>
</tr>
<tr>
<td>10.0 Conclusion</td>
<td>104</td>
</tr>
<tr>
<td>Notes</td>
<td>108</td>
</tr>
<tr>
<td>References</td>
<td>112</td>
</tr>
</tbody>
</table>

4. TRANSLATION THEORY AND BILINGUAL LEXICOGRAPHY | 114
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Notes</td>
<td>136</td>
</tr>
<tr>
<td>References</td>
<td>139</td>
</tr>
<tr>
<td>Chapter</td>
<td>Page</td>
</tr>
<tr>
<td>------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>5. SEMANTIC ISOMORPHISM</td>
<td>140</td>
</tr>
<tr>
<td>Notes</td>
<td>165</td>
</tr>
<tr>
<td>References</td>
<td>166</td>
</tr>
<tr>
<td>6. LEXICAL STRUCTURE AND LEXICOGRAPHY</td>
<td>167</td>
</tr>
<tr>
<td>1.0 Field Theory</td>
<td>167</td>
</tr>
<tr>
<td>2.0 Componential Analysis</td>
<td>171</td>
</tr>
<tr>
<td>3.0 A Comparative Example</td>
<td>177</td>
</tr>
<tr>
<td>4.0 Analysis of Closed Fields</td>
<td>184</td>
</tr>
<tr>
<td>Notes</td>
<td>195</td>
</tr>
<tr>
<td>References</td>
<td>197</td>
</tr>
<tr>
<td>7. TRANSFORMATIONAL GENERATIVE GRAMMAR AND BILINGUAL LEXICOGRAPHY</td>
<td>198</td>
</tr>
<tr>
<td>1.0 Preliminary Notions</td>
<td>198</td>
</tr>
<tr>
<td>2.0 Dictionaries as Models of Verbal Behaviour</td>
<td>202</td>
</tr>
<tr>
<td>3.0 The Revised Semantic Component</td>
<td>205</td>
</tr>
<tr>
<td>4.0 Precision in Lexicographical Description</td>
<td>207</td>
</tr>
<tr>
<td>5.0 Transformational Solutions for Lexicographical Problems: Verbs</td>
<td>210</td>
</tr>
<tr>
<td>6.0 Transformational Solutions for Lexicographical Problems: Adjectives</td>
<td>217</td>
</tr>
<tr>
<td>7.0 The Bilingual Dictionary as a Model of Lexical Competence</td>
<td>227</td>
</tr>
</tbody>
</table>
Chapter 8.0 Conclusion .................. 243
                      Notes .................. 245
                      References .................. 248

8. BILINGUALISM AND BILINGUAL LEXICOGRAPHY .................. 250
   1.0 The Problem of Definition .................. 250
   2.0 First Language Acquisition .................. 254
   3.0 Second Language Acquisition .................. 261
   4.0 Psycholinguistic Description of Bilingualism .................. 261
   5.0 Bilingualism in Lexicography .................. 266
                      Notes .................. 274
                      References .................. 276

9. CONCLUSIONS TO PART I .................. 278

PART II

LINGUISTIC METHODOLOGY

10. THE CORPUS .................. 282
                      Notes .................. 298

11. STRUCTURAL METHODOLOGY .................. 300
   1.0 Preliminary Considerations .................. 300
   2.0 Word Order .................. 313
   3.0 The Lexicographical "Word" .................. 318
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.0</td>
<td>Internal Article Structure</td>
<td>332</td>
</tr>
<tr>
<td>5.0</td>
<td>Sense Order</td>
<td>352</td>
</tr>
<tr>
<td></td>
<td>Notes</td>
<td>363</td>
</tr>
<tr>
<td></td>
<td>References</td>
<td>367</td>
</tr>
<tr>
<td>12.</td>
<td>DIFFERENTIAL LEXICOGRAPHY (A) GRAMMAR</td>
<td>369</td>
</tr>
<tr>
<td>1.0</td>
<td>Grammar and Lexicon</td>
<td>369</td>
</tr>
<tr>
<td>2.0</td>
<td>Grammatical Description</td>
<td>373</td>
</tr>
<tr>
<td>3.0</td>
<td>Gender and Number Problems</td>
<td>399</td>
</tr>
<tr>
<td>4.0</td>
<td>Syntactic Problems</td>
<td>414</td>
</tr>
<tr>
<td></td>
<td>Notes</td>
<td>434</td>
</tr>
<tr>
<td></td>
<td>References</td>
<td>437</td>
</tr>
<tr>
<td>13.</td>
<td>DIFFERENTIAL LEXICOGRAPHY (B) SEMANTICS</td>
<td>439</td>
</tr>
<tr>
<td>1.0</td>
<td>Semantic Preliminaries</td>
<td>439</td>
</tr>
<tr>
<td>2.0</td>
<td>The Formal Properties of the</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Equivalence Relation</td>
<td>442</td>
</tr>
<tr>
<td>3.0</td>
<td>Differential Equivalence Problems</td>
<td>457</td>
</tr>
<tr>
<td>4.0</td>
<td>Idiomatic and Figurative</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Equivalence</td>
<td>474</td>
</tr>
<tr>
<td>5.0</td>
<td>Language Level Equivalence</td>
<td>481</td>
</tr>
<tr>
<td>6.0</td>
<td>Practical Equivalence Problems</td>
<td>491</td>
</tr>
<tr>
<td>7.0</td>
<td>Conclusion</td>
<td>505</td>
</tr>
<tr>
<td></td>
<td>Notes</td>
<td>507</td>
</tr>
<tr>
<td></td>
<td>References</td>
<td>510</td>
</tr>
</tbody>
</table>
# Chapter 14: Lexicographical Presentation of "Faux Amis"

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1.0 A Deficiency in Bilingual Lexicography</strong></td>
<td>511</td>
</tr>
<tr>
<td><strong>2.0 Origins and Significance of Deceptive Cognates</strong></td>
<td>514</td>
</tr>
<tr>
<td><strong>3.0 Classification of Deceptive Cognates</strong></td>
<td>519</td>
</tr>
<tr>
<td><strong>4.0 Lexicographical Presentation</strong></td>
<td>538</td>
</tr>
<tr>
<td><strong>5.0 Conclusion</strong></td>
<td>546</td>
</tr>
<tr>
<td>Notes</td>
<td>547</td>
</tr>
<tr>
<td>References</td>
<td>548</td>
</tr>
</tbody>
</table>

# Chapter 15: The Treatment of Regionalisms in Bilingual Lexicography

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1.0 Regional Variation in Bilingual Lexicography</strong></td>
<td>549</td>
</tr>
<tr>
<td><strong>2.0 The Problem of Definition</strong></td>
<td>552</td>
</tr>
<tr>
<td><strong>3.0 Regionalisms in the Dictionary</strong></td>
<td>601</td>
</tr>
<tr>
<td><strong>4.0 Conclusion</strong></td>
<td>615</td>
</tr>
<tr>
<td>Notes</td>
<td>617</td>
</tr>
<tr>
<td>References</td>
<td>619</td>
</tr>
</tbody>
</table>

# Chapter 16: The Automation of Lexicography

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1.0 Introductory</strong></td>
<td>621</td>
</tr>
<tr>
<td><strong>2.0 Input Hardware</strong></td>
<td>625</td>
</tr>
<tr>
<td><strong>3.0 The Structure of Lexicographical</strong></td>
<td></td>
</tr>
<tr>
<td>Chapter</td>
<td>Page</td>
</tr>
<tr>
<td>---------</td>
<td>------</td>
</tr>
<tr>
<td>Material</td>
<td>629</td>
</tr>
<tr>
<td>4.0 File Structure Analysis</td>
<td>640</td>
</tr>
<tr>
<td>5.0 Computer Data File Structure</td>
<td>646</td>
</tr>
<tr>
<td>6.0 Automatic Generation of Phonetic Transcription</td>
<td>654</td>
</tr>
<tr>
<td>7.0 Data Retrieval</td>
<td>662</td>
</tr>
<tr>
<td>8.0 Automatic Inversion of Entries</td>
<td>675</td>
</tr>
<tr>
<td>9.0 Conclusion</td>
<td>696</td>
</tr>
<tr>
<td>Notes</td>
<td>700</td>
</tr>
<tr>
<td>References</td>
<td>702</td>
</tr>
</tbody>
</table>

17. GENERAL CONCLUSIONS | 704  |

APPENDIX

I CHRONOLOGICAL BIBLIOGRAPHY OF BILINGUAL FRENCH AND ENGLISH DICTIONARIES | 710 |

II SUMMARY OF FRENCH EXPRESSIONS EMPLOYING THE WORD "COUP", WITH ENGLISH EQUIVALENTS | 720 |

III FIELD LABEL ABBREVIATIONS USED IN CD2 | 727 |

IV SAMPLE OF DOCUMENTATION USED IN CD2 | 731 |

V SAMPLES OF PHONETIC DESCRIPTION IN CD2 | 740 |

VI INVENTORY OF CANADIANISMS RELATING TO CLIMATE | 741 |

BIBLIOGRAPHY | 751 |
<table>
<thead>
<tr>
<th></th>
<th>BIBLIOGRAPHICAL SOURCES</th>
<th>GENERAL LEXICOGRAPHY</th>
<th>BILINGUAL LEXICOGRAPHY</th>
<th>LEXICOLOGY AND LEXICOGRAPHY</th>
<th>GRAMMAR IN LEXICOGRAPHY</th>
<th>PRONUNCIATION IN LEXICOGRAPHY</th>
<th>REGISTER IN LEXICOGRAPHY</th>
<th>REGIONAL VARIATION</th>
<th>HISTORICAL LEXICOGRAPHY</th>
<th>COMPUTATIONAL LEXICOGRAPHY</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>752</td>
<td>753</td>
<td>769</td>
<td>774</td>
<td>782</td>
<td>785</td>
<td>786</td>
<td>788</td>
<td>792</td>
<td>794</td>
</tr>
<tr>
<td></td>
<td>page</td>
<td>page</td>
<td>page</td>
<td>page</td>
<td>page</td>
<td>page</td>
<td>page</td>
<td>page</td>
<td>page</td>
<td>page</td>
</tr>
<tr>
<td>TABLE</td>
<td>FEATURE MATRIX FOR BILINGUAL LEXICOGRAPHY</td>
<td>105</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>-------------------------------------------</td>
<td>-----</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SCHEMATIZATION OF THE DEVELOPMENT OF A BILINGUAL DICTIONARY</td>
<td>107</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>BILINGUAL SCALE OF TEMPERATURE: AN EXAMPLE OF A CLOSED FIELD</td>
<td>187</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>FEATURE MATRIX FOR THE BILINGUAL SCALE OF TEMPERATURE</td>
<td>189</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>FEATURE MATRIX FOR &quot;BREAK&quot;</td>
<td>241</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CATEGORIZATION OF BILINGUAL DICTIONARY USERS</td>
<td>302</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>AN EXAMPLE OF ABSOLUTE ALPHABETICAL ORDER</td>
<td>315</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>AN EXAMPLE OF TOPICAL ORDERING</td>
<td>316</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PHONETIC TRANSCRIPTION IN SELECTED DICTIONARIES</td>
<td>340</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>SUBCATEGORIZATION OF WORD CLASSES</td>
<td>378</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>NOMINAL CATEGORY LABELS IN SELECTED DICTIONARIES</td>
<td>385</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PRONOUN SUBCLASSES IN LEXICOGRAPHY</td>
<td>394</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>FRENCH GENDER DESINENCES</td>
<td>410</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CORRESPONDENCE BETWEEN FILE STRUCTURES</td>
<td>635</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CODING SYSTEM FOR BILINGUAL DATA FILE</td>
<td>650</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CORRESPONDENCE BETWEEN TYPOGRAPHICAL ENTRY AND COMPUTER DATA FILE</td>
<td>652</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>FLOWCHART FOR AUTOMATIC TRANSCRIPTION CYCLES</td>
<td>659</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TABLE</td>
<td>page</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>XVIII SCHEMATIC VIEW OF DATA FILE CONTAINING 49 ELEMENTS OF ACCESSIBLE INFORMATION</td>
<td>664</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>XIX PROCEDURE FOR RETRIEVAL OF DESIGNATED FIELD LABELS</td>
<td>673</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>XX FIRST INVERSION CYCLE</td>
<td>689</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>XXI SECOND INVERSION CYCLE</td>
<td>690</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>XXII THIRD INVERSION CYCLE</td>
<td>691</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FIGURE</td>
<td>PRIMARY SUBDIVISIONS OF GENERAL DICTIONARIES</td>
<td>PAGE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>---------------------------------------------</td>
<td>------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.</td>
<td></td>
<td>80</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.</td>
<td></td>
<td>83</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.</td>
<td>AN EXAMPLE OF DIFFERENTIAL EQUIVALENCE:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>&quot;lukewarm&quot;</td>
<td>191</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.</td>
<td>FORMAL EQUIVALENCE OF FRENCH AND ENGLISH</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ARTICLES</td>
<td>396</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.</td>
<td>A METHOD OF DEALING WITH MEANING DISCRIMINATION</td>
<td>448</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.</td>
<td>AN EXAMPLE OF DIFFERENTIAL EXTENSION:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>&quot;chou&quot;</td>
<td>459</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7.</td>
<td>AN EXAMPLE OF DIFFERENTIAL EXTENSION:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>&quot;bill&quot;</td>
<td>460</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8.</td>
<td>RAMIFICATIONS OF FRENCH &quot;compte&quot;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9.</td>
<td>LANGUAGE LEVELS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10.</td>
<td>PARTIALLY OVERLAPPING SEMANTIC FIELDS</td>
<td>517</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11.</td>
<td>THE FRENCH LANGUAGE IN CANADA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12.</td>
<td>THE ENGLISH LANGUAGE IN CANADA</td>
<td>604</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13.</td>
<td>FRENCH AND ENGLISH IN NORTH AMERICA</td>
<td>606</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14.</td>
<td>TOPICAL ORDERING OF ENTRIES</td>
<td>633</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AE</td>
<td>American English</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BE</td>
<td>British English</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BELL</td>
<td>Bellows' French and English Dictionary</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BORDAS</td>
<td>Dictionnaire Bordas</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CASS1</td>
<td>Cassell's Compact French and English Dictionary</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CASS2</td>
<td>Cassell's French-English English-French Dictionary</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CD1</td>
<td>The Canadian Dictionary/Dictionnaire canadien (1962)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CD2</td>
<td>Revised edition of CD1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CE</td>
<td>Canadian English</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CF</td>
<td>Canadian French</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CROW</td>
<td>Crowell's Spanish-English Dictionary</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DCE</td>
<td>Dictionary of Canadian English (Senior edition)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DCHP</td>
<td>Dictionary of Canadianisms on Historical Principles</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DPL</td>
<td>Dictionnaire pratique Larousse</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EF</td>
<td>European French (France)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FEW</td>
<td>Französisches Etymologisches Wörterbuch</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GASC</td>
<td>Gasc's French Dictionary (Student's edition)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GPFC</td>
<td>Glossaire du parler français au Canada</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HAP1</td>
<td>Harrap's Concise French and English Dictionary</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HAP2</td>
<td>Harrap's Shorter French and English Dictionary</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HAP3</td>
<td>Harrap's Standard French and English Dictionary</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HNS</td>
<td>Harrap's New Standard French-English Dictionary</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IF</td>
<td>International French</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Code</td>
<td>Dictionary Name</td>
<td>Description</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------</td>
<td>----------------</td>
<td>-------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LAR1</td>
<td>Larousse de poche</td>
<td>(french and English)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LAR2</td>
<td>Dictionnaire Larousse moderne</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>LITTRE</td>
<td>Dictionnaire de la langue française</td>
<td>(Littré)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OED</td>
<td>Oxford English Dictionary</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OXF</td>
<td>Oxford Concise French and English Dictionary</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PROB</td>
<td>Petit Robert--Dictionnaire de la langue française</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RHD</td>
<td>Random House Dictionary of the English Language</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SF</td>
<td>Standard French</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SL</td>
<td>Source Language</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TL</td>
<td>Target language</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WIII</td>
<td>Webster's Third New International Dictionary of the English Language</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WND</td>
<td>Webster's New Collegiate Dictionary</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
PART I
DESCRIPTIVE AND THEORETICAL ANALYSIS

CHAPTER 1
INTRODUCTION

1.0 The Nature of the Problem

Dictionary-making, in its various manifestations, has a venerable lineage. In the eastern civilizations there are indications of lexicographical activity as early as three thousand years ago, while in western countries modern lexicographers enjoy the benefits of a tradition which spans more than four centuries.

One of the curiosities of lexicography in French and English is the fact that its inception came as a response to the need to comprehend texts written in a non-native language, namely, Latin. In other words, the first dictionaries in these two languages were bilingual in the sense that they defined, in the vernacular, unusual or difficult terminology in Latin. It simply did not occur to pre-medieval scholars, or so it would appear, that anyone might require definitions or explanations of the native lexicon, and the reason for this state of affairs is quite elementary: any texts written in the vernacular up to the sixteenth century consisted only of words in common oral use; concepts for which the native language

1
had no terminology were expressed in Latin. Only when authors began to invent new terminology in the native language did it become necessary to produce some kind of explanatory glossary for the reading public. Thus, for example, Shakespeare had no access to an English dictionary; if he had needed one, he would have been obliged to compile his own.

1.1 Even if we take into consideration only those publications which conform to modern notions about dictionaries, lexicography in French and English has been practised for at least four hundred years. One might reasonably expect, therefore, that its exponents would have by now established a theoretical basis for the practical methodology which finds expression in published dictionaries. Yet this is by no means true. As U. Weinreich (1966:395) puts it, "Lexicography carried on in paradisiac innocence without questioning its own theoretical foundations," whereas in all other areas of linguistics considerable advances have been made during the past century. It is, in fact, only within the last two decades that any serious attempt has been made to subject lexicographical materials to theoretical linguistic analysis.

1.1.1 There are various reasons for this state of affairs. To begin with, dictionaries have traditionally been
compiled by men of letters whose scholarship and literary prowess qualified them for the arduous and demanding task of assembling and describing the entire lexicon of their language. Their output is usually considered, and often is, a work of art; this is why the controversial question of whether lexicography can be considered a science or an art may be reduced to a simple statement to the effect that it may be both, depending on the aims and skills of the lexicographer. The modern trend is for dictionaries to be prepared by trained linguists whose concern to apply linguistic research to lexicography brings a measure of scientific precision to the dictionary-making process.

1.1.2 Secondly, dictionaries evolved as the direct result of practical necessity and social pressures and such pragmatic considerations have continued to exert an influence on lexicographical output. To such an extent, indeed, that modern dictionaries still tend to draw their inspiration from traditional methodology instead of attempting to appeal to scientific tactics predicated on theoretical linguistic research. As a consequence, the only lexicographical discussion normally available is the preface material of the dictionary itself and this may vary from a page or two, acknowledging the collaboration of contemporaries or the debt owed to predecessors, to a
voluminous explanation of the methodology employed in the work. Scientific analysis is totally absent because the aim of lexicographers is generally to inform their reading public about the content of the dictionary and to instruct them on its most efficaceous use.

1.2 In an attempt to bridge the gap between theory and practice in monolingual lexicography, a few scholars have comparatively recently undertaken research which is described in articles in learned journals and a handful of book-length studies. The need to establish a theoretical framework for the practice of dictionary-making has also motivated groups of concerned linguists and lexicographers to participate in conferences devoted exclusively to practical and theoretical issues in lexicography and related fields. Yet bilingual lexicography usually plays a marginal part in these gatherings and in the scholarly output which accompanies them; this is because it is considered by some as an offshoot of monolingual lexicography (even though the reverse is true) and by others as a branch of translation (itself an ill-defined discipline within linguistics). The paucity of research in bilingual lexicography is such that only one book-length study has so far been undertaken, and even this one is more a summary of what contemporary linguists and a few lexicographers have to say about
dictionary-making (mainly monolingual works, in spite of the title) than a serious attempt to come to grips with the theoretical foundations.

1.2.1 Cogent arguments can be brought forward to defend the notion that bilingual lexicography is merely a subdivision of general lexicography. It does, after all, define the words of one language in terms of the words of another, including information on orthography, pronunciation, grammar, register, dialect and so on, all of which are elements of traditional monolingual lexicography. However, the resemblance is in fact quite superficial. Monolingual dictionaries, except in special circumstances, are designed for consumption by native speakers and therefore editors make certain assumptions about innate language competence. Bilingual lexicographers, on the other hand, are attempting to introduce their readers to a second language, either through translation or composition, and must be constantly aware of the fact that non-native speakers will require certain kinds of information which are generally taken for granted and not furnished in a definition dictionary. Since the methodology is so different, we may logically expect its theoretical basis to be correspondingly independent.

1.2.2 Similarly, the notion that bilingual lexicography is
a dependent category of translation is an excessively simplistic view. There are two kinds of translation--oral interpretation and written translation--both of which operate at the level of continuous text and assume various kinds of correspondence between the discrete lexical units of the original and its translated version. Bilingual dictionaries are concerned almost exclusively with independent lexical units as corresponding pairs in two languages. The correspondences are obtained on the basis of an objective analysis of distributional phenomena in each language and are an abstraction of contextual potentiality. Translation tends to be a creative and artistic activity which is initiated with the study and interpretation of the intent of the original author (or speaker), including stylistic nuances, and concludes with the production of a text (or utterance) which captures the total message and flavour of the original text. It may therefore be more accurate to claim that translators are to a large extent dependent on the precision which lexicographers bring to bear in establishing the lexical correspondences of their dictionaries. Once again, since the practical activities differ in such explicit ways, the theory upon which each is based should be correspondingly distinctive.

1.3 A popular controversy surrounding all translation
activities is the contention that translation is a theoretical impossibility and those who hold the opposite view have generally been unable to provide convincing evidence to support their position. In the succeeding chapters of Part I we will endeavour to substantiate the hypothesis that lexical correspondence between two languages, as manifested in bilingual dictionaries, is not only physically possible but also theoretically defensible. We will therefore attempt to make a case for the acceptance of bilingual lexicography as an important and independent discipline within linguistics, possessing its own unique theoretical framework set in a general theory of language, with a distinctive methodology ranging into technological applications.

1.3.1 Even within bilingual lexicography, considered as a self-contained study, there will be subdivisions relating to questions such as the languages which are being equated, the eventual users of the dictionary, the specific aims of the editors, and so on. The scope of this study is therefore necessarily restricted to a consideration of French and English lexicography, although much of what is presented is applicable to any language pairs. Similarly, because the source materials are basically those of a Canadian research project, the range of the investigation is further specified in terms
of Canadian English and Canadian French. As we shall observe in Chapter 6 of Part II, the question of dialect in bilingual lexicography is a crucial one which engenders its own particular methodology.

1.3.2 In proposing a theoretical basis on which to found the practice of bilingual lexicography it was considered essential to constitute a frame of reference which would enable the boundaries to be clearly delineated. Thus, for example, in the second part of this introduction we undertake a fairly detailed study of the terminology "lexicography" and "lexicology," since there is confusion in some quarters as to their precise definition. The inclusion of descriptive and historical accounts in this section is justified on the grounds that no accurate theoretical analysis can be adequately carried out in the absence of a complete appraisal of the physical materials which are subject to its influence. For this reason, Chapter 2 begins with a discussion of historical developmental factors which have contributed to present-day methodology, while Chapter 3 analyses descriptive and typological features which reflect the aims and scope of contemporary bilingual lexicography.

1.3.3 The succeeding chapters of Part I deal with the more properly theoretical aspects of bilingual lexicography. In Chapter 4 a detailed examination of translation theory
demonstrates the manner in which it resembles and differs from lexicography. Chapter 5 draws attention to the concept of "semantic isomorphism" and is an attempt to explain how meaning transfer is possible between languages because there is a sufficient degree of correspondence in semantic structures. Modern studies in field theory, componential and distributional analysis form the nucleus for a discussion of total lexical structure and its significance for bilingual lexicography in Chapter 6. The findings of transformational generative grammar with respect to lexical phenomena are examined in Chapter 7 and shown to have important consequences for bilingual dictionary-makers, and finally, in Chapter 8 we attempt to draw a lesson from psycholinguistic research on bilingualism and its pedagogical ramifications. In Part II there is a detailed analysis of the linguistic methodology which must follow as a consequence of the theoretical conclusions reached in Part I. Practical considerations include methods of data collection, the characteristics of formal lexicographical structure, differential problems in grammatical and semantic equivalence, regional variation and technological applications.

2.0 Terminology: "lexicography" or "lexicology"

If lexicography, as generally defined and accepted
by linguists, is the art (or science) of making
dictionaries, then the theory of which it is the
practical manifestation will require a different
terminology.

2.1 The term which suggests itself to denote the
theoretical considerations of lexical phenomena is
lexicology; but the word has proved far from satisfactory
because of inconsistent interpretations placed on it by
several generations of linguists. Saussure emphasized
that lexicology—the "science des mots" as understood by
both his contemporaries and predecessors—is an
inseparable element of "grammar" and not an external
adjunct of linguistics as the classical tradition would
have it:

Notre définition ne concorde pas avec celle, plus restreinte, qu'on en donne
généralement. C'est en effet la
morphologie et la syntaxe réunies qu'on
est convenu d'appeler grammaire, tandis
que la lexicologie ou science des mots en
est exclue. Mais d'abord ces divisions
répondent-elles à la réalité? . . . Est-il
logique d'exclure la lexicologie de la
grammaire? A première vue les mots, tel
qu'ils sont enregistrés dans le dictionnaire,
ne semblent pas donner prise à l'étude
grammaticale, qu'on limite généralement aux
rapports existant entre les unités. Mais
tout de suite on constate qu'une foule de
ces rapports peuvent être exprimés aussi
bien par des mots que par des moyens
grammaticaux.

2.1.1 The word lexicology has generally found favour among
European linguists as the term which denotes the theoretical study of the lexicon, mainly since the publication of G. Matoré's *Méthode en lexicologie* (1953). Matoré devoted a good deal of space in this work to a definition of lexicology which some earlier linguists, including J. Marouzeau,\textsuperscript{10} confused with lexicography. Thus, Marouzeau's entry for *lexicographie* (there is no separate entry for *lexicologie*) reads thus:

Le nom de lexicologie pour l'étude du vocabulaire et celui de lexicographie pour la science des dictionnaires.

It can be seen that the German and English equivalents provided are monovalent, but that Italian has dual translations which are presumably equivalent to the two alternative French entries. Only the final statement of the article takes hesitant steps towards separating the dual terminology, but the definitions are still very unsatisfactory.

2.1.2 Matoré (1953) considers lexicology to be a science which is sadly neglected and whose boundaries are confused with other branches of linguistics such as stylistics, morphology, syntax, and semantics, as well as
with other disciplines such as logic and psychology. He identifies the same distinction between lexicology and lexicography as that which separates ethnology and ethnography. Lexicography, he claims, is the analytical study of vocabulary and as such is a branch of linguistics; lexicology, on the other hand, is a synthetic study which appeals in large measure to sociology. This restrictive distinction is not generally accepted and prompted A. Henry, for example, in his Preface to the *Etudes de lexicologie française et gallo-romane* (1967) to dismiss Matoré's definition and conception of lexicology with the cryptic remark: "Il va de soi que dans mon titre le mot lexicologie n'a pas le sens que lui donne, par exemple, M. G. Matoré."

2.1.3 In a more recent publication, Matoré (1968) has slightly tempered his earlier remarks as he defines the two disciplines thus:

Le mot, en raison de la complexité de sa nature, peut être l'objet d'études très différentes: phonétiques, morphologiques, stylistiques, etc. Limitons notre examen aux deux disciplines qui considèrent le mot essentiellement du point de vue de la signification: ces deux sciences sont la lexicologie (ou sémantique), qui étudie les mots en groupe et qu'on peut considérer à certains égards comme une discipline sociologique, et la lexicographie, qui classe et définit les mots isolés (p. 13).

The division, according to this conception, is between
isolated words (lexicography) and grouped words (lexicology). His loosely appended parenthetical alternative "sémantique" does not really clarify the definition.

2.1.4 Other linguists have also drawn a careful distinction between the two terms, without, however, assigning lexicology to sociology. J. Casares (1950) gives his conception of the dual concepts in his treatise on the problems encountered in the preparation and publication of the *Diccionario Histórico de la lengua española* (1960):

Y de igual manera que distinguimos una ciencia de la gramática y un arte de la gramática, podemos distinguir dos facultades, que tienen por objeto común el origen, la forma y el significado de la palabras: la lexicología, que estudia estas materias desde un punto de vista general y científico, y la lexicografía, cuyo cometido, principalmente utilitario, se define acertadamente en nuestro léxico como el "arte de componer diccionarios" (p. 11),

which clearly places the burden of scientific and theoretical investigation on lexicology. B. Migliorini (1951) takes an identical position in defining the object of his study thus:

Parallela a questa distinzione va quella fra lessicologia e lessicografia: lessicologia è lo studio del sistema lessicale di una lingua, lessicografia l'arte di compilare i dizionari (p. 2).
2.1.5 L. Kukenheim (1966) demonstrates how the word lexicology has undergone a progressive narrowing in definition from the broad, comprehensive study of words in general to a more restricted "théorie sur la façon dont sont composés les dictionnaires" (p. 201). Lexicographers, according to Kukenheim, have paid scant attention to the theory underlying their craft:

Les premiers lexicographes ne se sont guères souciés de la théorie de leur métier: ils rangeaient les mots dans l'ordre alphanabétique ou d'après les sens, sans se préoccuper d'examiner d'autres possibilités. Ce n'est en effet que depuis une vingtaine d'années que nous pouvons parler d'une nouvelle branche de la linguistique vouée à des considérations sur la présentation de la matière lexicale (p. 201).

2.1.6 R. L. Wagner (1967) takes the same position while assigning lexicography to the domain of art:

L'art de confectionner des dictionnaires est une branche de la lexicologie (p. 8).

Both terms, says Wagner, are known as early as 1765 in the Encyclopédie in which lexicography is defined as "the orthography of modern French" (p. 17), since this was a major preoccupation of the 18th century Académiciens, and also as the "art of making dictionaries," an expression which was later to be changed to "science" by Larousse and Littré. In the same work, lexicology is defined as the "knowledge of words" (independent of style) and as a
science which "studies the material of words" (phonetics, orthography, classification, and so on) their semantic "value" and "etymology." This definition persisted at least up until Saussure's generation.

2.1.7 The enormous scope of the discipline is emphasized by Wagner (1967):

Ce terme désigne à la fois l'ensemble des enquêtes, des recherches ayant pour objet les unités lexicales, et les méthodes qui y président. Le champ de la lexicologie est vaste (identification et classement des mots, détermination de leur origine et des conditions dans lesquelles ils sont appelés à désigner quelque chose, étude de leur transmission, étude des valeurs de sens que les mots présentent dans leurs différents emplois) (p. 47).

2.1.8 P. Guiraud (1969:119-120) describes the three main divisions of language study as "phonologie" (the study of sounds), "lexicologie" (the study of words) and "syntaxe" (the study of constructions). The lexicon comprises three major subdivisions: "morpholexicologie," which examines the form of words independently of their function; "sémantique," the study of words from the point of view of meaning and of form as the meaning-bearing feature; and "lexico-stylistique," words studied in relation to expressive and socio-contextual features. The description goes on to examine the various further subdivisions of "sémantique," the object of the study in
question. Semantics, according to this conception, is a branch of lexicology and not merely another word for lexicology as Matoré (1968) implied. But, as Guiraud points out, there is little agreement on the terminology involved in this area of linguistics:

Il reste à fixer la terminologie encore flottante et à décider, en particulier, si le mot "sémantique" doit recouvrir l'ensemble de ces problèmes ou être réservé plus spécialement à l'un d'eux, ce qui est la tâche d'un congrès de linguistes (p. 121).

2.1.9 J. Rey-Debove (1970) has little confidence in the acceptability of either term. Lexicography suffers from the fact that it has been practised for centuries:

L'activité lexicographique se trouve dans une situation doublement défavorable: d'une part on ne sait pas en quoi elle consiste réellement, d'autre part elle n'offre pas l'intérêt de la nouveauté (p. 3).

Lexicology, on the other hand, is a branch of linguistics devoted to the study of lexical phenomena, which has the opposite disadvantage of being too novel to be universally acceptable:

Mais la lexicologie n'a pas trouvé sa place dans la linguistique moderne, faute de bases théoriques suffisantes (p. 3).

She identifies two main reasons for the failure of lexicology as a viable scientific study. The first is
the question of "lexical competence" which, unlike Chomsky's grammatical competence, is far from limitless (all speakers can master their own language with the knowledge of only a small percentage of its total lexicon). The second is the lack of a sound theory of semantics on which to base the description of the "signifié" portion of lexical units as a whole. Rey-Debove concludes that, all things considered, lexicography remains the object of greater interest to linguists:

Aussi bien, de la lexicologie et de la lexicographie, c'est la seconde qui retient aujourd'hui l'intérêt des linguistes, parce qu'ignorant ou repoussant pour des nécessités pratiques les problèmes théoriques constamment posés, elle construit tant bien que mal un modèle naïf, mais général et qui fonctionne (puisque la vente des dictionnaires n'a pas cessé d'augmenter) (p. 7).

2.1.10 A. Rey (1970) has devoted an entire volume to lexicologie. Although he does not examine the matter of terminology as such, he does make frequent references to the distinction between lexicology and lexicography. The latter is depicted at one point as "l'étude empirique des mots de la langue" (p. 1); the former as the "science du mot" and "tributaire de la sémantique" (p. 2). But whereas lexicography is clearly understood to be the practical application of the study of lexical phenomena,
lexicology has no such universally accepted description:

Cette situation paradoxale de fait linguistique contesté par les seuls linguistes a fait de la lexicologie une science incertaine, centrée sur un objet primordial, mais insuffisamment ou mal étudiée. Discipline centrale en linguistique, par ses affinités avec la grammaire des formes (morphologie) et la sémantique; discipline carrefour dans les sciences humaines et en sémiotique, cette lexicologie contestée est en outre indispensable à la plupart des applications de la linguistique, depuis la philologie et la traduction jusqu'à la confection des dictionnaires (p. 1-2).

Once again we are confronted with the enormous breadth of the field supposedly covered by lexicology, and the notion that lexicography is but a small offshoot of the larger discipline is reinforced. A glossary appended to Rey's study gives the following definitions of the terminology which concerns us here:

**Lexicographie:** Description du lexique (au sens large: morphèmes et lexèmes ainsi que leurs combinaisons codées) en une suite d'éléments isolés et ordonnés (des entrées) et comportant l'analyse synonymique de ces éléments (définitions).

**Lexicologie:** Etude du lexique en tant que système, et de ses éléments (p. 290).

2.1.11 For L. Guilbert (1969: 4-7) the production of a dictionary is, in effect, an attempt to describe the lexicon of a language, with the implication that basic linguistic problems will have to be dealt with and that a
theory of language will be an important constituent. But he expresses concern that the lexicographer's task is hampered by the conventional nature of lexicographical description—the arbitrary limitation of the number of entries, their conventional form, and so on. This leads him to make the following distinction between lexicography and lexicology:

Si la lexicologie, branche de la linguistique, tente de fournir une description scientifique des segments de l'énoncé linguistique, la lexicographie, prisonnière des règles traditionnelles résultant d'une longue pratique, considère les unités à classer non pas comme les éléments d'un ensemble mais comme des entités linguistiques définies conventionnellement et que le lexicographe A. Rey appelle "unité de traitement lexicographique" (p. 5).

Guilbert's assessment of the limitations imposed on lexicography by its own venerable lineage and conventional methodology thus coincides with the remarks made by Rey-Debôve (above, 1.3).

2.2 English-language linguists (both European and American) have contributed little to the elucidation of suitable terminology. H.A. Gleason (1967: 85) introduces a discussion on the relation between lexicon and grammar by attempting to define the two terms. The fact that he finds it necessary to clarify the terminology emphasizes the absence, among English-
speaking linguists, of a satisfactory description for the study of lexical phenomena. His definition separates lexicon, "a portion of the total language structure," from dictionary, "the descriptive instrument or product of linguistic research." In other words lexicology and lexicography, although these words are not actually used. Gleason takes to task the proponents of descriptive linguistics whose obsession with grammatical theory has caused them to relegate the study of the lexicon to an inferior level of investigation; it was Bloomfield's statement about the irregularity of the lexicon, he suggests, which caused linguists to seek solace in the more orderly and closed system of grammar. This excessive preoccupation with structural regularity has had its effect on lexicography:

To a very real extent we have made the dictionary exactly what Bloomfield suggests--a list of irregularities, the refuse of grammatical description, the repository of things put aside because they did not give a clear and incisive picture of an all-pervading "structure" (p. 87).

2.2.1 Like many other linguists, Gleason regrets the lack of discussion in the literature regarding lexicography and lexicology and admits to failing in that regard himself (p. 87). One searches in vain in English-language linguistics texts; rarely is either word to be
found in an index of subject headings, let alone as a major chapter title. Even the handful of texts devoted entirely to dictionaries or dictionary-making contain little information with regard to accurate terminology. In *Problems in Lexicography* (1967, eds. F. W. Householder and S. Saporita) the catchword is clearly *lexicography*; the word *lexicology* appears nowhere in the publication. Some contributors to the volume use *lexicon* for theory and *lexicography* for practice, but the majority employ the latter term indiscriminately for both. The *Manual of Lexicography* (1971, ed. L. Zgusta), which is the most complete treatise in English devoted solely to dictionary-making and dictionary theory, also uses *lexicography* to denote theory and practice. In its introduction mention is made of the comprehensive nature of lexicography:

... the lexicographer must take into consideration not only the whole structure of the language in question, but also the culture of the respective linguistic community in all its aspects, to mention only two outstanding fields of his immediate interest (p. 15).

Zgusta follows this statement with a passage quoted from the introductory editorial of *Leksikografičeski sbornik* (1957, No. 1) which reinforces his view:

The theory of lexicography is connected with all the disciplines which study the
Zgusta echoes the sentiments expressed by Gleason (above) and Weinreich\(^1\) with regard to the paucity of theoretical discussion on lexicography and announces that one of the main goals of his Manual is to correct this deficiency. Nowhere in the text, however, is there any attempt to define the terminology and we are left with the understanding that the term lexicography fulfills the dual task of denoting theory and practice.

2.2.2 In Mario Pei's *Glossary of Linguistic Terminology* (New York, Doubleday, 1966) the following entries are to be found:

**Lexicography:** The listing and describing of the words or morphemes of a language, particularly from the standpoint of meaning, with the possible addition of derivation and history.

**Lexicology:** The semantic or morphological study of the linguistic stock of a language, particularly as to content, meaning, or use of the individual forms; the study of the words in a language, their meaning and use, their derivation and history.

The first is a somewhat loosely defined description of the dictionary-making process; it emphasizes the practical and descriptive nature of lexicography and does not suggest that any theory is involved. The appearance of an entry for *lexicology* is surprising in view of the
absence of this word from most English-language texts; in general, the definition seems to concur with that of European linguists with respect to the scope of lexicology.

2.2.3 E. R. Hamp's *A Glossary of American Technical Linguistic Usage 1925-1950* (Utrecht/Antwerp, Spectrum, 1957) is, according to the author, designed to offer explanations of terminology either peculiar to America or used in a sense peculiar to America; terms which do not differ are not listed. Hence, the absence of an entry for lexicography implies that there is more or less universal acceptance of its domain. The term lexicology is entered thus:

\[
\text{lexicology: study of the lexicon; such lexicology often passes out of the field of linguistics into metalinguistics, as when the listing treats the kinds of morpheme arrangements usually known as "words" and "phrases."}
\]

This definition is more notable for what it does not say than for what it adds to our understanding of the terminology in North America; the implication of the quotation is that lexicology is outside the domain of linguistics proper.

2.2.4 *Webster's Third New International Dictionary* defines lexicography thus:
lexicography: 1. the editing or making of a dictionary. 2. the principles and practices of dictionary making.

According to the second definition, we might conclude that lexicography covers both the practice and theory of dictionary making. The entry for lexicology is:

lexicology: the science of the derivation and signification of words: a branch of linguistics that treats of the signification and application of words.

2.2.5 Etymologically the English word lexicology is derived from French lexicologie, and since the discipline is a comparatively new one we should expect the English definition to bear some resemblance to the original. The entry for "lexicologie" in the Petit Robert is:

LEXICOLOGIE nf [Ling] Partie de la linguistique, science des unités de signification (monèmes) et de leurs combinaisons en unités fonctionnelles (mots, lexies), souvent étudiées dans leurs rapports avec la société dont elles sont l'expression.

There certainly appears to be agreement with respect to the salient characteristics: "science," "branch of linguistics," "signification." But the Robert definition is much more careful in attempting to specify the actual units of study, which are not just "words" (a very controversial term in any case); it also alludes to the
sociological aspects of the discipline as Européan linguists see it. Under the entry lexicography the Robert includes the note that it is considered to be a branch of applied lexicology. But this assessment is not necessarily, acceptable for, as J. Dubois (1971: 63) points out, the entries in a dictionary are not, properly speaking, the "words" of a language but an abstraction which, in the case of most dictionary entries, is merely a graphic representation doing service for a variety of possible forms of the word. If this is so, then lexicography and lexicology are not related:

Les entrées lexicographiques ne sont donc pas les mots à proprement parler et la nomenclature ne se confond pas avec le lexique d'une langue. La lexicographie n'est pas une lexicologie appliquée (p. 64).

2.3 This analysis of some current views of the dichotomy lexicography/lexicology leads us to draw the following conclusions. Lexicology is generally agreed to be a branch of linguistics covering a vast area and concerned with all the aspects of lexical phenomena. For this reason it cannot be considered a suitable term to describe the theoretical base for dictionary-making which could only claim to be a small part of the entire discipline as it has been defined. If we choose to accept the notion that there is a branch of "applied lexicology" which deals with the theory and practice of
dictionary-making then we assume that the latter is the application of a competence model of language (langue) for, as Dubois observed, the basic elements of a dictionary are not related to the lexicon but are abstractions, what Rey (1970) refers to as "unités de traitement" or "mots lexicographiques." Hence the notion, expressed by some scholars, of an elementary division between lexicography (practical application) and lexicology (theory) is excessively simplistic and must be rejected.

The term lexicography finds universal acceptance as defining the practice of dictionary-making. For some authors, however, the word implies theory as well as practice and in view of the unique nature of the units of lexicographical treatment it seems appropriate to accept it as covering both. Other scholars have expressed concern at the limitations imposed on the word by the fact that for centuries it has been employed to denote the art (or science) of compiling dictionaries, and we have seen that dictionary-makers have, until very recently, been little concerned about a theoretical basis for their activity. However, since most modern lexicographical ventures are headed by practising linguists, and with the increase in published materials dealing with theoretical features in the dictionary-making process, there is no reason to suppose that the
term lexicography should continue to suffer from this drawback. In the discussion which follows, therefore, the word lexicography will be taken to mean either theory or practice, according to the context in which it is used.
Notes to Chapter 1

1 According to Ali M. Al-Kasimi (1977: 1-2; see Bibliography No. 3) the oldest bilingual dictionaries were made in Iraq about three thousand years ago as a response to a practical communications problem between Assyrians and Sumerians.

2 J. R. Hulbert (1968), p. 16 (Bibliography No. 2).

3 E. L. McAdam Jr. and George Milne (1963), p. vii (see Bibliography No. 2).


5 Notable are: the Colloquium on Lexicology and Lexicography held in Strasbourg, France in 1957; cf P. Imbs (1961); the Conference on Problems in Lexicography in Bloomington, Indiana, in 1960; cf. F.W. Householder and S. Saporta (1967); the Conference on Lexicography in English in New York, 1971; cf. R.I. McDavid and A. Duckert (1973); references in Bibliography No. 2.

6 See Bibliography No. 2, "Bilingual Lexicography."


8 The revision of The Canadian Dictionary/Dictionnaire canadien, eds. J-P. Vinay et al., Toronto: McClelland and Stewart, 1962, now in progress at the University of Victoria.


12 However, as early as 1913, Henry Sweet in Paper 31 of his Collected Papers, Oxford, distinguished between what he called "lexicology", which deals with the special facts of language such as the naming function of words, and "grammar", which is concerned with the general facts such as the methods of forming plurals. This view was later taken up by Otto Jespersen in The Philosophy of Language, W.W. Norton, New York, 1965, pp. 31-32, and considerably enlarged upon.

13 "The indifference which lexicography displays towards its own methodology is astonishing." U. Weinreich (1967: 26; Bibliography No. 2).
References for Chapter 1


CHAPTER 2

HISTORICAL DEVELOPMENT

1.0 The first lexicographical ventures were bilingual dictionaries which developed as the inevitable adjunct to a period of bilingualism. In their crudest form, dictionary entries consisted of glosses—translations of rare words and commentaries on unusual semantic content; but it was only a matter of time before some scholar recognized the advantages to be gained by gathering such annotations in one place, where they could be consulted easily and rapidly, although facility of consultation was a later development of alphabetically-ordered lists.

The respective histories of English and French lexicography are well documented and a serious study of the chronological development of bilingual lexicography can be found in Anderson (1971). In this chapter, we undertake an analysis of the developmental factors which have influenced modern approaches to bilingual dictionary-making; the main emphasis is placed on characteristic elements of bilingual lexicography rather than on the individuals who motivated them.
1.1 Arrangement of Entries

The concept of a device to permit communication between two language groups was a direct result of contact between different linguistic communities, leading to a state of bilingualism. But the device was not, in its primitive form, a dictionary as we now know it. The first hesitant steps towards a "dictionary" were taken in the 7th and 8th centuries with the appearance of so-called glosses, consisting of explanatory remarks in the margins of grammars and Latin teaching primers. Their modest aim was the elucidation of otherwise obscure Latin passages in the text being studied. Medieval lexicographical output, if it can be described as such, manifested itself in the spawning of a more organized form of the glosses, which came to be called glossaries and are the real precursors of modern dictionaries. They followed parallel lines of development in England and in France and arose from a need for a more readily accessible word list. But the notion of a logically ordered sequence to facilitate consultation apparently did not suggest itself to those early lexicographers; on the contrary, the first glossaries were little more than a gathering of glosses from many different texts and primers. More often than not, they consisted of a hectic conglomeration of unordered, even unrelated, lexical items whose only common factor was the exegesis of Latin
elements in the vernacular.

1.2 A first attempt at introducing some kind of order to the contents of a glossary was inspired by the prior existence of the Latin *vocabularia*—word lists arranged in ideological fashion as a teaching-aid. It was not until the 15th century that the notion of alphabetical ordering presented itself to the medieval copyists, who began by classifying words in first-letter order only and eventually in second- and even third-letter order. In the earliest period of lexicography, the modern vexed question of "absolute order" versus "sense order" was of little consequence, for the number of words was strictly limited and the source language, Latin, was devoid of the compounds and phrasal elements which many modern lexicographers still refuse to treat as entries.

1.3 The earliest-known ordered dictionary is the *Medulla Grammaticus* (1500) which comprises a crudely-alphabetized Latin-English word list; Robert Estienne's somewhat more sophisticated *Dictionarium seu Linguae latinae thesaurus* (1531) was a similar undertaking for French and Latin. Although William Caxton published his *Vocabulary in French and English* somewhat earlier (1480), this work more properly belongs to the pre-dictionary stage. It is little more than a commercially-oriented phrase book with scant attention paid to orderly presentation for ease of
consultation. The era of French and English dictionaries was heralded by the *Dixionary frensche and englesche*, first published in 1567 by Lucas Harryson; it is a one-way work, with French entries alphabetized according to classical Latin tradition which uses V in place of U and W, and I for J. Second- and third-letter order is somewhat haphazard; for example, NO- words precede NI- words, some SE- entries follow SC- entries while others precede. There is rarely any attempt to alphabetize words beyond the fourth letter.  

1.4 This system of ordering was imitated by all immediately succeeding bilingual lexicographers of the 17th century, including Randle Cotgrave (*A Dictionarie of the French and English Tongues*, 1611) and Guy Miège (*A New Dictionary, French and English, with another English and French*, 1677), although the latter distinguishes graphically between U and V and alphabetization is consistent even beyond the fourth letter. Miège also broke new ground and established a model for many later editors by attempting word-family grouping under a primitive head word, the derivatives also often appearing in their normal alphabetical order accompanied by a cross-reference. Nor did dictionary order undergo any major revision in the 18th century, since Thomas Nugents' *A New Pocket Dictionary of the French and English Languages* (1767) still combines I and
J entries as well as U and V entries, and although K entries are now distinguished, W and X are still absent. The first attempt at differentiating U-V and I-J entries was by Louis Chambaud (*Nouveau dictionnaire français-anglois et anglois-français, 1776*); but this distinction only applies to first-letter-order.

1.5 By the 19th century the vacillations of alphabetical order and classical combinations had been regularized, and the only differences observed in the ordering procedures of general dictionaries are those relating to word-family grouping as opposed to absolute alphabetical order—a dichotomy which, as we shall see in the detailed discussion on this topic (Ch. 11 in Part II), remains unresolved in the last quarter of the 20th century. User convenience remains uppermost in the minds of dictionary-makers and each system adopted for ordering the entries has its own particular advantages and drawbacks.

2.0 Canonical Form of Entries

The entries in Harryson (1567) established a pattern for the canonical form of dictionary entries which endured for the next four centuries. His penchant for capitalized initial letters is still employed in a number of 20th century dictionaries, to the continued confusion of users who are thereby prevented from learning whether the word is capitalized or not in running text. For
bilingual dictionaries this information is indispensable since two languages do not necessarily share the same philosophy with regard to capitalization.

2.1 Miège (1677) chose to present his main entries in upper case characters, once again a fashion imitated by succeeding generations of lexicographers and still used by some today. Although this procedure, designed to give prominence to the entry, did nothing to solve the problem of distinguishing between capitalized and non-capitalized words, he did add one refinement—the use of lower case characters for sub-entries. In his preface he describes this arrangement as his "Method," a first attempt at etymological or topical grouping as a logical interruption to absolute alphabetical order. His system also called for the use of cross-references in case the user should look for the word in its strictly alphabetical order; this, again, is the procedure still used by modern proponents of "word-family" grouping, although Miège's dictionary is somewhat inconsistent in applying it. Miège was obviously very conscious of user-convenience, for he also made use of different type fonts in order to make various elements of the article stand out—roman for definitions, italics for translations—in addition to upper and lower case roman for entries. Font variation is still the most commonly used device in
bilingual dictionaries to emphasize the separation between SL and TL. Indeed, apart from the development of more varied type fonts (such as boldface, small capitals, phonetic characters), there have been few advances in typography since Miège's day.

2.2 The canonical form of entries has not undergone any radical change since the 16th century. Harryson (1591) presents his French verbs in the infinitive; if the conjugation is irregular, the infinitive entry (with its translation) is followed by certain of its principal parts, also translated:

\[ \text{Hair,} \quad \text{to hate} \]
\[ \text{Hai,} \quad \text{hated.} \]

Yet the provision of morphological variants was inconsistent. Harryson, for example, omitted to give the principal parts of *avoir*. Succeeding lexicographers, starting with Hollyband (1593), also listed verbs under the infinitive. The principal parts of irregular verbs, if they were provided, were given as first person singular examples:

\[ \text{Vouloir, je veux, je voulus, j'ay voulu, je vouldray, to will.} \]

Miège (1677) was the first to produce a two-way dictionary and it is to him that we must turn to find the
system employed for English verbal entries. They were given in the full infinitive form:

    TO CANCEL,

with derivatives following as sub-entries in accordance with his grouping "method." Thus:

    Cancelled,
    A Canceller,
    A Cancelling,

and so on.

2.3 Chambaud (1776) recognized the need for non-English speakers to be provided with more details on the morphological variants of verbs and therefore presented his verbal entries as infinitives, followed by the present participle and preterite for weak verbs and by present participle perfect and preterite for strong verbs, thus:

    To ABET, ETTING, ETTED.
    To ABIDE, ING, ABODE, ABID.

Nugent (1767) broke with tradition by omitting the "to" from his English verb entries; the motivation was very likely space economy since he was producing a pocket edition; he also omitted principal parts even of
irregular verbs. But it was almost a century later that his system of entry-without-preposition was generally adopted for English.

Nineteenth and twentieth century lexicographers have varied only slightly from Chambaud's practice in the presentation of verbs. The "to" of the infinitive is considered redundant and dropped; regular verbs are presented only in the infinitive, while those whose only stem variation consists of a consonant doubling are so marked; irregular verbs are given with the requisite variable forms:

- *abet [-tt-]*
- *abide, abided or abode.*

2.4 The presentation of other lexical forms did not seem to warrant the same consideration by early dictionary-makers. It was not until Cotgrave (1611) that any attempt was made to show gender inflexions for adjectives:

- *Abhorre: m. éé: f.*

The canonical form of the entry had hitherto always been exclusively masculine and Cotgrave's addition of the feminine, together with a marking symbol, was truly an innovative approach.

2.5 One final remark on entries is called for by a remarkable 18th century attempt at formalizing the
presentation of English verbal phrases. Even today many lexicographers have failed to recognize the importance of these lexical units and they are more often than not relegated to the status of examples under a primitive entry. Chambaud (1776) gives a very thorough treatment of such verbs, as reference to the entry *come* will indicate. Presented in formal alphabetical order as self-contained entries, the list included forms which are found only in the most exhaustive dictionaries of the twentieth century:

- **TO COME ABOUT** [to change, to come around]
- **TO COME AT or BY** [to obtain, to reach]
- **TO COME IN FOR**
- **TO COME OUT WITH** [= to give vent to]
- **TO COME UPON.**

Chambaud was perhaps the first lexicographer to conclude, from the existence of self-contained equivalent forms in the other language, that English verbal phrases are usually lexical units with a unified semantic value and that therefore they merit consideration as full-status entry words.

3.0 **Grammar Identification**

Primitive dictionaries, like the glossaries which spawned them, were essentially word lists devoid of what
might be called "lexicographical frills." Their purpose was simple and clearly-defined: to enable readers of a foreign language text, namely Latin, to find the meanings of words and phrases which were expressed as broad translations in the vernacular. On the few occasions when homographic form might cause ambiguity, it was assumed that the contextual setting of the definition would identify the appropriate lexical item. For this reason, early lexicographers had no reason to identify the grammatical category to which their entries belonged, and we find in Estienne (1531) the following kind of article:

Louer, laudare.
Lourd, Secours, vel Secors, Obtujus.
Poil, Villus, Pilus.

3.1 The first dictionaries matching French and English apparently adopted the same philosophy, for neither Harryson (1591) nor Hollyband (1593) gives any grammar indication whatever. Cotgrave (1611) deemed it appropriate to identify the gender of adjectives:

Detrimenteux: m. eu/e: f. Hurtfull,
dommageous; . . .

and nouns were somewhat inconsistently noted as "m" or "f":


Enharnachement: m. The harneis, or furniture; al/o, a harneis/ing, of a horse.

Pourtraicture: f. A pourtraiture; or as Pourtraict

Sometimes grammar was omitted altogether.

Guerrite: A kind of defence in a rampier; see Guerite

Septieme. The seventh.

Taige. A kind of vine; q Rab.

3.2 It was almost a century before any improvement was made on the presentation of grammatical information. Miège (1677) imitated Cotgrave's gender indications with only a small typographical change--the enclosure of the gender notation in parentheses:

BAGUE (f.).

In view of the fact that dictionaries were by this time generally two-part works, with the assumption that users might wish to generate texts in the non-native language, the absence of grammar indications is quite surprising. But finally Boyer (1699) recognized the need for them when he established a pattern still followed by many twentieth-century lexicographers. French nouns are entered with S.M. or S.F.; adjectives in both languages are followed by Adj., with gender indications for French
where needed; verbs are generally identified as transitive or intransitive using the notation: Verb. Act., Verb Neut., which were subsequently abbreviated to v.a. and v.n. and are still used in some 20th century works: 15

To MOVE, Verb. Act.

Musseau, S.M.

Possible, Adj.

3.3 Nugent (1767) was the first bilingual lexicographer to give systematic treatment to grammatical categories and his method has persisted, largely unaltered, for more than 200 years. All verbs are entered as v.a. or v.n., reflexives as v.r., French nouns as s.m. or s.f., English nouns as s.; other notations are: prep. (preposition), ad. (adverb), a. (adjective). Locutions of the type Mettre en cage are left unlabelled and appear under the particularizing word (cage), thereby being reduced to the level of examples under that headword. 16 The only variations observed in 20th century dictionaries are the use of n. for nouns, the notation adv. for adverbs and adj. for adjectives, and the use of the more modern verbal terminology v.tr., v.intr. and v.pron. 17

4.0 Pronunciation

A very early sign that pronunciation would become an
element of bilingual dictionaries was the appearance in Hollyband (1593) of a brief description of the pronunciation of the French vowel a (under A) and the initial group qu- (under Q). Cotgrave (1611) went considerably further than this by adding a grammatical appendix containing a detailed discussion of the sounds of French. It is a general discussion giving suggestions for the articulation of vowels, consonants (including nasal consonants) and commenting on elision and stress placement.

4.1 Boyer (1699) reverted to Hollyband's imperfectly formulated phonetic presentation by putting his descriptions in the body of the work. Under each initial letter he gives a very full discussion of the various possible pronunciations, according to phonetic environment, of all letters in both French and English. A crude attempt at phonetic notation is often given:

B S.M. The Second Letter of the Alphabet is sounded bé in French

Lexicographers of the 18th and 19th centuries apparently saw no reason to enlarge upon this kind of approach to pronunciation, other than correcting errors and adding examples of their own. The standardization of orthographical systems also occasioned a few changes to the remarks made by lexicographers in 16th and 17th
century dictionaries.

4.2 By the 20th century the advances made in phonetic research and the appearance of more sophisticated systems of notation began to make themselves felt in lexicography. Although some dictionary-makers, such as Gasc, continued to consider pronunciation as a non-lexicographical feature and eliminated it from their work, the majority took steps to include more or less complete phonetic coverage. We have observed that their predecessors were content merely to provide a few notes on the pronunciation or orthographical elements, either in the body of the work or in special appendices; modern lexicography now called for precise phonetic indications for every entry.

4.3 Bellows' 1910 edition probably represents the first steps taken by a 20th century bilingual lexicographer to implement this requirement. His method of notation was what he describes as "phonetic spelling" (p. 6) which he placed between square brackets immediately following the entry; it was at once both novel and ingenious. In the absence of any more accurate method of transcription, Bellows developed a system which he claimed to be operable for all but "one word in 20" (p. 6). Thus, for example, his notation $^i$ for the phonetic sound [i] works for long "i" but not for short "i" as in the word
"simple." The difficult "th" sound of English is channelled into its phonetic realizations by means of the marks "fort" for [θ] and "faible" for [ð].

4.3.1 Bellows excuses the inaccuracies of the system on the grounds that it is better than none at all which is what his predecessors virtually provided their readers with. He is concerned, for example, that the unstressed "y" suffix of English ably, pity, fully, etc. should be rendered by the notation "é" rather than "i" because the slight exaggeration of é is much less unpleasant than that of i, which has an air of caricature about it, and raises a smile at the speaker's expense" (p. 6).

4.3.2 Bellows was apparently under the impression that English speakers had less need of phonetic representation for French than French speakers for English, for the English lexicon is quite fully treated whereas the French contains only a few exceptional pronunciations. This fact is in sharp contrast to all the other indications in the dictionary which suggest it was designed for English-speaking users (all preface notes, labels and so on in English). Although he fails to mention it in the preface remarks on pronunciation, it is probably because his "phonetic spelling" is largely based on French pronunciation and uses French orthography and diacritics; in a majority of cases, therefore, his phonetic notation
would add little to the already phonetically-oriented French orthography.

4.3.3 The phonetic representation is also novel for its inclusion of English primary stress, marked by means of an acute accent following the stressed segment; it is also marked in the orthography; secondary stress is not marked:

\textit{Lib\'erate} \quad \text{[lib\'e-rète].}

Words which are identical in both languages, and which appear in the French section only, usually contain the English pronunciation only; in exceptional cases, both may be given:

\textsc{LICHEN} \quad \text{[lee-kenn]} \quad -- \quad \text{laïke-nn},

although we note in this case that Bellows has omitted to indicate stress. Some French entries considered worthy of phonetic treatment are:

\textsc{LIAIS} \quad \text{[lee-ay]}

\textsc{POIDS} \quad \text{[pwha]}

\textsc{ENNEMI} \quad \text{-E} \quad \text{[enn-mee]}.\quad

But the absence of notable pronunciation deviations such as \textsc{oignon}, \textsc{lis}, \textsc{six}, \textsc{oeuf} and \textsc{événement} is inexcusable, and there are also some serious errors:
AGNÉAU  [ine-yo]
BESOGNE  [pr. be-zoine]
BERCEAU  [bare-so].

French nasals are sometimes indicated:

ESSAIM  [m nas.]
ENSEMBLE  [em as an nasal]

and certain phonetic exceptions are noted:

PNEUMATIQUE  [pr. p]
ADONIS  [pr. the s].

Aspirate "h" is shown by the mark "$p" appearing before the entry; it is even somewhat redundantly used in the English section, even though only a small handful of words are unaspirated. Under the entry liaison in the French section Bellows offers a small treatise on the rules for French liaison and explains his system for marking lack of liaison by means of a double period.

4.4 'Inaccurate and inadequate as it was, Bellows' treatment of pronunciation was a major advance over all of his predecessors and paved the way for the more sophisticated representations which soon followed. The establishment of the International Phonetic Alphabet persuaded many authors to give a complete pronunciation guide in the preface and a full phonetic transcription of
every entry in both sections (or at least in the non-native language for monodirectional works). Their decisions with regard to the variety of English or French represented probably affected the pronunciation of several generations of language-learners, since the European-dominated bilingual dictionary "industry" provided the best southern British as the English model and the so-called "Parisian" for French. North America's 220 million English speakers and 10 million Francophones were obliged to wait until 1962 and the publication of The Canadian Dictionary/Dictionnaire canadien to find their dialects represented. 21

5.0 Ordering of Senses

The notion of applying a structured internal order to the contents of a dictionary entry is a comparatively recent innovation, usually attributed to Samuel Johnson, who, in his Plan of a Dictionary makes the following claim:

In explaining the general and popular language, it seems necessary to sort the several senses of each word, and to exhibit first its natural and primitive significations, as

To arrive, to reach the shore in a voyage. He arrived at a safe harbour.

Then to give its consequential meaning, to arrive, to reach any place whether by land or sea; as, he arrived at his country seat.

Then its metaphorical sense, to obtain anything desired; as, he arrived at a peerage.
Then to mention any observation that arises from the comparison of one meaning with another; as, it may be remarked of the word arrive, that in consequence of its original etymological sense, it cannot be properly applied but to words signifying something desirable; thus, we may say a man arrived at happiness, but cannot say without a mixture of irony, he arrived at misery.  

Dr. Johnson was concerned that "the popular part of language . . . admits of many distinctions and subdivisions," and his insistence on marking these classifications established a precedent followed by dictionary-makers of all languages.

5.1 In his study of a number of entries selected from bilingual dictionaries of the 16th to 18th centuries, Anderson (1971) points to similarities in the arrangement of equivalents, noting that many early lexicographers turned to Robert Estienne's *Dictionnaire francois-latin* (1531) as a model for their own work. Thus, Estienne's entry abaissement is largely imitated by Harryson, Hollyband and Cotgrave in the use of particularizing words. Hollyband (1549), for example, has:

\[
\text{Abaissement de voix ou de son état, a speaking soft, a decay of estate.}
\]

Cotgrave (1611) presents a lengthy list of synonymous English equivalents for the same entry word:

\[
\text{Abaissement. a debasing, abasing, dejecting, humbling, prostrating.}
\]
bending downwards, bringing downe, making low; also a stooping; decreasing, falling, waxing low.

The punctuation in this entry can be considered a first attempt to classify different semantic levels, although even within the categories marked off by semi-colons further distinctions of meaning may be detected. But this dictionary was clearly designed with English-speaking users in mind, and the need to indicate the contexts in which each of the equivalents might be used was not under consideration.

5.2 Chambaud (1767) went a step further than any of his predecessors by assigning two base meanings to abaissement and following each of them with a series of descriptive phrases indicating current usage. The primary meaning "diminution de hauteur" is illustrated by several translated sentences:

l'abaissement du mur . . . , the lowering of the walls,
l'abaissement des eaux, the abatement of the waters.

The secondary meaning is given as "abjection, humiliation" in a moral sense:

Un parfait Chrétien doit se plaire dans l'abaissement, A true Christian ought to delight in humiliation.
The publication date of Chambaud's dictionary, some twenty years after the appearance of Johnson's *Plan*, seems to suggest that he was influenced to some degree by the latter's innovative approach to lexicography. In fact, for more than two centuries Johnson's categorization has been the basis for entry structure in almost all monolingual dictionaries as well as the bilingual dictionaries which tended to be modelled on them. The classical historical dictionaries arrange entries beginning with the "natural and primitive significations," developing from these to the "consequential meanings" and finally the "metaphorical senses," the entry may be concluded by the addition of remarks arising from the "comparison of one meaning with another."

5.3 A few 20th century dictionaries have deviated from this established model to experiment with user-oriented convenience and logical organization based on frequency or alphabetical arrangement. Yet the most widely-used bilingual dictionaries are clearly modelled on the historical organization of monolingual works, even though this is rarely the most convenient arrangement for ease of consultation.

6.0 Meaning Discrimination

The modern lexicographical notion of disambiguating the various senses of a word by means of contextual
material and discriminatory labels is comparatively recent. Early lexicographers chose to use definitions in the target language, accompanied by numerous examples, to clarify different semantic levels within the framework of each entry. But, as the following entry from Randle Cotgrave's 1611 dictionary illustrates, there was at first no systematic attempt to provide the entry with a logical internal structure:

Bachelier: m. A bacheler; a youth of sixteene, or eighteen yeares of age; also he that hath paas'd Master in a trade, but is not yet sworne of the Companie; also, a Batcheler of Art; also, a title of gentrie inferiour to Banneret, and superiour to Es/cuyer; a young gentleman that a/pires unto Knighthood, and the priviledge of bearing a Banner in the field; also, the Lord of a caste, fort, great house, or Place, that is derived from (but in iuridiction equall with) an Earlėdom, Vicountie, or Baronie.

The first meaning, that of "a young unmarried man," corresponds to the usual first sense provided in modern dictionaries; it is presumably the most frequent acceptation. Then follows a designation, unknown in the modern language, relating to trade guilds of the Middle Ages. The second modern meaning is referred to in the brief remark "a Batcheler of Art," while the third (young knight) draws several lines of explanation thereby indicating the importance of this concept in Medieval society.
6.1 For words of lower semantic content, however, Cotgrave and his immediate successors did little more than provide lists of unexplained equivalent synonyms and it was not until the end of the seventeenth century that an attempt was made to discriminate between the various possible shades of meaning. Abel Boyer's 1699 dictionary contains a preface which begins by acknowledging the debt he owed to his main predecessors, Cotgrave and Guy Miège, and goes on to take them to task for not providing sufficient explanatory material. By way of illustration, Boyer's gloss for the English verb "to move" contains the following subdivisions:

1. (to stir) mouvoir, remuer.
2. (to shake, in a proper or figurative sense) mouvoir, branler, ebranler, remuer, agiter.
3. (to stir up, egg on, sollicite, persuade) mouvoir, emouvoir, exciter, inciter, animer, pousser, porter, exhorter, solliciter.
4. (to affect or touch) emouvoir, toucher, attendrir.
5. (to provoke to anger) faire mettre en colère, facher irriter, provoquer.
6. (to cause laughter) faire rire.
7. (to propose, speak for something) proposer une chose, en faire la proposition ou l'ouverture, la mettre sur le tapis, la mettre en avant.
8. (to stir, be in motion) se mouvoir, se remuer, faire un mouvement, être en mouvement, bouger, branler, etc. 26
This highly sophisticated arrangement of the various senses of a difficult English word was the model for later lexicographers who have changed only the typographical layout of the entry; most modern dictionaries merely employ an abbreviated form of Boyer's parenthetical material.

6.2 In 1776, Louis Chambaud published his French and English dictionary in which, for the first time, discriminatory labels in parentheses precede each translation equivalent. It is also worthy of note that these labels are in the source language:

CABARET, f.m. [taverne, maison ou l'on donne à boire & à manger pour de l'argent] Tavern, house of entertainment. Cabaret à bière, An Ale-House. Un cabaret borgne (mauvais et peu frequente) A blind tavern, a poultry place.

6.3 By comparison, the modern entry for cabaret, even as illustrated by the voluminous HNS, differs little from its 17th century predecessor except with regard to modern derivative meanings and typographical procedure:

cabaret s.m. 1. (a) = public house, tavern, wineshop; F: pub; (b) inn, eating-house; (c) (small fashionable) restaurant; night-club, cabaret. 2. (a) liqueur-stand; (b) tea-, coffee-, chocolate-) set, service (with its tray); cabaret; (c) Nau: fiddle; (d) Fr. C: tray. 3. Bot: (a) asarabacca; (b) c. des murailles, Venus's navelwort; (c) c. des oiseaux, fuller's teasel. 4. Orn: (North European) Linnet.
7.0 Contextual Examples

Present-day bilingual lexicographers subscribe to the view that the user's interests are best served by providing as much contextual environment as possible in order to illustrate the variety of translations which different situations will demand. Even the use of discriminatory labels does not necessarily set up a one-to-one correspondence between a SL entry word, and its multiple equivalents in TL, since different settings may require (or at least permit) a number of possible renderings. Mediaeval lexicographers were not unaware of this state of affairs and often furnished a wealth of contextual material, "locutions figées" and proverbs.

7.1 Thus, for example, Louis Chambaud includes, in his very extensive entry for the word affaire, the following:

Il a poussé l'affaire d'une assez vigoureuse manière, He has carried on the affair with a high hand.

Voyez un peu la belle affaire! A fine matter truly to make such a pother about!

C'est là le bon de l'affaire, That's the glory of it.

C'est une affaire, Tis hard to do; Tis trouble.

Avoir affaire d'argent, To have occasion for money.

While Cotgraves' entry abandon incorporates a proverbial
expression, labelled as such:

Abandon fait larron; Prov. Things carelessly left, laid up, or looked unto, make them theves that otherwise would be honest; we lay, fa't hind, fa't find.

Chambaud's treatment of abandon includes the adverbial locution à l'abandon which he renders as "at sixes and sevens" and further illustrates thus:

Tout est à l'abandon, Every thing is at sixes and sevens or at rack and manger.

7.2 Modern dictionaries have found little cause to improve upon the methodology of these predecessors in this regard, as the following entry from CASS exemplifies:

abandon nm Relinquishment, surrender, cession; abandonment, the state of being forsaken, forlornness, destitution; ease, lack of restraint. A l'abandon, at random, in confusion, at sixes and sevens, (Naut.) adrift; laisser ses enfants à l'abandon, to neglect one's children; se livrer avec a. à, to indulge in; un champ qui est à l'a., a field left to run wild.

8.0 Cross-references

Most 20th century dictionaries employ a system for referring the reader to other related material which will assist in reinforcing the meanings and equivalents provided. Sometimes the cross-reference directs the user to synonymous information; in other cases it indicates an
essential relationship between a word and its derivatives which may be separated by the exigencies of alphabetical ordering. The practice has a venerable history, since the earliest lexicographers used cross-references to a greater or lesser degree. For example, Cotgrave's French entry bail ends with the note "See garde"; Boyer's article aboyer refers to an alternative orthography abayer. Cotgrave terminates a curiously uninformative article on the word badin with a proverb which he does not translate but for which he refers the reader to the entry where it is explained: looke Escolier.

8.1 Miège was perhaps the most thorough of early lexicographers in his use of cross-references which constituted an important feature of his "Method" and enabled alphabetical order to be breached in order to maintain morphological relationships. Even verb parts are listed as entries with a note referring the reader to the base word:

Etant. V. Etre
Été (from the Verb Etre) V. Etre.

Moreover, Miège improved on his precursors by employing a consistent, abbreviated form for the cross-reference—V. (= voir), instead of the numerous words indiscriminately used by other authors—"see," "looke," "as," "also," "seeke," "cherchez," "voyez," etc.
The preface to a dictionary, although often sadly neglected by the readers for whom it is written, has always been considered by lexicographers to be an indispensable introduction to their work. It may vary from a brief statement of acknowledgement to contributors and colleagues, to an extended essay setting forth the author's philosophy of translation and containing a wealth of useful linguistic information. Although some earlier prefaces contain information on the authors' rationale, Cotgrave's is probably the first to set forth in some detail a lexicographer's aims and methods. For example, he justifies the inclusion of rare words by indicating that they are the result of a careful study of contemporary output:

Il te supplie bien fort, si tu trouves icy quelques mots qui sonnent mal auz oreilles, ou me/mé qui n'y ayent encore jamais sonne, de croire, qu'ilz ne sont point de son invention, mais recueilliz de la multitude & diversité de noz Auteurs, que possible tu n'auras pas encore luz, & qui, tant bons que mauvais, désirent tous d'etre entenduz. 27

9.1 Miège, some 60 years later, devoted three pages to his preface which contains, in addition to an acknowledgment of his predecessors' work and an indication of his motives for producing the present volume, a fairly detailed account of his so-called "Method" and the extent
to which it improves upon the methodology of earlier
dictionaries. Thorough as his treatment may have been,
however, Miège was severely criticized by another later
lexicographer, Boyer, who detected faults of inconsist-
tency and an excessive reliance on Biblical allusions for
contextual examples. What seemed to have become a fetish
for prefatorial commentary on earlier dictionaries, was
elaborated upon by Boyer's inclusion of a statement
regarding his main sources--the monolingual French
dictionaries of Furetière and of the Académie. His
preface, moreover, was followed by a list of abbrevia-
tions and special symbols employed in the work, thereby
establishing a model for all later dictionaries.
Although some of his predecessors had included some
material of this nature, none had applied such thorough
treatment to it.

9.2 Boyer, in turn, suffered severe criticism at the
hands of Chambaud whose preface proclaimed:

... Boyer's Dictionary, even the last
dition, is full of enormous faults, as fal/e
genders, fa\textit{\l}e con\textit{\l}tructions, forged
express\textit{\l}ions and barbarous English.

Chambaud goes on to take issue with specific items of
Boyer's methodology, such as his morphological word
grouping which is confusing to the young user, and
justifies his own inclusion of a large variety of
language registers on the grounds that all levels of society should be able to communicate effectively with each other.

9.3 Boyer maintained that the preface of a dictionary was the author's opportunity to set forth his own view of the aims of lexicography and to rationalize them by identifying the weaknesses in earlier productions. Succeeding lexicographers continued to make statements on their own methodology, but judiciously refrained from passing judgement on that of their predecessors or contemporaries. In modern bilingual dictionaries a typical preface, which may also be subsumed in an "introduction," contains a brief statement justifying the publication, followed by abbreviation lists, special symbols, a description of phonetic and orthographic content and various other materials of pedagogical significance.

9.4 The amount of detail provided may be a function of the audience to which the work is directed. HAP3, for example, in spite of its voluminous dimensions, contains a single page of preface in the 1940 revised edition of the English-French volume; that page is devoted to a brief mention of the perennial problems of English hyphenation followed by a few remarks on typographical procedure and a statement on the sources of new
The preface to the 1972 revision of the French-English section (HNS) is expanded to 5 pages with the addition of a French language version, indicating that non-English speakers may also have some use for its contents. Apart from a justification for the revision, however, the information contained in it is only marginally different from that of the earlier version; the editors expect their readers to have little need for instruction in the use of their dictionary or in the structure of its languages.

9.5 Other dictionaries, however, clearly have a younger or less sophisticated public in mind as their prefaces incorporate a variety of information of a purely pedagogical nature. GASC, for example, includes a 3-page section entitled "Observations," devoted to some enlightening insights into the nature of translation and ending with a table of equivalent weights and measures. CD1 has gone further than any other dictionary in this regard with an introductory section totalling 28 pages of detailed information on the orthography, phonology and syntax of both languages and including an extensive verb table for French and remarks on the regional variations of Canadian English and Canadian French. Nor are the English and French prefaces translations of each other; the section written in French is a discussion of English materials.
linguistic problems for Francophone users, and vice versa for the English section. This kind of approach to prefatorial information emphasizes the true bipartite nature of a dictionary destined for users of both language communities.

10.0 From this brief review of the chronological development of bilingual lexicography it can be seen that the most important and fundamental methodological features were well established by the end of the 18th century. Estienne's French-Latin list appeared to provide the basic model for the first two-language dictionaries, with crude alphabetical order and a basic one-to-one correspondence of lexical units. Harryson's productions of 1567 and 1591 implemented more or less identical procedures for the first (one-way) French and English dictionary although, as we have seen, he systematized the canonical form of the entries by issuing verbs in the infinitive and providing run-on material where appropriate. As early as 1593, Hollyband had taken the first hesitant steps towards phonetic representation which Cotgrave, eighteen years later, enlarged upon with details on French pronunciation incorporated in the body of the work. Cotgrave also introduced novel features such as gender inflexions and extensive use of cross-referencing. The notion of the pedagogical value of
bilingual dictionaries had already made its appearance.

10.1 In 1677 the first two-way dictionary was created by Guy Miège. Among other things, this dictionary experimented with word order by introducing "word family" grouping with related alphabetical order and cross-references for detached entries. Variations in the use of type fonts and a full preface containing a description of his lexicographical "Method" made this 17th-century dictionary as complete as many 20th-century works. Boyer's 1699 dictionary implemented Miège's methodology and systematized the representation of abbreviated grammatical categories; he refined the application of disambiguating material and enlarged upon phonetic detail and prefatorial information.

10.2 The only changes effected by 18th-century lexicographers were relatively minor ones: Nugent removed the "to" from English infinitive verb entries in 1767; Chambaud (1776) followed Dr. Johnson's "Plan" with regard to internal entry structure and extended entry status to phrasal verbs. In the 19th century, authors seemed only to find it necessary to improve on layout, having recourse to much improved typography in order to enhance the appearance of the complex lexicographical material. Bellows (1873) and Gasc (1876), representative of this period, introduced a variety of non-lexical material as
pedagogical adjuncts to the more properly lexicographical fabric of their dictionaries. Twentieth-century dictionary-makers were able to benefit from the findings of linguistic research, but their contribution to the development of bilingual lexicography will probably be regarded as one of consolidation and refinement, rather than originality and inventiveness. Mansion's 1934 dictionary, a work of monolithic proportions, is notable for its vast coverage; but on close analysis it cannot be said to have added much which was not already at least implied, if not fully implemented by lexicographers of the three preceding centuries. Perhaps the main claim of the 20th century to innovation is in the application of computer technology to production methods, although this approach to lexicography changes little in the actual methodology and the resultant output; the technological contribution here is in the area of accuracy and consistency, attributes which the machine exacts from its users to the eventual benefit of dictionary users.
Notes to Chapter 2

1 Cf. J. and C. Dubois (1971), p. 34.

2 For French the most authoritative and complete study is B. Quemada (1967); it includes a chronological bibliography, detailed methodological analysis and a small section on bilingual dictionaries. Other reliable sources are G. Matoré (1968), A. Rey (1970), R-L. Wagner (1967). The history of English lexicography has not been treated in such depth in a single volume, but various titles may be consulted in order to observe the whole perspective: Murray (1971), Starnes and Noyes (1946), Starnes (1963), Friend (1967), Hulbert (1968). For bilingual lexicography of French and English covering the 15th to 18th centuries, see J. Anderson (1971). A chronological bibliography of French and English dictionaries is given in Appendix 1.


4 Matoré (1968), p. 50.

5 See detailed discussion in Part II, Chapter 11.

6 Eight years later Estienne produced the first bilingual dictionary based on the vernacular as SL; his Dictionnaire francois-latin, autrement dit les Mots français avec les manières d'user d'icieux, tournés en latin (1539) is also notable for marking the first occurrence in French of the word "dictionnaire."

7 Anderson (1971: 37) quotes the example of Cordouannier followed by Cordouannerie, and Crocheteur followed by Crocheter, although these may be examples of an early attempt at arranging derivatives in logically dependent order.

8 The order AVARIE, AUBADE, AUJOURD'HUI, AVIRON, mentioned by Anderson (1971), p. 103, exemplifies this inconsistency.

9 For example: Dictionnaire des synonymes et des antonymes (1975), Fides, Montreal (edited by Hector Dupuis and Romain Légaré); Gasc (1876, 1925).

10 Cf. Lexis, Dictionnaire de la langue française (1975), Larousse, Paris, (edited by Jean Dubois); Dictionnaire Beauchemin Canadien (1968), Librairie Beauchemin, Montreal; Bellows (1873, 1951).
His practice of capitalizing the initial letter of these "derivatives," however, perpetuates the difficulty referred to above.


See the discussion on this subject in Chapter 7 below and in Part II, Chapter 11.


See further examples quoted in Anderson (1971), p. 93.

Cf. CD1, CASS, HAP3.


For further discussion see Part II, Chapter 11.

See the detailed discussion on this topic in Part II, Chapter 15.

Plan of a Dictionary (1747), p. 22.

Ibid., p. 27.

Ibid., pp. 118-182.

For detailed discussion see Chapter 11 in Part II.

Boyer (1699), adapted from the entry "to move," quoted in Anderson (1972), p. 80.

Ibid., "Au favorable Lecteur François," reverse of dedication page.

Ibid., unpaginated preface.

CD1, for example, claims the pressing need, in an officially bilingual country, for "an authoritative dictionary"; J-P. Vinay et al. (1962), p. vii. In the
first edition of HAP3, Mansion's preface opens with remarks on the lack of a good bilingual dictionary; while acknowledging the contributions of "Elwall, Bellows, Boiëlle, Tarver and others," he justifies his work by adding that "it is nonetheless true that no French and English dictionary hitherto available has proved adequate to the needs of the serious student"; Mansion (1934), p. v.

30 Mansion (1940), p. v.; there is also an appeal to readers to provide authors with instances of errors.

31 "Separate entries for feminine forms of adjectives and verb conjugations have also been largely suppressed, as it is considered the users of Harraps New Standard French and English Dictionary are sufficiently familiar with them to know without hesitation which headword to consult"; Ledésert (1972), p. vi.

32 Gasc (1876), p. 3, pp. 3-6.


34 See Part II, Chapter 16.
References for Chapter 2


CHAPTER 3

DESCRIPTIVE ANALYSIS

1.0 **Typological Classification**

Dictionaries may be classified in a number of ways. Y. Malkiel has attempted a typological classification of general dictionaries based on distinctive features, by which is meant the decomposition of dictionaries into a number of basic characteristics, the various possible combinations of which are consequently ordered and classified in much the same way as the distinctive features of phonological analysis from which the procedure is borrowed. Yet it will readily be appreciated that the degree of overlap which such a system implies would lead to a classification in which almost every dictionary would belong to a unique category, for each lexicographer emphasizes different aspects and is motivated by his own particular goals. The exercise of establishing a typology of this kind is not, however, altogether without merit: it enables us to pinpoint some of the essential attributes of dictionaries, to explain why some are neglected at the expense of others and it provides us with a useful tool for the analysis of bilingual lexicography.
1.1 Al-Kasimi (1977) summarized various attempts at typological classification of dictionaries and concluded that their main weakness was that they are global (relate to all dictionaries and not to a given class, such as bilingual dictionaries) and have "minimum usefulness or productivity." He proposed, instead, a typology based on "source," "scope" and "purpose," which are precisely the parameters employed by those he criticized. The only difference is that his classification relates to bilingual dictionaries only and not to all dictionaries. In his "new typology" he proposes seven "contrasts": the intended users; literary versus spoken language; production as opposed to comprehension; dictionaries for human, as opposed to machine, processing; historical/descriptive; lexical/encyclopedic and general/special. Although the typology is intended to cover bilingual dictionaries, the last three contrasts, as he himself suggests, are rather more appropriate to monolingual dictionaries, and in general the classification adds little, if anything, to existing typologies.

1.2 In the discussion which follows there is no attempt to establish a "new typology"; instead, the important descriptive features of bilingual dictionaries are analysed in order to demonstrate to lexicographers those elements which must be carefully considered before the
compilation of a dictionary is undertaken. The analysis thus forms the basis for everything which is dealt with in succeeding chapters.

2.0 **Aims and Objectives**

In terms of their physical properties, monolingual and multilingual dictionaries followed parallel lines of development from the end of the Middle Ages. We have seen that, in the case of French and English, historical evolution spawned a tradition of monolingual Latin lexicography, and that it was only after a period of one-way, and eventually two-way, bilingual dictionaries that monolingual vocabularies were produced in the vernacular. But the aims and objectives of both forms were clearly differentiated at an early stage. Monolingual dictionaries were, from the outset, oriented towards a reading public already well-versed in the language under consideration; their aims were therefore conditioned by the need to clarify and exemplify rare and technical vocabulary, and to provide some kind of norm for general vocabulary as a guide to their users. They were, in effect, merely recording, classifying and elucidating the lexicon on the basis of documentary evidence. Bilingual dictionaries, on the other hand, having broached the field of living languages found themselves addressing a different kind of public, with different objectives and
using different procedures. One of the basic properties separating the two forms, for example, is the content of the word list itself; from the earliest stages, the bilingual lexicons concentrated on current everyday language and there was no hesitation about introducing neologisms which were rejected by their more normative counterparts, for this kind of dictionary has always stressed practical vocabulary, the kind with which the user is likely to be confronted in a bilingual situation. In this regard, claims B. Quemada (1968:41), multilingual dictionaries are a more reliable indicator of contemporary usage, in the synchronic sense, than their diachronically-oriented monolingual counterparts which generally record linguistic phenomena with a time lag of several decades. This statement is to be taken advisedly, however, for as G. Matoré (1968:58) points out, the mother language in early dictionaries was only employed as a function of the opposite language; in most instances, the foreign words were selected on the basis of the difficulty which they posed for translation into the mother tongue. For dictionaries of the modern period (post 17th century), however, Quemada's claim is probably valid.

2.1 The general aims of bilingual dictionaries have undergone very little change from those of the first
medieval vocabulary lists which matched Latin and a modern language. If the claim made by Zgusta (1971:294) that "the basic purpose of a bilingual dictionary is to coordinate with the lexical units of one language those lexical units of another language which are equivalent in their lexical meaning" is correct, then the aims have not changed at all, only their orientation. For, as Zgusta goes on to point out, the difficulty in coordinating the lexical units varies in proportion to the differences between the languages involved; these differences may be anisomorphic in nature, that is, related to cultural diversity, or they may be purely linguistic. The earliest dictionaries related simple lexical items in two languages on a one-to-one basis, without attempting to disambiguate homonyms or suggest synonymic development.

2.2 But the requirements which early lexicographers set themselves were somewhat different from those of later dictionary-compilers. Their task was merely to elucidate rare or difficult Latin terminology by applying an appropriate gloss in the vernacular. In this respect the first bilingual dictionaries were consulted for much the same reasons as monolingual dictionaries, namely, in order to find the meaning of a new or unusual word. When Latin was eventually replaced, as the source language, by a modern foreign language the basic purpose of the
bilingual dictionary remained largely unchanged. The earliest French-to-English dictionaries were compiled with the express purpose of interpreting legal documents written in French without actually translating them. The statutes of the English parliament continued to be written in French right up to the reign of Henry VIII, even though the 1362 Act of Parliament established English as the working language of the law courts. These early bilingual dictionaries were therefore simply a variety of monolingual work in which the "semic equation" was balanced in two languages instead of one. Whereas the source language (SL) and target language (TL) of a monolingual dictionary are one and the same, their bilingual predecessors replaced SL by a foreign language.

2.3 The goals of bilingual dictionaries altered with the conditions which gave rise to them. Once the need to understand and interpret the foreign language text was developed into a requirement that it be actually translated in its entirety into the mother tongue, the compilers of bilingual dictionaries were bound to consider all words which might be encountered in any given context. From a restricted list of "hard words," the bilingual lexicon thus emerged as a prototype of modern extended word banks. At the same time, the monodirectional dictionary was born. Yet, from William
Caxton's pre-dictionary of 1480 it was almost 200 years before a two-way dictionary was produced when Guy Miege published his New Dictionary, French and English, with another English and French (1677). Social and political conditions, as we have seen, were directly responsible for the phenomenon of bidirectional dictionaries, and hence for a new goal in bilingual lexicography.

2.4 At this point the goals of monolingual and bilingual lexicography can be said to have followed quite different paths. For while the definitions and synonyms provided in the former are not necessarily items which can replace the entry word in equivalent contexts, the TL equivalents of a bilingual dictionary are supposed to do just that. The first bilingual lexicographers, lacking the benefit of modern linguistic research, proceeded as if Gusta's statement (above) were literally true and attempted to match lexical units of language A with lexical units of language B on the assumption that the equivalents in both languages covered more or less identical semantic fields. But, as numerous investigators have pointed out, the real world is viewed and interpreted differently by different cultures, and bilingual lexicography has to take this into account when suggesting translation equivalents. This point of view has been put into practice, with varying degrees of success, by most modern
bilingual dictionaries and constitutes yet another aim, namely, the delineation of the limits of semantic equivalence by the use of discriminatory labels.

Any teacher of second languages is familiar with the frustrating results of failure to comply with this aim. Dictionaries, especially the pocket-sized editions (often the most available from an economic point of view), which merely list a series of unmarked possible translation equivalents, invite confusion and despair. Even languages as closely related as French and English are poles apart in the degree to which semantic fields overlap (as we shall see in Chapter 5), and the structure of dictionary articles must reflect this phenomenon.

2.5 A final goal, which is common to all modern dictionaries, but of particular importance to bilingual lexicography, is ease of consultation. As the content of dictionary articles increased in size and complexity, so the requirement of a carefully ordered internal structure became indispensable. It is an unfortunate fact of bilingual lexicography, however, that this requirement is seldom met. Space considerations and the concommitant financial aspects of commercially-oriented lexicographical ventures have more often than not dictated that clarity in setting out should be sacrificed to the more pressing need to confine as much material as possible
within the smallest possible page space. The result of this procedure is evident to any translator who has expended precious minutes scanning the longer entries of large dictionaries such as Harrop's Standard (HAP3). One recent dictionary which has broken new ground in this regard is Crowell's Portuguese-English Dictionary, in which the format of entries is entirely governed by user convenience. Long, rambling articles, usually compiled by analogy with monolingual historical dictionary entries, have been replaced by a logical arrangement in which every syntagmatic example is located on an indented separate line, alphabetically ordered and in boldface type. Although publication costs might preclude this system for many dictionaries, it is clear that more attention could be given to the internal structure of bilingual articles.

3.0 Choice of Languages

The next dimension of a typology of bilingual lexicography will clearly involve a statement on what languages are used. As early as 1751, d'Alembert's Encyclopédie makes this point in the article "dictionnaire"; his division of bilingual dictionaries into those treating a dead language and those treating modern languages still holds good. Indeed, it is used by Quemada (1968:38) in the most recent and comprehensive
study to date of French lexicography. In a sense, however, all dictionaries are bilingual in that they offer alternative "codes" for the entry word. Thus, for example, if the left-hand side of a dictionary presents dialect forms, the lexicon of an older period of the language in question, or technical vocabulary, argot, and so on, while the right hand side furnishes appropriate equivalents or definitions in the standard modern form of the same language, a bilingual situation exists. The point is very clearly emphasized by the title of a recent French dictionary, *Dictionnaire français-français des mots rares ou précieux*, and by the author's prefatory remarks: "La richesse même de notre langue fait qu'un grand nombre de mots français nous sont pratiquement inconnus."

3.1 Clearly, however, a distinction has to be drawn between what we are accustomed to defining as monolingual dictionaries and those which are bilingual dictionaries in the true sense of the word. We may therefore employ T. Knudsen & A. Sommerfelt's definition of monolingual dictionaries as "dictionaries in which the words and the definitions of them belong to the same language," where "language" is taken to mean the system of communication employed within a given speech community and not comprehended by the members of any other separate speech community. Multilingual dictionaries, therefore, are
defined as "dictionaries in which the entries (in SL) are expressed in terms of another language (TL)." Bilingual dictionaries, according to Quemada's description, constitute a sub-category of multilingual dictionaries which are also subdivided into "dictionnaires hétéroglosses" and "dictionnaires homoglosses," the latter class including those which treat ancient varieties of a language and its regional forms.

3.2 Restricting our analysis to bilingual dictionaries which, in our definition, belong to the class of "dictionnaires hétéroglosses," it will be seen that the category can be subjected to further subdivision on the basis of the kinds of language treated. The first division would be that which concerns dictionaries treating a dead language and a living language and dictionaries treating two modern languages. Other divisions possible are (a) between two European languages, (b) between a European and an "exotic" language, (c) between two "exotic" languages and (d) between a natural language and an artificial language. These categories could be further classified in terms of more precisely defined language family groupings and most, if not all, are analysable according to the various categories which are the subject of the discussion which follows and which bears on a particular class of
languages, namely, French and English. The schematic representation in Figure 1 summarizes the typology to this point.

Figure 1. Primary subdivisions of general dictionaries

4.0 Directional Criteria

An important consideration which enables a distinction to be drawn between different kinds of bilingual dictionary is the direction of translation. From the historical discussion in Chapter 1, it is clear that the first dictionaries involving French and English
were single-volume works proceeding from the foreign to the native language. The terms "monodirectional" and "unidirectional" used by a number of scholars to denote this kind of dictionary are somewhat equivocal. Some scholars employ them in the sense just mentioned, namely, to indicate that the dictionary concerned provides one series of translations, A → B, and not the inverse, B → A. Dictionaries which provide both series are, for these investigators, "bidirectional" or "two-way." Other scholars use the same terms to specify dictionaries which are structured for "the use of native speakers of only one of the two languages" (monodirectional), or for "the use of native speakers of both languages" (bidirectional). For still others, the term "bipartite" signifies either A → B and B → A dictionaries, or dictionaries for native speakers of both languages.

4.1 This ambiguity could be resolved by agreeing that, since most bilingual dictionaries consist of two parts, it is a redundancy to refer to them as "bidirectional"; the term could then be employed in Steiner's sense. In order to dispel any confusion which exists because of the conflicting terminology, this study will employ the term "bidirectional" (or "bipartite") for dictionaries addressing speakers of both languages, while the term "monodirectional" will refer to dictionaries which
address speakers of only one of the languages treated. Unless otherwise stated, we assume the bilingual dictionaries under discussion consist of two parts, a French-English and an English-French, that is, that they are two-way dictionaries.

4.2 The directional parameter, therefore, is an important classificatory feature and clearly the lexicographer has this feature uppermost in his mind when he sets out to construct his dictionary. Mansion, for example, refers to the need to determine "what classes of readers and students we proposed to serve" by the work. There is, in other words, a pre-determined flow of information, dictated by the prospective users for whom the dictionary is destined and expressed by the explication of one language in terms of the other. The schema referred to in Figure 1 above may therefore be further refined, as in Figure 2, to include the various possible combinations: one-way and monodirectional, one-way and bipartite, two-way and monodirectional or two-way and bipartite.

4.3 The classification to which a bilingual dictionary belongs may sometimes, but by no means always, be found expressed directly by the author in his preface. In the section on phonetics in HAP3, for example, Mansion states that "For the benefit of those users of the dictionary to
whom English is not the mother tongue, the pronunciation has been indicated. From this we infer that the dictionary was compiled for users of both languages and is therefore bidirectional. The subtitle of the English-French section of OXF clearly announces that it is "An aid to the speaking and writing of French specially adapted for English-speaking users," leading us to assume that this is a monodirectional work. In the preface to CD1, we are informed that the dictionary is a translation tool specially geared to speakers of the two official languages of an officially bilingual country. Yet the majority of dictionaries give no information at all about the intended users of the work and the question of directionality must be determined by inspection; this
is the case with CASS2, the various Larousse dictionaries, BELL, and GASC.

4.4 In spite of the claims made by some authors for the directional features of their dictionaries, it is usually more reliable to inspect the body of the work itself in order to determine the manner in which entries are presented to various kinds of user. A number of simple tests will provide the necessary information. For example, an A → B type dictionary which provided no phonetic information, would presumably be intended for native speakers of A who would have no need of pronunciation aids. If the same dictionary, in its B → A section, furnished phonetic indications with the B entries, this would confirm the fact that the work was conceived for native speakers of A, for whom the phonetic transcription of B would be a useful, even necessary, adjunct.25 But the treatment of pronunciation is by no means a sure indicator of directionality. In the preface section on pronunciation in his dictionary, J. E. Bellows claims that the "phonetic spelling (thus) shown after a French word is, of course, intended for the English user, and vice versa."26 An indicator which is probably more reliable than phonetic information is explanatory material: a dictionary A → B and B → A in which all discriminatory labels and rubrics appear in language A would clearly be
compiled for speakers of that language and would be classified as monodirectional. Examples of dictionaries exhibiting this directional feature are GASC, CASS2. The following entries from CASS2 serve to illustrate the point:

*surcharger*  v.t.  to overprint (a stamp)
*overprint*  v.t.  Surcharger (a stamp)

5.0  **Relational Criteria**

In two-way bilingual dictionaries, it is to be expected that a relationship will exist between the two parts. Such a statement may appear, at first sight, to be axiomatic; but on closer examination it is apparent that this is not necessarily the case. There are some dictionaries, for example, in which the two sections were compiled independently by different authors or teams of editors, resulting in discrepancies in methodology and bulk, and a general impression of imbalance. Even when both sections have been compiled by the same author, there may be inconsistencies due to the emphasis placed on translations in one direction as opposed to the other, or as a result of the need to depend on native speakers of the other language for material in the section translating into that language. Thus, although Mansion proclaims the bipartite nature of his dictionary, the physical dimensions alone of the English-French volume
suggest that this is not entirely accurate.

5.1 But far more important than this physical relationship is the semantic interdependence of the two parts of a bilingual dictionary. This is a measurable dimension which is determined by inspection of equivalents. For example, if a word $x$ of the source language A has a translation $y$ in the target language B, then in the other section where B becomes the source language, the word $y$ should be found as an entry, with $x$ one of its equivalents. This is an essential relationship which serves as a bond between the two parts of the dictionary; it has the immediate effect of ensuring a reasonable structural balance and offering, within the physical limits of the work, a degree of semantic closure.

5.2 On the pedagogical level, it ensures that the user has the opportunity to verify the fitness of an equivalent he has chosen by reference to the additional information which can be gleaned from the corresponding entry of the other section. This is not to suggest that every word offered as a translation of entries in language A should necessarily appear in the opposite section as an entry of language B. Geometric harmony is not entirely feasible, for there will be circumstances which dictate otherwise. For example, a dictionary which offers translations of Canadian-French forms in the
French-English section, may not see fit to offer these same forms as equivalents in the English-French section.  

The author's decision with regard to such words is often dictated by the climate of opinion in linguistic regulatory circles, rather than by any personal value judgements. By the same token, certain lexical units are without equivalence in the opposite language; for example, the word *capésien* as a translation of a non-existent concept in English.

5.3 For the large majority of entries in the dictionary however, the relational characteristic is valid and is an important classificatory feature. Its usefulness as a means of maintaining equilibrium is not diminished by the foregoing remarks, because elements of this nature, especially those involving lacunae, tend to be statistically evened out in both languages.

6.0 Classification by Range

This criterion is employed by Malkiel (1967) and by S. Katre (1965) to denote the volume and spread of the lexicographical material. It is measured in two dimensions which reflect the relative density of the entries: (a) the breadth of coverage, which is a count of the number of individual lexical units; (b) the depth of coverage, or the degree of semantic development for each lexical unit. Although it should, in theory, be
possible to compare the number of entries in a given
dictionary with a calculated total of the possible
lexical items available in the language, in practice it
is impossible. To begin with, there is no general
agreement on what constitutes a lexical unit. Some
dictionaries insist on single-word entries, admitting
only those compounds which are hyphenated or otherwise
totally lexicalized; others admit as entries all words or
groups of words which relate to a conceptual unit.

6.1 If the computation of the real total number of
entries is in doubt, the estimation of the theoretical
total is even more so since its very boundaries have yet
to be defined. For example, is such an ideally
complete lexicon to include archaic forms as well as all
neologisms, regionalisms as well as the standard
language, the full range of registers, foreign words and
expressions, technical vocabulary and jargon? The
lexical boundaries are somewhat fluid and open to varying
interpretations. However, most lexicographers base their
word-lists on statistical frequency, whether derived from
actual word-count studies or based on intuition and
experience; the breadth of coverage is therefore a para-
meter which varies with the definition of frequency. For
some, this will admit only those items which have a
statistically determined frequency of occurrence greater
than 5, and would generate only basic vocabularies, such as the *Français fondamental*,\textsuperscript{34} the *Frequency Dictionary of French Words*,\textsuperscript{35} or the *Semantic Frequency List*.\textsuperscript{36} At the opposite end of the scale, the *Trésor de la langue française* bases its word list on a concordance generated from some 70 million occurrences; the large majority of the entries therefore belong to the low frequency area of the statistical distribution.\textsuperscript{37} Recent studies on lexicometry by Sayard and Richards (1970) have revealed other dimensions which may be brought into play in the process of vocabulary selection: they are termed "indices of coverage" and relate to the ability of a word to replace other words in a given context (inclusion); its tendency to form compounds (combination); its capacity to act as a defining element, that is, as a basic semic unit in dictionary definitions (definition); and the degree to which it forms extended meanings (extension).

6.2 With respect to bilingual dictionaries, the range of coverage varies from the few thousand entries of a pocket-sized edition to the more than one hundred thousand entries of the two-volume *Harraps Standard French-English Dictionary*. All are, in a sense, scaled-down versions of an ideal and as yet undefined total lexicon. Within certain limits, therefore, it is possible to assess the breadth of coverage but only in a fairly vague, if not
arbitrary, fashion.

6.3 The depth of coverage is equally difficult to determine in a strictly objective and empirical manner. The number of senses provided for a given word varies from one monolingual dictionary to another and bilingual lexicographers are inconsistent in their adaptation of the semantic structure. The traditional approach is to duplicate the historically-ordered arrangement of the former, although some steps have been taken towards a logical article structure motivated by user convenience.\(^{38}\) It is possible to evaluate the depth of coverage according to number of meanings by inspection, but the only clearly-defined boundaries will be those separating the succinct and often excessively truncated material of pocket editions from the meandering, encyclopedic entries of more important dictionaries; the same grey area of uncertain distinction lies between the two extremes and defies accurate assessment.

6.4 The minimum requirement for depth of coverage is that all the basic definitions of a word be accounted for, but the entries of a bilingual dictionary must proceed a good deal further and show the correspondences which obtain in certain syntagmatic combinations. For example, the fact that an element \(x\) of language A has definitions \(x_1, x_2, x_3, \ldots\), which are rendered in
language B by \( y_1, y_2, y_3 \ldots \), does not mean that in all situations \( x_2 \) will correspond automatically to \( y_2 \). There is a limit to the coverage which can be given to these extensions and it is clearly a measurable variable in different dictionaries.

7.0 **Lexical Concentration**

Closely allied to the notion of depth of coverage is the dimension of lexical concentration—that is, the degree to which information that is not, strictly speaking, lexical material is incorporated within the body of the work. For monolingual dictionaries, it can be argued that such elements as proper names, historical or biographical vignettes and the like properly belong in an encyclopedia, the task of the dictionary being to define the meanings of words rather than to elaborate on their social or historical associations. Conversely, the task of a bilingual dictionary being to offer equivalent forms in another language, it is evident that some material of this kind does qualify for inclusion. It is important, for example, that the user be informed that the equivalent in French of **Hugoesque** is **hugolien**, that the **Thames** is **la Tamise**, that **William** is equivalent to **Guillaume**, that **Nepalese** is rendered in French by **Népalais** when referring to a person, by **népalais** when used adjectivally and by **népali** (if the language is
meant, and so on. In many instances, the word itself may have no intrinsic value as a translatable unit and yet occur in a number of important derivational forms. No one would have any reason to seek a translation of the proper noun Parkinson, and yet the medical possessor of this name is responsible for an extensive morphological series which requires treatment:

(1) **maladie de Parkinson** (unaltered noun) - Parkinson's disease
(2) **maladie parkinsonienne** (derivational adjective) - Parkinson's disease, Parkinsonia
(3) **parkinsonien, -ne** (derivational noun) - male/female afflicted with Parkinson's disease
(4) **parkinsonisme** (derivational noun) - Parkinsonism
(5) **parkinsonia** (derivational noun) - Parkinsonia.

The word Stalin has proved no less productive in both languages:

(1) **Staline** (proper noun) Stalin
(2) **stalinien, -ne** (derivational adjective) Stalinist
(3) **stalinien** (derivational noun) Stalinist
(4) **stalinisé, -e** (derivational adjective) Stalinized
(5) **stalinisme** (derivational noun) Stalinism.

7.1 Toponymic information may be restricted to those items which have a different form in the two languages. and, in the case of French, those which possess a gender
which is likely to create problems for the English user. Patronymic material may likewise be treated in this way. Other kinds of information that may be found in bilingual dictionaries include equivalent weights and measures, acts of parliament, book titles, honorific titles and terms of address, government agencies, international organizations, common abbreviations, verb tables, grammatical and phonetic indications, regional variations, proverbs, and so on. Without specifying whether he is referring to monolingual or polyglot dictionaries, Malkiel (1967: 14-15) labels the incorporation of such data as an "infiltration" and a "dilution" of normal lexicographical requirements. Yet if the aims of the "ideal" bilingual dictionary of Haas (1967) are to be met, that is, a dictionary which "would anticipate every conceivable need of the prospective user," then these supposedly non-lexical items may rather be considered to be minimum requirements.

7.2 The physical location of the material is a moot point. The tenor of Malkiel's argument suggests that he objects to the relegation of encyclopedic information to supplements, while Haas considers its placement in separate sections as "wasteful of the time and energy of the user." Yet the lexical status of individual items should first be considered. It is obvious that the words
derived from *Voltaire* should appear in the body of the dictionary in their normal order since they constitute clearly lexicalized elements. On the other hand, proper name equivalences of the type *William/Guillaume* may be better situated in a separate listing since it is not always in the best interests of translation that such words should be transposed.

7.3 Grammatical information is normally subsumed under the entry to which it belongs. English verb parts are usually simple enough to appear as run-ons to the entry, but their French counterparts are so complex as to warrant separate treatment. Phonological information usually takes the form of phonetic or quasi-phonemic transcriptions accompanying the entry, but occasionally more detailed information is incorporated in the preface to be used in conjunction with the transcriptions. Abbreviations which have become lexicalized: *capésien*, *onusien*, *NDPer*, may appear in the body of the dictionary, while their more synthetic counterparts remain in a separate listing.

7.4 There are various permutations and combinations of the material which is included under the heading of lexical concentration. The manner in which it is treated constitutes a feature enabling further typological classification.
8.0 **Structural Criteria**

Dictionaries may be categorized according to their physical structure, the two most obvious features being the form and sequence of the entry words themselves and the arrangement of the information contained in the entry. The canonical form of the entry word is fairly standardized—infinitive for verbs, masculine with run-ons for adjectives, singular with run-ons for nouns, and so on. Variations are observed in the quantity of information given with the entry; for example, some dictionaries present the infinitive only of the verb and assume the reader's ability to generate the principal parts. This is normal procedure for monodirectional works, where the native speaker of one of the languages is presumed to have no need for morphological props in that language; but the same dictionary may well eliminate run-on material from the foreign-language section as well. 44 The majority of French and English dictionaries present tables for French irregular verbs as a separate section of the preface. 45 One dictionary has practised sound pedagogical philosophy by a triple reinforcement process: the principal parts of all irregular verbs occur as run-ons to the entry; they also occur as separate entries with a cross-reference to the full entry; and every verb, regular or otherwise, has a number attached to it which refers to the verb table in the
preface where the appropriate conjugation form can be obtained. 46

Ideally, a bidirectional dictionary will also provide run-on material in the English section but few do so in practice, 47 assuming, perhaps accurately, that the French user is better able to cope with the vagaries of English conjugations than his Anglophone counterpart is with French. Modern lexicographical practice seems to be calling more and more for a relaxation in the rules governing the status of entry words. At one time the lexicographical "word" was a single graphic form; items consisting of two or more graphic elements were normally treated as sub-entries or examples under the heading of one of the elements. The notion of "word = conceptual unit" is, however, gaining acceptance, especially in monolingual dictionaries and in some bilingual works whose editors point to equivalence features to justify the bestowal of entry status on certain lexical groups. 48

8.2 The physical appearance of the entry word also distinguishes different kinds of dictionaries. All use some procedure for making the word stand out from the body of the material—usually boldface type, but sometimes colours, indentations, vertical lines, and various symbols. The irritating habit, followed by some editors, of starting each entry with a capital letter, 49 or
placing the whole entry in capitals, has mercifully become unfashionable, one hopes because it has been amply demonstrated to be a contributing factor in linguistic confusion and error. Alphabetical ordering is obviously the most commonly used system for arranging words in a form which is suitable for reasonably quick consultation. Yet it is a purely conventional device based on the arbitrary disposition of the letters of the alphabet and not necessarily superior to any other classification which might suggest itself, such as the ordering by phonetic form used in rhyming and inverse dictionaries. The organization of word lists according to centres of interest is the basis for the so-called analogical or ideological dictionaries and, although some restricted vocabulary lists designed for pedagogical purposes are occasionally classified in this way, all major bilingual French and English dictionaries can be assumed to belong to the alphabetical category. There are, however, subdivisions within the latter: the order may be strictly alphabetical, with no deviations permitted under any circumstances, or it may be quasi-alphabetical, with groupings of related words attached as sub-entries to a base word chosen as the main entry.

8.3 The following would be a representative word order for these kinds of dictionary:
### 8.4 Occasional rupture of pure alphabetical order is not necessarily a hindrance to the user, as the last example demonstrates; the search for expressions involving *devant* is more likely than not to lead to the article, and success can be assured by the use of cross-references in other locations. The same dictionary, for example, gives as entries the principal parts of irregular verbs.

<table>
<thead>
<tr>
<th>HNS: 54</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>double</td>
<td>adj</td>
<td></td>
</tr>
<tr>
<td>doublé</td>
<td>adj</td>
<td></td>
</tr>
<tr>
<td>doublé</td>
<td>s.m.</td>
<td></td>
</tr>
<tr>
<td>double-as</td>
<td>s.m.</td>
<td></td>
</tr>
<tr>
<td>double-bang</td>
<td>s.m.</td>
<td></td>
</tr>
<tr>
<td>double-bacassine</td>
<td>s.f.</td>
<td></td>
</tr>
<tr>
<td>double</td>
<td></td>
<td></td>
</tr>
<tr>
<td>double-face</td>
<td>s.f.</td>
<td></td>
</tr>
<tr>
<td>doublement¹</td>
<td>adv</td>
<td></td>
</tr>
<tr>
<td>doublement²</td>
<td>s.m.</td>
<td></td>
</tr>
<tr>
<td>double-fond</td>
<td>s.m.</td>
<td></td>
</tr>
<tr>
<td>double</td>
<td></td>
<td></td>
</tr>
<tr>
<td>double-quatre</td>
<td>s.m.</td>
<td></td>
</tr>
<tr>
<td>doubler</td>
<td>v.tr.</td>
<td></td>
</tr>
<tr>
<td>double-scull</td>
<td>s.m.</td>
<td></td>
</tr>
<tr>
<td>double</td>
<td></td>
<td></td>
</tr>
<tr>
<td>double-un</td>
<td>s.m.</td>
<td></td>
</tr>
<tr>
<td>doubleur</td>
<td>s.m.</td>
<td></td>
</tr>
<tr>
<td>doubleuse</td>
<td>s.f.</td>
<td></td>
</tr>
<tr>
<td>double-zéro</td>
<td>s.m.</td>
<td></td>
</tr>
<tr>
<td>CD1: 55</td>
<td></td>
<td></td>
</tr>
<tr>
<td>devant</td>
<td>prép.</td>
<td></td>
</tr>
<tr>
<td>devant</td>
<td>n.m.</td>
<td></td>
</tr>
<tr>
<td>au-devant</td>
<td>adv.</td>
<td></td>
</tr>
<tr>
<td>au-devant de</td>
<td>loc.</td>
<td>prép.</td>
</tr>
<tr>
<td>de devant</td>
<td>loc.</td>
<td>prép.</td>
</tr>
<tr>
<td>par devant</td>
<td>loc,</td>
<td>prép.</td>
</tr>
<tr>
<td>par devant</td>
<td>adv.</td>
<td></td>
</tr>
<tr>
<td>devanture</td>
<td>n.f.</td>
<td></td>
</tr>
</tbody>
</table>

---

8.4 Occasional rupture of pure alphabetical order is not necessarily a hindrance to the user, as the last example demonstrates; the search for expressions involving *devant* is more likely than not to lead to the article, and success can be assured by the use of cross-references in other locations. The same dictionary, for example, gives as entries the principal parts of irregular verbs,
supplying a cross-reference to the infinitive form under which the form will be found. For example:

pu, pus, put  cf. POUVOIR.
pus'  cf. POUVOIR.

8.5 Unfortunately, the zeal of some lexicographers has persuaded them to push the "word family" concept to the limit, resulting in fractured word order and reduced ease of consultation; the system presupposes a reader with a fairly sophisticated knowledge of etymological relationships. Thus, for example, the following word order may be found in BORDAS. 56

| têde | adj. |
| tièdement | adv. |
| tiédasse | adj. |
| tiédeur | n.f. |
| tiédissement | n.m. |
| tiédir | v.tr. |
| attiédir | v.tr. |
| s'attiédir | v.pr. |
| attiédi | adj. |
| attiédissement | n.m. |

Cross references are usually supplied for words which are placed far from their normal strictly alphabetical position; but in the case of attiédi,
attiédir, s'attiédir and attiédissement, no such indication is given and only a reasonably knowledgeable user or mature native speaker would know where to find the words.

8.6 The internal structure of bilingual dictionary articles may likewise be used as a classificatory feature. Some lexicographers imitate monolingual dictionaries and the sense order most likely to be found is historical or etymological combined with frequency of use. In this case, the article material will be organized into subdivisions corresponding to those of a monolingual work. For example, the article crosse in HNS demonstrates the following internal arrangement:

1. [Eccl] stick, (bishop's) crook.
2. [Sport] stick, club (for games).
3. crook, bent portion (of anything).
5. [Culin] knuckle.
8. [Techn] (other technical uses and expressions).

The same article in the monolingual French dictionary Le Petit Robert (PROB) orders the senses thus:

1. [Eccl] bâton pastoral (d'évêque).
2. [Sport] baton recourbé utilisé dans certains jeux; 
3. bout recourbé.
5. [Culin] crosse de bœuf.
7. [Arme] partie postérieure d'une arme à feu.
8. (other expressions).

For the same entry, other more concise editions of bilingual dictionaries either use a reduced version of the same order or else appear to have no systematized arrangement at all, as the following examples demonstrate (the numbering refers to the numbered senses of HNS):

CD1: 1. 2. 7. 
OXF: 1. 3. 2. 7. 3. 4
GASC: 3. 1. 7. 2. 8. 4
LAR. MOD: 1. 2. 8. 4. 7. 6. 8.
CASS: 1. 7. 2. 6
LAR PRAT: 1. 7. 6. 2
BELL: 1. 2. 7.

Other internal arrangements are possible. The senses may be ordered on a logical basis, developing from the concrete towards the abstract; they may be listed conventionally, according to the alphabetical or other order of the subject fields of the meanings; finally,
they may be completely unordered, or only partially and inconsistently ordered. 60

9.0 Discriminatory Features

An extremely important attribute of bilingual dictionaries is the use of discriminatory labels to reduce equivalence ambiguities. 61 Productions with limited budgets which are forced to cut corners by reducing space consumption are unlikely to be able to comply with this requirement; the price of inadequate treatment, in this regard, is a state of utter confusion such as that which reigns between the covers of pocket editions. The untrained user who consults the paperback version of Larousse's French-English English-French Dictionary (LARFL) will not be very much enlightened, for example, by being informed that the word mess may be rendered in French by any of the following:

plat; mess; ordinaire; ration; popote; sachis; désordre. 62

If he happens to select plat, reference to the French-English section will inform him that this word means "dish." But if his choice is popote he will search in vain for information in the other section where it does not appear.

9.1 Most reputable bilingual dictionaries offer adequate
meaning discrimination in the form of rubrics denoting subject matter—[Philos], [Math], [Numis], etc.; explanatory rubrics—[person], [fête], etc.; exemplifying material—to stock up on (flour, oil, etc.); and language level labels—[Fig], [Pej], [Pop], etc. A few dictionaries go a step further and provide warnings of deceptive cognates (faux amis).

9.2 The use of discriminatory labels, as we have observed above (Chapter 2), arose from the practice of early lexicographers of including definitions of entries before their translations. Gradually these definitions gave way to briefer particularizing words and finally they were reduced to the basic elements of a semic analysis which assures the process of disambiguation. In monodirectional dictionaries these labels appear in the language of the speech community for which the work is destined; in bidirectional dictionaries, since it is obviously impractical to provide labels in both languages, they should appear in SL.

The extent to which disambiguating material is incorporated in a dictionary is a measure of its utility. Without it, the phenomenon of differential lexicology constitutes a major stumbling-block in the translation process and the dictionary user is left to consult monolingual dictionaries in order to make the correct
selection from an array of unspecified equivalents such as those illustrated above.

10.0 Conclusion

On the basis of the foregoing discussion, several selected 20th-century dictionaries have been submitted to an analysis in order to determine those features of a proposed model which they display. In the matrix in Table I a plus sign indicates that the dictionary possesses the feature so marked, a negative sign that it lacks it and a blank that no serious judgment could be made. It will be observed that some dictionaries in this schema are positively marked for what might be considered pairs of mutually exclusive features, such as monodirectional/bidirectional. This is to be taken as implying that the work concerned has not treated the feature in an exhaustive fashion, that it sets out, for example, to be bipartite but lapses into monodirectionality—a situation which was illustrated in 4.0 above. No two dictionaries depicted in the matrix share the same overall features, which effectively emphasizes the claim made in 3.0 with regard to the impossibility of categorizing dictionaries according to a feature matrix.

Some of the features, as we have seen, involve an element of choice on the part of the lexicographer; he can decide in advance, for example, the kind of audience
<table>
<thead>
<tr>
<th>Feature Matrix for Bilingual Lexicography</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Table I</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Monodirectional</strong></td>
</tr>
<tr>
<td>------------------------------------------</td>
</tr>
<tr>
<td>HNS</td>
</tr>
<tr>
<td>HAP3</td>
</tr>
<tr>
<td>CASS2</td>
</tr>
<tr>
<td>CD1</td>
</tr>
<tr>
<td>LAR1</td>
</tr>
<tr>
<td>LAR2</td>
</tr>
<tr>
<td>DPL</td>
</tr>
<tr>
<td>GASC</td>
</tr>
<tr>
<td>BELL</td>
</tr>
<tr>
<td>OXF</td>
</tr>
</tbody>
</table>
he wishes to address, the manner in which he will order the entries; the internal structure of the entries and so on. Some choices will leave the way open to various options while others may restrict the author to a particular line of development. These selectional criteria suggest the development of a model (Table II) in which the "ideal" bilingual dictionary is depicted in terms of a "flow-chart" which illustrates the options available to the author and consequently the features which a dictionary governed by such options will possess.

In these first three chapters we have restricted our attention to the descriptive features of bilingual dictionaries, tracing the developmental factors which have contributed to contemporary lexicographical methodology and analysing the latter in order to expose to view the elements which are of importance to this study. These findings will now serve as a frame of reference for a theoretical analysis of the foundations of bilingual lexicography.
TABLE II
SCHEMATIZATION OF THE DEVELOPMENT OF A BILINGUAL DICTIONARY

LANGUAGES

monodirectional ← DIRECTIONALITY → bidirectional

SL speakers TL speakers

SL labels TL labels

LEXICAL RECIPROCITY

LEXICAL CONCENTRATION

DISCRIMINATION

didactic ← RANGE → technical

frequency counts
concise edition

utility
extensive edition

ENTRY STRUCTURE (external)
quasi-alphabetical
x-referencing
etymological

ENTRY STRUCTURE (internal)
statistical logical

Historical
Notes to Chapter 3

1 Y. Malkiel (1967) and (1959).
3 Ibid., p. 20.
4 Ibid., p. 20.
7 J. Darbelnet (1971), p. 92; J-P. Vinay and J. Darbelnet (1958); see also the discussion in Chapters 4, 5 and 6 below.
8 However, computer-based projects in which printing is accomplished by computer-generated photocomposition should achieve a degree of economy which would enable otherwise costly typographical layout to be more fully investigated.
9 This question is dealt with in some detail in Chapter 11 of Part II.
12 Avant-propos, p. 5.
14 Quemada (1968), p. 38.
15 J. D. Anderson (1972), pp. 5, 7; G. E. Roffe (1955) employs the terms "one-way" and "two-way" in this sense, pp. 2-7.
16 R. J. Steiner (1970), in a glossary of "special terms" used in his study, pp. 13-14.
It might be profitable to borrow, from mathematical or scientific terminology, a number of appropriate terms to emphasize the distinction. Thus: "bicentric," "bifocal," "bifurcating," "bigeneric," "bilateral," "bilinear," "binary," "bipolar," "bivalent," could all be adapted for the definition.

Preface to HAP3, p. v.

Preface to HAP3, p. viii.

However, certain features in the body of the work tend to refute the accuracy of this description.

Preface to the English-French section, p. i.

Preface, pp. vii-viii.

This is the case for OXF and confirms the authors' claim that it is a dictionary for "English-speaking users"; preface, p. iii.


OXF is an example of this kind of "fractured" dictionary in which each part was compiled in vacuo by independent teams of editors.

HAP3, preface p. viii.

It contains 1,539 pages as opposed to 997 pages in the French-English volume.

This question is examined in detail in Part II, Chapter 15: "The Treatment of Regionalisms."

See detailed discussion in Part II, Chapter 13.


Malkiel (1967) identifies the features which hinder its definition as the "essential instability of its lexicon" and the "incompatible scopes of even the best lexicographical projects so far launched, as a result of sharply divergent objectives and techniques"; p. 7-8.

Based on the work of A. Gougenheim (1964) et al.

36 H. Eaton (1940).

37 Cf. preface to Dictionnaire de fréquences (TLF); also Ch. Muller: "Le Trésor de la langue Française et la statistique lexicale." Travaux de linguistique et de littérature, 11 (1973), pp. 86 ff.

38 Cf. Part II, Chapter 11 on "article structure."

39 See below, Chapters 4, 5 and 6; also Part II, Chapter 13.

40 Haas (1967), pp. 46, however, takes to task the editors of OXF for including, in the French-English section, the words fauteuil-voltaire, voltarianisme and voltarien, -ne while omitting the proper name from which they are derived. On the other hand, if dictionaries are to include all proper nouns, the majority of which are identical in both languages, large quantities of redundant material will be produced; there comes a point, quite clearly, when the user must be credited with the basic intelligence to make some decisions for himself. Haas also overlooks the fact that the derivation of voltaire, fauteuil-voltaire is, in fact, given in the body of that entry (OXF, p. 877).

41 Haas (1967), pp. 45-46.


43 Haas (1967), p. 46.

44 This is true of OXF, which gives no principal parts in the body of the dictionary although it supplies all feminines and plurals of adjectives even when regularly derived, as most are.

45 Notable exceptions to this are HAP3 and HNS which instead give all the principal parts with the entry; cf. the entry chœur in HNS.

46 CD1 and its revised version, CD2. The same numbering convention (to direct the reader to appropriate verb tables) is also employed by Crowell's Spanish-English and English-Spanish Dictionary, G. A. Gilhoff (ed.), New York, Thomas Crowell, 1967.

47 Cf. CD1, CASS, HNS. No verb principal parts whatever are supplied in OXF and DPL. The LAR2 provides them only in the English section, relying on verb tables for French verb parts.
See the full discussion on this question in Chapters 11 and 12 of Part II.

Cf. GASC.

Cf. BELL.


Cf. HNS, DPL, CASS2, CD2.

Examples are CD1, GASC, BELL, OXF, LAR1.


CD1, entries on pp. 123-4.


See detailed discussion in Part II, Chapter 11.

Cf. CROW.

Also used by CROW.

Cf. OXF.

The question is studied in depth in Chapter 13 of Part II.


See the detailed analysis in Chapter 14 of Part II.

Discussed in detail below, Chapters 4 and 5.
References for Chapter 3


CHAPTER 4

TRANSLATION THEORY AND BILINGUAL LEXICOGRAPHY

1.0 It has been assumed as axiomatic, so far in this discussion, that the transfer of meaning from one language to another is a physical possibility which is confirmed by the existence of bilingual dictionaries, translators and equivalent texts in pairs of languages. Yet even the reality of the latter is not incontrovertible proof of the theoretical possibility of translation. In fact, the overwhelming weight of linguistic opinion, based on modern theories of language, leans heavily in favour of a position which refutes the possibility of translation; everything which has emerged from semantic studies, from research in syntax and morphology, seems to deny translation a viable position in a general theory of language. ¹

Until comparatively recently translators were content to translate, and readers to digest the outcome of their activity, with little attempt to justify on theoretical grounds what was happening. Lexicographers, likewise, have practised their "art" (or science) for centuries without showing the least concern for the
theoretical basis for their endeavour. Nevertheless, translation and bilingual lexicography have progressed, if not flourished, over many centuries and this leads to the somewhat self-evident conclusion that the theory, even if it has not been openly expressed, has always been potentially available and intuitively sensed by practitioners. It remained for linguists like E. O. Nida, R. Catford and G. Mounin to articulate the framework of a theory for translation which would answer the many questions posed by their colleagues and by the disparagers of the theoretical possibility of translation. But in spite of all the activity generated by renewed interest in the theory of translation, bilingual lexicography continued to find itself relegated to the position of a sub-category of translation, or as one practical manifestation of the wider science of lexicology. Although a few isolated endeavours have been made to improve both its status within the total discipline of linguistics, as well as its craftsmanship, no serious attempt has yet been made to formulate a theory which would ensure these goals and at the same time account for the unique character of an occupation which had been practised for many centuries before linguistics itself was identified as a science. In this chapter we will attempt to isolate those features of bilingual lexicography which distinguish it from the
related activity of translation and which justify its existence as an autonomous field of study.

1.1 Bilingual lexicography and translation appear, on the surface, to be intimately related, even interdependent, operations. In terms of a theory, it could even be argued that the former is simply a small part of the latter and that the theory of translation supports the physical functioning of translation as well as accounting for the existence of bilingual dictionaries. On the other hand, there is some justification for the proposition that translation is the physical output of a general theory of bilingual lexicology, that it is an activity which is only practicable because we are able, *grosso modo*, to effect the matching of pairs of lexical units from different languages (lexicography) and to extend the phenomena of equivalence from the level of discrete "words" to the level of phrase, sentence and text (translation and interpretation).

1.2 Upon closer inspection, however, the supposed similarities between bilingual lexicography and translation can be largely discounted. In the first place, the physical activities involved are totally different; the former is essentially a process in which the inventory of lexical items from one language is coordinated with equivalent elements from the lexical inventory of another
language. But these lexical items are treated in isolation, deprived of the semantic and syntactic environment in which they operate and of the associative network to which they belong; the lexicographer selects from the environment only those elements which are essential to the equivalence relation which he is attempting to establish. It is thus basically a classifying and recording activity which calls upon the experience of the lexicographer and the intuitions of bilingual individuals as well as monolingual native informants for each language.

1.3 Conversely, translation takes as a point of departure an already existent text in one language, that is, an acceptable combination of its lexical elements comprising a message which is to be first interpreted by a mediator (in the person of the translator) and then reproduced as an identical message in the other language. The translator reads and interprets the message and when he fully comprehends its meaning proceeds to create a new text comprised of the symbols of a different language, so arranged that the import of the original remains unaltered. We may therefore conceive of translation as a process which combines comprehension, interpretation and linguistic creativity. But there are a number of variables which come into play during the
process of translation. The nature of recorded or spoken information as presented to the translator is such that affective, emotive or stylistic values inherent in the message influence different people in different ways, with the result that the same original text is differently interpreted by translators from diversified backgrounds and of unequal dispositions; the outcome may well be a message which varies considerably from the intention of the original.

2.0 No such variables are evident in lexicography which is concerned only with words as discrete units and with their meanings and translational equivalents. However, as Saussure has pointed out, the "word" has no independent existence; its semantic function is determined by the total network of relations that it may enter into with other lexical elements. These relations, Catford maintains, may be formal, that is, related to purely physical structure (morphological similarities, functional activity and so on), or contextual, that is, dependent on the situations in which they occur (in utterances, texts, etc.). The same relations are equally valid for bilingual lexicography which uses them to establish total meaning and equivalence. But whereas lexicography derives the equivalent pairs from an analysis of their occurrences in various contexts,
usually by way of distributional analysis (whether consciously applied or not), translation creates a context in which the lexical elements of the target language may be arranged in such a way that they constitute the same message as that expressed by the elements of the source language.

2.1 The formal relations between lexical units are an important part of bilingual lexicography, even though the main concern is semantic equivalence. Thus, for example, phonetic information may be provided for pedagogical motives which dictate that the non-native user needs assistance with this structural aspect of language transfer; and other formal relations between words are used in identifying traditional lexicographical features, such as the canonical form of the entry word, morphological variants and functional grammatical categories. Similarly, syntactic information is employed to exemplify some of the potential strings of the language; these are the practical realizations, regrettably provided in small quantities, of otherwise abstract entities. In cases where structural syntactic equivalence is diversified, this kind of information is crucial although it cannot be provided in anything like an exhaustive manner because dictionaries are limited, both by space considerations and by their immediate
goals, to exemplifying only a few of the possible realizations in context of the lexical units of which they are composed. Bilingual lexicography is thus concerned with contextual relations to the extent that they are a source for determining intrinsic meaning or are necessary to exemplify differential equivalence between pairs of strings with the same overall meaning. 6

2.2 Translation is obviously concerned with the same formal and contextual relations, but to a different degree and in totally different ways. The translator is always presented with a completed text, composed (or spoken) by a native speaker and comprising various formal and contextual relations which are imposed by the semantic and structural features of the language in question. Translation thus presupposes a text and a mediator endowed with the competence (in the linguistic sense) to comprehend and interpret the message contained in the text, to evoke equivalent forms at the level of lexical unit, syntagm, sentence and ultimately paragraph or text, and to arrange them in the manner appropriate to the structure of the target language. Without elaborating on the problems of arriving at an objective assessment of the success of the translation (faithfulness of the transferred message, effect on the reader/listener, stylistic features and so on), it is clear that
the process implies, theoretically at least, a mediator equipped with the knowledge of the total structure of both languages (more particularly of the target language). Moreover, if oral interpretation is involved, this knowledge must include phonological structure. We are therefore forced to conclude that a person possessing all the required attributes will be a very rare phenomenon and must be considered as a theoretical ideal. 7

2.3 A text in a given language is composed of real elements of the language, combined according to rules which determine the meaning conveyed by the text; this meaning depends, in turn, on various contextual factors. In other words, a translation to be carried out on such a text is operating at the level of *discours*, whereas the discrete elements of the original text, as well as its translation, when taken in isolation (as in a dictionary), belong to *langue*. This is why access to a bilingual dictionary and a knowledge of the syntax of a pair of languages are alone insufficient to accomplish an acceptable translation—the dictionary deals in abstractions while translation is, by definition, a practical and concrete activity. 9 In the course of this practical activity, lexical structure is exposed to all the influences which may be brought to bear during the
communication act:

... dans l'exercice concret du langage, la sémantique lexicale dépend non seulement de la combinatoire élémentaire des signes, de la nature des sens fondamentaux, "nucléaires," mais de toute la situation de communication, de l'énonciation (conditions de production du texte, personnalité de l'auteur, intentions du message), des conditions de l'énoncé (parlé ou écrit, spontané ou non, destiné à un lecteur, à un auditeur ou à plusieurs ...), du contexte linguistique, des circonstances de l'échange d'information, etc. Chaque mot et chaque expression, outre son signifié moyen, apporte ainsi sa physionomie unique (pour celui qui l'emploie et pour celui qui doit le lire) ou particulière (pour le groupe de ceux qui participent du même univers mental et affectif). 10

Clearly none of these factors enter into lexicography which is necessarily a totally objective activity in which equivalence relations must be established in the absence of the extralinguistic influences which are an essential part of parole.

2.4 Many of the difficulties encountered by lexicographers in compiling bilingual dictionaries are not necessarily problems for translators because the latter are not dealing with words in isolation, nor with abstractions, but with context and reality. 11 The totality of the text to be translated is taken into consideration, and the determination of the choice of a word may be influenced by contextual features which are
physically quite remote from the particular element under examination. While the individual lexical unit retains its abstract values (semantic, quantitative and so on), which in effect were first derived from textual occurrences, there are situational modifications brought about by its association with other lexical items in a given context—modifications which very often cannot be predicted or accounted for within the confines of a dictionary.

Thus, the translation of a text may capture the message and stylistic intent of the original in its totality and yet, if a comparative analysis of the discrete elements of the texts is undertaken, the word pairs may appear to be only partial equivalents which would never occur in this form in a conventional bilingual dictionary. In theory, the equivalents are selected on the basis of their semic affiliation which, as we have seen, is an abstract quality not necessarily realizable in concrete situations. When all the contributory parameters are brought to bear—disparate lexical structure, syntactic disjunction, anisomorphism, cultural disparity and so on—the real potential output of pairs of lexical units which are proposed in the dictionary as equivalents may have the effect of driving them even further apart.
It would appear that there are various degrees of equivalence. We may speak of global equivalence between a text and its rendering in another language when, to all intents and purposes, the totality of the message of the original has been conveyed to a reader/listener of another language and provokes in him the same reactions as those which would be aroused in a reader/listener of the language of the original. This, presumably, is the general aim of translation. Within the same hypothetical text we are discussing there will be large units of equivalence corresponding to sentences and even whole paragraphs which might be referred to as extended equivalence and which are still the chief concern of translation. When we examine the smaller clusters of lexical units, subgroups of the sentence or clause, we are dealing with syntagmatic equivalence which may take the form of phrasal groupings in normal distributionally determined arrangements, the predictable elements of language, or else set phrases (locutions figées) which function as though they themselves were lexical units. These elements may be composed of anything from two to a dozen "words": là douloureuse/the bad news; parlour games/jeux de société; se mettre sur son trente-et-un/to get all decked out. They are an important part of bilingual lexicography as well as of translation, although the former usually restricts itself to exemplify-
ing the less predictable strings belonging to this category. Finally, in the overall spectrum of equivalence there are the individual lexical units or "words," the headwords of a bilingual dictionary, and the raison d'être of lexicography; the relationship between the lexical units of a pair of languages is lexical equivalence and, as we have seen, it means one thing in a dictionary and another in the act of translation.

3.0 The foregoing discussion leaves little doubt that the practical activities of translation and bilingual lexicography are quite distinct and independent, although both are related to theoretical considerations which have much in common: Translation theory has been the object of a great deal of research which we will now examine briefly in order to determine to what extent lexicography, which remains a largely untouched area of investigation, can learn from it.

3.1 G. Mounin (1963) depicts translation as a phenomenon of language contact which means that it is therefore a special case of bilingualism. However, he does not view the role of translation theory as shedding light on various problems of general linguistics, but rather the reverse: the solutions for the problems of translation theory as a whole should be sought in the findings of modern linguistic theory, especially structural and
functional linguistics. Nevertheless, his examination of the latter only seems to provide him with further evidence against the theoretical possibility of translation and leads him to conclude that there can be only two acceptable positions. Either (a) we accept in toto current linguistic theory and relegate translation to the status of a non-scientific adjunct to the study of language, thereby confirming earlier extremist views on translation. Or (b), on the grounds that translation is not a physical impossibility, we may question the validity of linguistic theory in those areas that conflict with translation. Eventually Mounin opts for a more moderate position, however, and concludes, after a careful examination of both points of view, that contemporary linguistics would define translation as a process which achieves varying degrees of success in terms of the act of translating.

3.2 R. Catford (1965) likewise drew upon general linguistic theory in attempting to formulate a linguistic theory of translation, which he regards as a branch of comparative linguistics since it is concerned with relations between languages. According to his broadly-defined conception, translation is "the replacement of the textual material in one language (SL) by equivalent textual material in another language (TL)."
His emphasis on the **textual** aspect of translation sets it apart from lexicography. On the other hand, the term **equivalent** is considered by Catford to be central to the whole question of translation and he deals with it in some depth, describing it as an "empirical phenomenon," and also as a "functional process"; clearly this falls within the domain of bilingual lexicography. He draws a distinction between **textual equivalence**, which is any portion of a TL text considered to be the equivalent of a SL text under certain conditions and verifiable by commutation, and **formal correspondence**, which applies to approximate equivalence at the level of grammatical categories.

3.3 E. A. Nida (1964) advances on a fairly broad front, incorporating with linguistic theory various metalinguistic sciences such as anthropology, sociolinguistics, psychology, ethnology and so on. He maintains that the basic problem of translation is the "fundamental fact that languages differ radically one from another." But the power of this axiom, long upheld by the disparagers of translation theory, has recently been considerably moderated by linguists and anthropologists alike who are now generally agreed that the similarities in man's existence far outweigh the divergences, thus paving the way for a "basis for communication." Adopting a
moderate course between the extremist views of translation as an art and as a science. Nida attempts to bring the findings of linguistic research to bear on theoretical considerations with particular emphasis on the problem of meaning. Like Catford, he analyses the principles of correspondence in translation, distinguishing between "formal equivalence," which "focusses attention on the message itself, in both form and content," and "dynamic equivalence," where the translation "aims at complete naturalness of expression." These terms correspond, more or less, to literal and free translation respectively, the former being an integral part of bilingual lexicography.

4.0 Meaning and equivalence thus play a dominant part in translation theory, just as they do in bilingual lexicography. Early semantic theory, which depicted language as a repertoire of words each uniquely corresponding to a thing or a concept, was shaken by Saussure's conception of meaning as an interdependent network of lexical elements; the meaning of a given word, according to Saussure, depends on the totality of its relationships with the other words associated with the reality designated by that word; meaning is determined by an analysis which finds expression in negative terms: it is what all other word meanings are not.
4.1 The notion that meaning is part of a system of lexical interrelation and that, in the linguistic sign, the link between acoustic image and concept is of far greater complexity than had hitherto been suspected, only serves to spread further confusion over the issue of the theoretical status of lexical equivalence and translation. In fact, it was proposed by Mounin that the Saussurean conception of the sign provided a satisfactory answer to the problem of why word-for-word translation is usually totally unsatisfactory; it is the difference in semantic extension which precludes a one-to-one correspondence between the respective signs of two different languages. 29

This disparity is most clearly demonstrated by natural bilinguals who are unable to translate a text word-for-word because the nature of their bilingualism is such that they possess two independent but complete language systems; while they can, with relative facility, transform a message in one language into an equivalent message in the other, they do so by means of paraphrase, that is, they convert entire sense groups and avoid the necessity to match up individual lexical items. 30 The process involves the selection of leading referents in the text (nouns, verbs and so on) and the building up of a restructured context appropriate to the manner of expression in the second (target) language. The reason literal translation is difficult for natural bilinguals is that
"individual semantic items are not commensurate." 31 Also, in the context of phrase, sentence and paragraph, individual lexical units tend to be dissociated from their individual referents and to submit to a more generalized referent indicated by the sense of the whole. This is exactly the kind of meaning which translation has to contend with, whereas lexicographical meaning is restricted to discrete elements of the lexicon which are "frozen," so to speak, at the very moment of their being plucked from the real context which determines their meaning. In bilingual lexicography a referent is either implied or else overtly indicated by various conventional methods (explanatory rubrics, subject labels and so on); every entry word has at least one referent and the translator selects the elements of his translation more or less on the basis of the meanings supplied in the dictionary, although always with the proviso that meaning in context is transmutable.

4.2 Bilingual lexicography is thus concerned with word-for-word equivalence of the very kind that is rejected by translation as unsatisfactory. For example, a lexical element like French canard, a generic term with respect to the various referents which it may denote, is provided in a dictionary with English equivalents which are specific in terms of the sense area covered by the French
word (duck, rag, rumour, false note, sugar lump). In translation a reference to the word canard entails a search of the context in order to determine which of the multiple referents is appropriate, although the operation may not end there because the equivalent given in the dictionary may not be the term called for by the context. This kind of disparity between the way in which referents are classified in different languages is sometimes referred to as "differential lexicology" and is a vital component of bilingual lexicography as well as translation.  

4.2.1 In the word-for-word situation of bilingual lexicography the translation equation may be represented thus:

\[ X \rightarrow S \rightarrow Z \]

where X and Z are the respective equivalent signs in two languages and S their common referent. However, in view of the foregoing remarks, it will be necessary to refine the equation in order to account for the differently organized referential systems of the languages:

\[ X(x_1) \rightarrow S \leftarrow Z(z_1). \]

In this vector diagram we include the independent referent
possessed by each sign and suggest that the direction of the arrows should be inwards, that is, oriented towards the generalized referent which is the means whereby translation becomes possible. Where one of the languages has a conceptual space corresponding to a symbol and its referent possessed by the other language, the classic situation of lacunae brought about by cultural divergences, the vector diagram would revert to its original equation form; in this case the equivalent, if given, would be a definition or explanation of the missing concept or even the word itself borrowed from the other language.

4.3 Although much attention is usually given to the lacunae which seem to set languages apart with respect to their lexical equivalence, the phenomenon is in reality quite limited except where genetic links are remote. Non-equivalence in the symbolization of relations is the rule, however, at the level of text, sentence, phrase and syntagm and this is why word-for-word translation is unacceptable in translation theory and why in a dictionary it is an abstraction which becomes modified in real situations. Thus, while we may say, in a dictionary, aller = to go, the syntagmatic realizations of the verb do not share the simple equivalence implied by the dictionary entry:
il va = he goes, he does go, he is going, he (habitually) goes;
he went = il est allé, il alla;
il est allé = he went, he has gone, he did go, he has been going;
il allait = he went, he used to go, he would go, he kept going.

These structural variations are, of course, an important part of the competence of the translator; dictionaries tend to assume this kind of knowledge and normally go no further than providing the principal parts of the verbs, leaving even the naive learner of the language to generate the forms required by the context in which they are to be employed.

4.3.1 Similarly, there are a number of semantic features which may range over the whole sentence or text and which are an important consideration in translation—negation, interrogation, emphasis, exclamation and so on; also stylistic features—register, variations in sentence structure for literary purposes and other similar devices.

4.3.2 Lexical lacunae are caused by two situations: either the concept does not exist in the reality experienced by the speakers of the language in question and therefore no symbol has been assigned to the conceptual space occupied by a symbol in the system of another language; or conversely, the concept is
discernible or in the process of becoming known to the language group although it has not yet penetrated the linguistic system of the language. Lacunae, in both translation and lexicography, are resolved by paraphrase, but whereas the former must provide a single lexical unit or string in a text, the latter is more flexible and may supply a number of alternatives in the form of synonyms, circumlocutions or definitions which eventually build up a semantic "picture" of the missing concept. Sometimes, also, the lacuna is eliminated by the simple process of adopting the symbol as well as its referent directly from the language which possesses it, or else by creating a calqued version if this is possible. While dictionaries may exercise a degree of this flexibility of definition, translation demands a more positive, ne varietur approach. In fact, dictionary meanings more often than not undergo considerable modification in context where "the words are often not used in their dictionary sense at all, but in the metaphorical, enlarged, pregnant and other off-senses which the system of oppositions and interplay of meanings within the system permits."34

5.0 This brief comparison of the respective features of translation and bilingual lexicography makes it clear that we are in fact dealing with two independent disciplines which are nevertheless linked by their mutual
association with a theoretical base which might be called the theory of **bilingual lexicology**. Such a theory would cover all aspects of translation and bilingual lexicography as well as general semantic theory and the related studies of bilingualism and second language acquisition. In the remaining chapters of this section we will discuss those portions of the theory which are of special significance to bilingual lexicography under the headings of "Semantic Isomorphism," "Lexical Structure," "Grammatical Theory" and "Bilingualism."
Notes to Chapter 4

1 Cf. G. Mounin (1963: 8): "L'activité traduisante pose un problème théorique à la linguistique contemporaine: si l'on accepte les thèses courantes sur la structure des lexiques, des morphologies et des syntaxes, on aboutit à professer que la traduction devrait être impossible. Mais les traducteurs existent, ils produisent, on se sert utilement de leurs productions. On pourrait presque dire que l'existence de la traduction constitue la scandale de la linguistique contemporaine.

2 U. Weinreich, for example, was dismayed that an important dictionary like Webster's Third should have occasioned so little theoretical study: "It is disconcerting that a mountain of lexicographical practice such as an unabridged dictionary of English should yield no more than a paragraph-sized molehill of lexicological theory"; "Webster's Third: A Critique of its Semantics," IJAL, 30 (1964), p. 408.

3 "Translation is a process by which a spoken or written utterance takes place in one language, which is intended and presumed to convey the same meaning as a previously existing utterance in another language"; C. Rabin, "The Linguistics of Translation," Aspects of Translation (Studies in Communication), London: Secker and Warburg, 1958, p. 123.

4 F. de Saussure (1960): "La partie conceptuelle de la valeur (du mot) est constituée uniquement par des rapports et des différences avec les autres termes de la langue"; p. 162.


6 To this extent we can therefore claim that bilingual dictionaries are 'just as concerned with definition and the "semic equation" as their monolingual counterparts; but in the former, the definitions are subsumed under the guise of equivalence relations. In a manner of speaking, therefore, the translational equivalents of a bilingual dictionary constitute another kind of metalanguage definition.

7 "It can even be doubted whether total 100% bilingualism exists; it would mean that a person could use either language in any situation, with the same facility and correctness and with the same proficiency as the native speakers. (And if such cases exist, it is hard to see how they could interest the linguist, because interference phenomena would by definition be excluded.)"; Hans Vogt, "Language Contacts," Word, 10, 2-3 (1954), p. 369.
Rey suggests that, although the first reaction to this situation might be that dictionaries are not well made, the true facts are that translation and dictionaries are quite different: "... c'est que les opérations nécessaires à la traduction se font sur du texte ou sur des paroles: sur du discours ... les dictionnaires ne peuvent tenir compte que d'une infime partie des réalisations concrètes de la langue"; p. 21-2.

"L'opération traductrice est la transformation d'un texte dans une langue en un texte d'une autre langue capable de produire sur un lecteur (ou un auditeur) tous les effets --informatifs, affectifs, esthétiques--du texte de départ. Cette définition correspond à un programme maximum"; ibid., p. 20.

"A textual equivalent is any TL text or portion of text which is observed on a particular occasion ... to be the equivalent of a given SL text or portion of text," (p. 27).

"A formal correspondent is any TL category which may be said to occupy, as nearly as possible, the 'same' place in the economy of the TL as the given SL category occupies in the SL," (p. 32).

Mounin's answer to this old controversy is as follows: "On peut, si l'on y tient, dire que, comme la médecine, la traduction reste un art--mais un art fondé sur la science"; op. cit., p. 16-17. Another linguist gives this explanation: "At least two factors will always inhibit the formulation of any general theory of translation: translation is partly an exercise in the art of writing as well as a field of comparative applied linguistics, and no one can determine where the first ends and the second begins"; Peter Newmark, "An Approach to Translation," Babel, 1 (1973), Vol. 19, p. 13.


Ibid., p. 160.

F. de Saussure (1960).

Ibid.

G. Mounin (1963:27): "les mots n'ont pas forcément les mêmes surfaces conceptuelles dans les langues différentes." This question is studied in more detail in Chapters 5 and 6.


Ibid., p. 135.


This kind of symbolic representation is employed by Rabin (1958: 125).

References for Chapter 4


1.0 The practice of bilingual lexicography reposes on the premise that it is possible to produce an array of equivalent terms corresponding to the wordlist of any given language. If a one-to-one correspondence existed between these equivalent forms in any given pair of languages, the task of the lexicographer would be elementary and the goal of machine translation much nearer than it presently appears to be. The hard facts are, however, that translation equivalence is little more than an approximation in varying degrees of precision depending on the linguistic and cultural contiguity of the two languages. Zgusta (1971) discusses the concept of the "anisomorphism" of languages which refers to "the differences in the organization of designata in the individual languages and ( . . . ) other differences" (p. 294). Linguists and anthropologists have discussed these phenomena for several generations and presumably lexicographers have taken them into account, albeit intuitively, in the production of their dictionaries. Clearly the degree to which any pair of languages can be
said to be semantically isomorphic, that is, to possess convergent semantic structures, will vary considerably. Even closely related languages like French and English are far from being semantically isomorphic in the ultimate sense, although there may be a considerable amount of overlap brought about by close cultural relationships and by a generally equivalent view of the real world. At the opposite end of the scale of semantic isomorphism are language groups for whom the material, extralinguistic world presents an entirely different perspective which is reflected in a high proportion of virtually untranslatable lexical elements.

1.1 Early notions of translation theory included the assumption that all humanity possessed a universal view of the real world and that experience was classified mentally on similar logically-based lines and expressed in similar terms. The philosophical treatises of von Humboldt, pursued and refined by later generations of linguists, reversed this view by postulating that, on the contrary, man's view of the world was fashioned by his language. 1 This is clearly a further impediment to the theoretical possibility of translation, for even if two languages possess their separate labels A and B for an identical concept it is by no means certain that one can therefore make the claim that A = B. No two languages
categorize experience in an identical manner; each divides conceptual space into elements corresponding to the unique view of the world shared by the members of that linguistic community and consequently assigns labels to those elements in a more or less arbitrary manner. The so-called Sapir-Whorf hypothesis claims that the language of a given community is its "world" or its "social reality"; in Sapir's terms, "the worlds in which different societies exist are different worlds and not merely the same world with different labels attached."² Lyons (1969) adds a further degree of refinement to the debate by proposing that "the language of a particular society is an integral part of its culture and (that) the lexical distinctions drawn by each language will tend to reflect the culturally-important features of objects, institutions and activities in the society in which the language operates" (p. 432).

1.2 The fact that different languages employ widely differing linguistic structures to depict the same invariable physical reality is usually cited by proponents of the Humboldtian view as evidence that it is not the physical world which fashions linguistic structure but linguistic structure which determines the way in which the physical world is viewed.³ Another proof is offered by the fact that a given individual may change
his life-style, his experiences and attitudes, but not
the linguistic medium through which he expresses these
phenomena. Mounin (1963:57) maintains that the
structural stability of language, in contrast to the
mutability of human experience, is particularly marked at
the level of the community where, for example, our
intellectual enlightenment does not hinder us from using
such mythologically-based expressions as le soleil se
lève the sun rises or from continuing to classify
(linguistically) chauve-souris with souris even though we
are aware that the physical truth of the former and the
zoological genetics of the latter negate the interpreta-
tions applied by less well-informed ancestors.

1.3 The findings of cultural anthropology, sociology and
ethnology, according to Mounin, also prove that it is not
always the same external world which linguistic structure
is being called upon to analyse. In this case, the task
of expressing the reality of one civilization in terms of
the language of another may well be a physical
impossibility, as illustrated by Nida's example of a
culture such as Maya for whom the concepts denoted by the
terms fig tree and vine are unknown, or the Amazon sub-
tropical tribe for whom the notion of desert is a
complete lacuna. Mounin maintains that most studies on
the subject tend to confuse the two issues, namely that
of the different view of the same reality and that of the
different methods of describing different physical
"worlds." But it is not altogether certain that he is
correct in drawing this distinction; the former is
clearly present in any given pair of languages whereas
the latter is restricted to culturally disparate
languages and not, as he suggests, a feature of even
relatively homogeneous cultures. He commits the same
error for which he censures Vinay and Darbelnet (1958)
when he quotes the existence in French of words such as
yard, gallon, dollar, mark, rouble to prove that
different "cultural worlds" apply also to the European
languages which supplied (and absorbed) these terms. 8
All of these terms actually relate to an identical
reality—the need to measure physical quantities. The
fact that they are divided by different groups into
different units with different labels is merely another
confirmation of the notion that identical physical form
is observed and analysed differently by separate language
groups. Similar remarks apply to Mounin's comments on
cheese and bread terms, 9 for while another European
language may lack appropriate labels for richly
diversified forms of these elements, their existence is
by no means totally inconceivable in the same way that,
for example, the religious notion of consubstantiation
might be for an Australian aborigine.
1.4 In view of this line of thought, it might be better to postulate a sliding scale of discrimination or anisomorphism, ranging from a disparate attitude towards the distribution of physical reality (such as the colour scale or measurements) to the non-existence in one of the cultures of the concept under consideration. Furthermore, by the application of fairly elementary logic to the vexed question of different views of the real world or differences in cultural attitudes, it may be demonstrated that the supposed impossibility of translation in this area may not, after all, be valid.

Although the wealth of terms manifested by a language to portray a given phenomenon seems to set that language apart and point to an unusual approach to the analysis and description of reality, the same kinds of differences can often be observed within the same language. Thus, for example, while the astonishing number of Eskimo terms for "snow" or the complex Hanunóo colour terminology seem to reveal a distinguishing feature of those languages, similar differences in levels of experience of the real world can be detected among different speakers of the same language. In Canada, European language speakers living in northern regions have a much more varied set of expressions for describing "snow" than those who live in southern cities or on the milder south-west coast where there is only the word snow itself plus a few euphemistic
or pejorative terms. As Mounin points out, moreover, skiers employ a wider range of technical "snow" terms than non-skiers. 10

1.5 Field theory in linguistics is of great significance to the question of translatability because it provides probably the most tangible evidence for illustrating the anisomorphism of languages. Linguists and anthropologists have a penchant for using relatively closed series such as colour terminology or military titles and kinship terminology to demonstrate this phenomenon; but even the establishment of highly complex networks of associated lexical items has done little to convince modern linguistics that the answer to the problem of meaning is to be found in semantic fields. Some skepticism is expressed by linguists who do not feel that the lexicon lends itself to dissection and consequent rearrangement into structural models on an analogy with those of morphology and phonology; the argument usually put forward to support this view is the fact that the domains selected for analysis are probably special cases and not therefore valid for the lexicon as a whole.

2.0 English and French would generally be rated as two distinct languages exhibiting a relatively high degree of semantic isomorphism based on their linguistic and cultural proximity. Yet it is not difficult to find
examples which illustrate clearly that even these two languages are heterogeneous and pose a significant number of translation problems. They, too, possess large numbers of "culture-bound" words and expressions for which the other language can offer, at best, only a definition or an approximation by way of circumlocution. In the area of culinary arts, for example, there are differences of extraordinary complexity. The same meat carcass viewed by French and English butchers would be sliced up and named in two totally different manners resulting in an almost impossible situation for the translation of meat cuts.

2.1 According to the classical Saussurian definition, the "word" links a concept and an acoustic image. This conception of the linguistic sign, while emphasizing the autonomy of the word and of the concept which it designates, at the same time sheds light on the phenomenon of anisomorphism which sets different linguistic communities apart from each other and adds a measure of difficulty to the translation process. For, as Hjelmslev pointed out, the semantic structure of a given language is not determined by appealing to a physical description of things, of which each community has its own conception, but rather by studying the kind of structure which links signifiants, the linguistic portion of the sign, and
signifiés, the conceptual portion. Vinay and Darbelnet (1958: 28) illustrated this relationship by pointing to cases where a single signifié is served by different signifiants in two languages: "knife" and "couteau" in the context *table knife/couteau de table*, and others where the signifiants appear, superficially at least, to be interchangeable whereas their signifiés do not coincide: "bread" and "pain"; English bread is neither similar in appearance nor nearly as important nutritionally and culturally as French bread.

2.2 At still another level of disparity there are numbers of words and expressions in one language which denote a concept totally unknown in the other. The French educational title *C.A.P.E.S.* (= *certificat d'aptitude pédagogique pour l'enseignement secondaire*) has no exact equivalent in English and the situation is further complicated by the fact that the term has spawned a number of even less translatable derivatives: *capésien, capésienne, capès*. Sometimes a language will adopt a missing concept and its label from another language and incorporate the word so completely into its own structural system that it acquires features unknown in the donor language. Thus the American English word *drugstore* was taken unchanged by French as a masculine noun and eventually treated morphologically to create the
forms drugstorien and drugstorienne. But more often than not the lacunae remain and translation is relegated to the level of approximation or definition.

3.0 The concept of semantic isomorphism, which is obviously of great significance to bilingual lexicography, requires a clear understanding of what constitutes the anatomic structure of the element under discussion which, for the sake of convenience, we will refer to as the "word." In Saussurian terms, the linguistic sign is viewed as an indivisible unit linking a concept with its linguistic realisation, that is, the conjunction of signifié and signifiant represented schematically thus:

![Diagram of linguistic sign](image)

The diametrically opposite vertical arrows of this conception illustrate the interaction of the two basic elements of the sign along the language-thought continuum;
the diagram also emphasizes the respective autonomies of the word and of the concept which it designates. From the point of view of interaction between languages, it would also be appropriate to modify the model to include a horizontal plane, as suggested by Vinay and Darbelnet (1958: 29), which would reflect the transmission of information from one sign to another.

3.1 Since lexicography concerns itself with the discrete elements of language, the signs, it is appropriate to study their theoretical structure in isolation in an attempt to understand the varying degrees of semantic isomorphism which exist between pairs of languages. The linguistic portion of the sign, from a lexicographical point of view, is composed of certain formal features: the visual image is represented by the canonical form of the lexical unit appearing as a dictionary entry; the acoustic image is also usually represented in phonetic transcription or phonemic notation; there are usually in addition a number of ancilliary formal characteristics such as grammatical irregularities, internal hierarchical relations and so on. Complementing the formal characteristics of the sign are its functional features: syntactic structures, distributional characteristics, levels of usage, domain of extension, etc. Finally, there are quantitative features, expressed by frequency
distribution and general statistical structure, which lexicographers, knowingly or otherwise, include in their dictionary articles.

3.2 Semantic theory has built upon Saussure's model by postulating that the relationship between signifiant and signifié is the only characteristic of immediate interest to linguistics; considerations relating to the problems of conceptualization, auditory perception, affective impressions and so on are considered to be extra-linguistic. If the function of the signifiant, therefore, is to identify and particularize the concept which it is said to represent, then this is at the heart of lexicographical methodology which attempts to ensure the efficacy of the bond between the two elements of the linguistic sign, whether it be by means of definition and example (as in monolingual dictionaries) or through the provision of translation equivalents (in bilingual dictionaries).

3.3 Under ideal conditions exact equivalence can occur between pairs of signs only if there is a single signifié shared by two signifiants; this is precisely the same set of conditions as that which applies to synonymy within languages, although perfect synonomy is something of a rarity. In practice, precise equivalence between languages occurs only in three situations: in scientific
or technical language where the lexical units are usually scrupulously defined in order to avoid any ambiguity, often by international convention and frequently by means of the same terminology or using words derived from the same base; for example, \textit{astronautics/astronautique}.

Secondly, with function words whose static role in the grammar of the language precludes the development of multiple senses usually associated with high frequency words; however, the disparate syntactic features of individual languages more often than not restrict this kind of equivalence to the presence of the words in their simplest relationships: \textit{avant/before, après/after}, but d'après/according to and avant tout/above all. Thirdly, there are certain pairs of low-frequency words whose uncommon use denies them the possibility of extending their sphere of meaning: \textit{duenna/duègne, hearsé/corbillard}.

3.4 The so-called "semiotic triangle" can be adapted to apply to the phenomena of simple equivalence, the monovalent cases, in the following manner:
The diagram should be conceived of as a three-dimensional structure (triangular pyramid) in which the shaded surface represents one language (say French) and the dotted surface another (English). We assume a common significé located at the apex and a common referent at one corner of the base; the usual broken lines indicate the indirect relationship which is supposed to exist between the referent and its symbolic representation, while the unbroken lines link the other elements of the linguistic sign. The heavy line joining signifiant and signifier suggests that the relationship may be either direct (etymological link, borrowing) or indirect (non-related signs).

3.5 As we have seen, however, such perfect semantic isomorphism is the exception rather than the rule and the model will therefore have to account for the more usual phenomenon of non-overlapping semantic extension; in other words, those cases where one of a pair of words possesses meanings not shared by the other. If each triangular pyramid of the schema depicts a complex consisting of a shared referent and concept, then a symbol which also relates to other referents and concepts will require linkages which might be shown thus:
In this diagram a word a in language A has equivalents x, y and z in language B; these equivalents are not synonyms because they do not share the same referent and concept.

3.6 The schematization is necessarily complex because it reflects a state of affairs which is far from simple. However, in the interests of clarity, and since the pyramidal structure is repetitive and always constant, it would be appropriate to represent each nucleus of meaning equivalence by a simple structure which will enable a clearer description of various kinds of multiple relation. In the diagrams which follow the clear circles represent a lexical unit of language A and the solid circles correspond to various equivalent forms in language B:

\[ a = x \text{ (1:1 correspondence)} \quad a = x, y, z, \ldots \quad \text{(synonymic equivalence)} \]
4.0 At this point it would be appropriate to apply the terminology of Greimas (1966) in order to clarify the true nature of the various structures depicted in these diagrams. Greimas starts from de Saussure's depiction of semantic structure (the "word" exists only by virtue of its relationship with other words in the system and is defined in terms of its relations with and differences from other words) and describes the various kinds of association between lexical elements as follows. There are relations de conjonction (in which words share some common feature) or relations de disjonction (where it is the differences, rather than similarities, which distinguish words). At the formal level (signifiants) relations are based on phonemic and distributional oppositions: back/pack (/bæk/-/pæk/); to bend one's back, but not *to bend one's pack; and so on. At the level of the signifiés relations are based on similarities
and semantic oppositions:mètre/litre are both measurements but are applied to different kinds of matter.

4.1 According to this analysis, if two terms are linked at the signifiant level by disjonction and at the signifié level by conjonction we have an example of synonymy. Between languages, where the same kind of relationship exists, we would refer to it as equivalence. The opposite relationship, conjunction of the acoustic and graphic symbols and disjunction of their signifiés, describes homonymy. But since conjunction can be of two kinds (acoustic and graphic) we must distinguish between two kinds of homonymy: homophony and homography. In the theory of bilingual lexicography the first of these subcategories is irrelevant because groups of homophones in one language (sceau, seau, sot, saut, etc.) are linked to equivalent forms in the other language which do not share this feature (seal, bucket, stupid, jump, etc.). Homography, on the other hand, is a vital part of all lexicography.

4.2 The traditional approach to homography in lexicography is to employ exponents to differentiate between lexical units of identical form but different meaning: bit¹, bit², bit³, etc. However, the distinction between three or four entries labelled "bit" which are homographs and a single entry "bit" containing several clearly
differentiated senses is far from clear, nor is the methodology applied to them anything like consistent.\textsuperscript{12} To a very large extent the decision is based on the subjective intuition and experience of the lexicographer; the original etymological relationship which may once have linked some or all of a group of lexical units has often become so vague with respect to their modern definitions that they are no longer conceived of by native speakers as polysemous elements of a single word. By contrast, a pair of homographs not originally related may have developed senses which are so alike that they are considered to be examples of polysemy and listed under a single entry. Z gusta (1971:77) quotes the case of French errer = "to be mistaken" (< Latin errare) and errer = "to wander" (< Latin itinerare); English ear = "hearing organ" (< Old English ēare) and ear = "head of corn" (< Old English ēar). In some dictionaries these pairs are listed as independent entries while in others they appear as separate senses under a single headword. The latter case, in fact, reflects the way in which the words are perceived by all native speakers but the small minority who are aware of their etymological roots.

4.3 Another sub-category of homonymy is needed to account for groups of homographs, frequently numbering up to four or five, which are simply categorical variants of
the same word: short (adj), short (adv), short (n) and short (v). Such words are usually distinguished by their differing syntagmatic and paradigmatic realizations and constitute examples of "partial homonymy" which must be carefully defined, in bilingual lexicography, for the sake of non-native speakers.

4.4 If homonymy, then, is described as the conjunction of signifiants and the disjunction of signifiés, we may well ask how this differs from polysemy which commands the same description. The distinction between the two is not altogether clear and is not enhanced by the sometimes inconsistent methodology of lexicographers. There is even one school of thought which suggests that every distinct meaning of a "word" requires that it be listed as a separate lexical unit, thereby throwing into even greater confusion the long-standing debate on the autonomy of the "word." Since what we refer to as "polysemy" is a highly frequent feature of semantics, as a glance at any dictionary will reveal, this approach would result in an exaggerated proliferation of the number of dictionary entries, with groups of homographs numbered up to as many as 50 or 60. The notion, logical as it may appear to be, runs counter to the natural tendency of the system towards economy and a fine balance between redundancy and entropy in the interests of
effective communication. Moreover, there is a close relationship between the polysemic subdivisions and their separation into discrete units cannot be justified even on purely linguistic grounds. The dividing line between homonymy and polysemy can be clearly determined by etymology and therefore the theoretical basis for the proposed division of "words" by sense is untenable.

4.5 Strictly speaking, and certainly from a lexicographical point of view, it might be preferable to consider that homonymy is not a part of meaning and that the dictionary-maker's practice of separating homographs by exponent numbers is really another way of saying that these are completely disparate lexical units. The fact that an accident of historical development has resulted in identical forms must be treated as the practical manifestation and not allowed to cloud the theoretical issue which calls for such forms to be treated as separate and unrelated elements of the system. In his analysis of the limits of translatability, Catford (1965: 93-103) refers to homonymy as "shared exponent" which may occur at the level of morphology, perhaps the only true homonymy (-s as plural morpheme as well as verbal paradigm morpheme), as well as in the lexicon. He points out that native speakers of English intuitively regard such pairs as bank (financial institution) and
bank (edge of a river) as separate items and not merely two meanings of a single item. To what extent this "intuition" is conditioned by external factors, such as educational texts and lexicographical methodology, is not entirely clear but the phenomenon does lend weight to our contention that homonymy is not a semantic feature but simply a practical problem which lexicographers are called upon to resolve.

4.7 It is, of course, possible to have homonymy between languages. This almost invariably manifests itself as homography since the phonological systems of any given pair of languages rarely coincide exactly enough to produce homophones. There may be a kind of partial homophony as in the case of a bilingual person with an imperfect control of the phonetic realizations of the second language which prevents him from discriminating between similar-sounding word pairs: pole/pôle [poul/poːl]; broom/brume [bruːm/bRym] and so on. The more frequent phenomenon, homography, may be of various kinds: (a) true cognates, which may be of identical or nearly identical form, where conjunction of form is matched by conjunction of meaning; (b) deceptive cognates, where conjunction of form is paralleled by partial or complete disjunction of meaning: English versatile (= skilled in many things; fickle) and French versatile (= fickle);
(c) unrelated pairs whose similarity is purely fortuitous: English fat (= plump) and French fat (= conceited). 13

5.0 We may summarize the foregoing analysis with respect to a theory of bilingual lexicography in the following way. In translation, the unit of semantic correspondence between two languages is called equivalence which, in its simplest terms, is the matching of lexical units from one language with equivalent forms in the other language. These matching pairs (or multiples) ideally share a common referent and signifié but differ in their signifiants. Under ideal conditions the theory would require that every pair of equivalent terms would be of this kind, that is, that perfect one-to-one correspondence would exist between them; but in practice perfect equivalence is very rare and restricted to clearly defined technical terminology and infrequent words. Using the same notation as that described above (3.6), the various kinds of equivalence can be schematized thus:

(a) Simple one-to-one correspondence.

5.1 This basic correspondence can be extended to account for the theoretical case of a group of synonyms in one language which are exactly equivalent to a group of
synonyms in another language:

(b) Complex one-to-one correspondence.

Complex one-to-one correspondence is thus described as a nuclear cell composed of several signs each of which is directly linked to every other sign in both languages. Again, this represents a theoretical ideal which is rarely, if ever, realized in practice.

5.2 The situation in which a single lexical unit of language A corresponds to several lexical units in language B could be described as "inter-language synonymy." Under perfect conditions this kind of equivalence would link a signifiant and one signifié of a language to several signifiants and the same signifié in another:

(c) Simple multiple correspondence.
5.3 Finally, when a single lexical unit in language A is equivalent to several lexical units in language B, each of which has an entirely independent meaning, we have an example of complex multiple equivalence in which there is no conjunction of signifiants and multiple conjunction of signifiés for each of the linked pairs of signs:

(d) Complex multiple correspondence.

5.4 Homography is eliminated from this description on the grounds set forth above that it is not part of the semantic theory of translation; the correspondence of signs between languages is the result of fortuitous phonetic development and is merely a "surface" manifestation which in reality corresponds to discrete "deep structure" elements.

5.5 The degree of semantic isomorphism existing between languages is thus a measure of the correspondence of their respective semantic structures, which in turn depends largely upon what Lyons (1969:55) refers to as "the distinctions and equivalences which are of importance in the culture of the society in which the
language operates. " Since the same is true even of dialect groups within the same language we cannot expect to find a high level of semantic isomorphism between languages no matter how much cultural reciprocity is present. This is why the last kind of equivalence, complex multiple correspondence, is by far the most frequently-occurring phenomenon in bilingual lexicography and the one which, as we shall observe in Part II, brings about the greatest difficulties in its methodology.
Notes to Chapter 5

1 In the words of André Martinet (1960): "A chaque langue correspond une organisation particulière des données de l'expérience. . . ." (p. 16); "Une langue est un instrument de communication selon lequel l'expérience humaine s'analyse différemment dans chaque communauté" (p. 25).


4 Harris (1954), p. 151.


7 Ibid., p. 64.

8 Ibid., pp. 64-65.

9 Ibid., pp. 65-66.

10 Ibid., p. 193.

11 "Le chien recevra une description sémantique tout à fait différente chez les Esquimos, où il est surtout un animal de trait, chez les Parsis, où il est animal sacré, dans telle société hindoue, ou il est repris comme paria, et dans nos sociétés occidentales dans lesquelles il est surtout l'animal domestique. . . ."; Louis Hjelmslev, "La Stratification du langage," Word X, 2-3 (1954), p. 176.

12 See the detailed discussion in Chapter 11 of Part II, Section 3.0.

13 This question is analysed in detail in Chapter 14 of Part II.
References for Chapter 5


CHAPTER 6

LEXICAL STRUCTURE AND LEXICOGRAPHY

1.0 Field Theory

The investigation of the structural aspects of the lexicon was initiated by Saussure with his regrettably brief incursion into the realms of "associative networks" in which words are linked to other words of the system in a vast constellation of phonological, morphological and semantic relationships. But whereas the purely formal affiliations between words are easily identifiable and have been systematically studied and classified, the infinitely complex nature of the lexical system has largely escaped global analysis. Structural linguistics has more or less abandoned the attempt to reduce the lexicon to a systematic structure, although many linguists have examined minute portions of it in an effort to uncover some underlying orderly arrangement which would parallel the other structural aspects of language.

1.1 These studies of isolated areas of the lexicon, usually referred to as semantic fields (or lexical fields), have generated intense interest on the part of
linguists although they have done little, if anything, to resolve the problem of global structure. On the other hand, the results of this research are of vital concern to lexicography in its endeavour to perfect the "semic equation" and to incorporate in dictionary articles the features of meaning which are essential to effective definition (in monolingual dictionaries) and translation equivalence (in bilingual dictionaries). G. Mounin (1972) considers that a good deal of progress has been made since A. Meillet rejected the existence of a structured system for the lexicon parallel to those of phonology and syntax: "On a certes beaucoup et bien travaillé à percer le secret de ces petits groupes, mais dans bien des directions, structuration lexicale, sémantique, logique, statistique, distributionnelle, informationnelle, dont on n'aperçoit pas encore vraiment l'unité" (p. 164).

Saussure's approach was taken up by other early twentieth-century scholars who proposed that a number of conceptual domains could be isolated which correspond to a very clearly-defined structural arrangement of lexical units. While this organization may be seen to be the solution to the problem, in the opinion of most modern linguists the method lacks scientific precision since the notion of starting with a concept and building up its associative network is necessarily an individual affair.
related to personal knowledge and experience as well as psychologically-based mental associations. Later investigators approached the question of lexical fields with a more rigorous attitude to its methodology and a studied attempt to contain it within the boundaries of pure linguistics.

1.2 G. Mounin (1963) studied the notion of semantic fields with a view to finding in them solutions to the problems of translation. They provide the most tangible evidence, he claims, for the fact that every separate linguistic system analyses the real world in its own distinct manner—distinct between different languages as well as between different developmental stages of the same language.

1.3 The well-known illustration of languages which present hundreds of individual terms for a given concept (such as "snow" in Eskimo) while lacking a generic term, or of the multiplicity of systems for depicting the colour spectrum (which is demonstrably a universal phenomenon), are usually cited among evidence which emphasizes the differences in lexical structure between languages and the theoretical impossibility of translation. If it were possible to demonstrate that global lexical structure is based on semantic fields of this kind, then it would be clear that each individual lexicon
would be composed of "mosaics" of terms whose sub-
categorizations and surface manifestations would almost
never coincide; it would be possible, further, to
demonstrate that lexical equivalence would be very rare,
even when word pairs belong to the same field in each
language.  

1.4 Fortunately several factors combine to refute this
hypothesis. In the first place, argues Mounin, transla-
tion is possible; secondly, it is by no means certain
that the totality of the lexicon may be subjected to
structural analysis of the kind which seems to produce
disparate lexical systems—in fact, the valuable studies
so far undertaken in the area of semantic fields appear
to be isolated fragments of those elements of reality
which lend themselves to such decomposition. Hence,
apart from a few well-tried zones in the lexicon, colour
terminology, kinship terms, military titles and so on,
the semantic system seems to retain the appearance of a
loosely-organized mass of a vast number of distinct
elements. On the other hand, P. Guiraud has proposed
that lexical fields may be composed of "non-coordinated"
relations between the elements which make it up and that
we should not seek parallel organizations with the
closely-coordinated systems of phonology and syntax in
which the constituent elements each have a crucial
2.0 Componential Analysis

Another approach to dealing with lexical structure is the proposal that each lexical unit constitutes a complex of elements of meaning analogous to the distinctive features of the phonological system. A number of investigators have used this hypothesis as a starting point and Mounin is of the opinion that componential analysis, as it is called, provides the most scientific evidence yet available to support the theoretical possibility of translation. In field theory it was supposed that the minimal units of meaning were "words" (or morphemes or Martinet's "monèmes"), but componential analysis seeks to decompose the semic structure of these units into yet smaller elements. According to this theory, various combinations of these minimal elements of meaning constitute the units we think of as "words" or some other more abstract entity of the lexicon in a Katz and Fodor type theory. Yet this kind of analysis may be criticized on the grounds that the components of meaning thus isolated are non-linguistic units. Thus, for example, when B. Pottier decomposes the supposed semic features of "chairs" into elements which differentiate them according to various physical criteria (arms, back, legs, number of seating-places and so on),
this is really an extra-linguistic analysis of things, not linguistic elements. A. J. Greimas takes issue with this conception of what Pottier labelled as "sèmes," arguing that his features "avec dossier," "avec bras," "pour s'asseoir" and so on are merely further lexical items, signs designating real things, and do not therefore constitute the minimal elements of meaning which this kind of analysis demands. Similarly, when L. Hjelmslev suggests that much of the groundwork of componential analysis has already been accomplished by lexicographers and their monolingual dictionary definitions, he is overlooking the fact that the lexicographical descriptions are contrived for purely functional purposes and do not necessarily reflect the true semantic content of the lexical unit. G. Mounin (1972) has illustrated this objection by establishing the components of various dictionary entries according to the definitions supplied for a number of terms dealing with human habitations; the resulting elements were not only inconsistent between the dictionaries consulted, but also vague and imprecise. He proposes that the true components can only be isolated by a conceptual analysis in which the hierarchical structure is established on the basis of logical oppositions of the type [± animate], [± male] and so on.
2.1 However, while this kind of analysis with its scientific rigour is worthy of a theoretical approach to semantics and lexicology, it may not be appropriate to practical lexicography which operates at the functional level. In a monolingual dictionary there is little to be gained by decomposing the elements of meaning into atomistic units which are comprehended intuitively by the native speaker if his attention is directed to the somewhat broader elements of conventional metalinguistic definition. Likewise, in a bilingual dictionary these primitive semic elements may be subsumed under the more comprehensive labels which serve to differentiate senses. What lexicography may learn from the componential approach to lexical structure, therefore, is that definitions must be precise and totally unambiguous.

2.2 The basic problem of lexicology is the attempt to reduce the complex, amorphous mass of the lexicon to a structured system and, within the limits of certain clearly-defined and closed fields, this structure may be determined. But the structure of the lexicon is not uniquely semantic. Various analyses reveal that there are also morphological parameters which play an important part in overall lexical structure, forming nuclear complexes known as "morpho-semantic fields."  

2.3 The method of determining the physical nature of
these fields is to establish an exhaustive inventory of lexical items designating a specific concept and to carry out an analysis which is both distributional and componential and which leads to a hierarchical classification based on a very primitive "deep structure."

Pursued to its logical conclusion, this kind of analysis would establish a finite number of basic concepts or "semes" which would constitute the total semantic deep structure of the entire lexicon and enable all surface manifestations to be generated according to a set of transformational rules. The notion of a semantic base of this kind is the focus of attention of generative semantics, although scholars in this area have not so far given a very specific account of how the system would operate in practice. P. Guiraud (1967), however, has approached the question from the point of view of statistical parameters and proposed a system composed of 32 primitive "sèmes" which might correspond to notions like animate/inanimate, human/animal, male/female and so on, and which enter into various combinations, according to mathematical distributions, in order to generate all the theoretically possible permutations and combinations of the lexicon. Just as the phonological system of a language is composed of a finite set of phonemes which enter into various acceptable combinations to generate the lexical units of the language, so the hypothetical
set of basic semic elements would obey certain rules specifying the permissible combinations.

2.4 The concept of morpho-semantic fields was illustrated by P. Guiraud (1969) in his study of the notion "coup" which is particularly fertile ground for the production of morphological and derivational forms and would seem to constitute one of the primitive semic elements in the structure proposed above. It is possible to reconstruct the lexical sub-system which it generates on three discrete levels: (a) morphological (chic, chique, chiquer); (b) phonetic, where it can be demonstrated, for example, that the majority of French verbs with the sense "donner un coup" are formed on onomatopeic bases ([tık/tak/tɔk], [trık/trak/trɔk], and (c) semic, where specific components of meaning differentiate and group various lexical units derived from the basic seme "coup." From an analysis of the thousands of words which make up this morpho-semantic field it becomes evident that there is, in fact, a system which would appear to be generated from a base etymon corresponding to the notion of "coup" and which depends upon a relatively small number of simple transformations.17

2.4.1 The first subdivision separates the members of the field into categories corresponding to various kinds of
"blow," according to whether they are **specific**, **repeated** or of a **general** nature. Specific blows are further subdivided into the manner or result of the blow: "on a fixed point," "into small pieces," "on a flat surface" and so on, giving rise to a number of series such as: cime, pic, etc.; coupon, quignon, etc.; tampon, tache, aplatir, etc. Repeated and general blows may be similarly categorized so that "surface structure" words appear to be the result of three or four clearly formulated transformations at most:

![Diagram](attachment://diagram.png)

2.4.2 This approach to the analysis of lexical structure appears to be eminently workable and has the decided advantage of answering many of the objections referred to above, especially with regard to the purely subjective compilation of lexical items according to psychologically-based associations. Moreover, the notion that there is a "deep structure" status for such constellations of meaning, with a finite set of transformations which would
generate the surface manifestations, conforms neatly to a transformational-generative theory of language. One is therefore tempted to pose the question of whether this kind of lexical structure might constitute a language universal and whether, in the case of a positive response, a comparative study would shed light on some of the enigmas which surround translational equivalence. It would obviously be revealing in a number of ways, not the least of which would be an answer to the question whether it is the system of each language which determines the manner in which reality is viewed by the speakers of each language, or whether it is the reverse. Furthermore, if there appears to be some parallel between the equivalent structures of pairs of languages, it might be possible to identify and formalize the divergences which impose serious difficulties on the translation act.

3.0 A Comparative Example

It is beyond the scope of this chapter to undertake an exhaustive study of a morpho-semantic field in two languages. Instead, the remarkable results obtained by Guiraud (1969) have been taken as the model for a similar investigation in English so that the results can be compared with the original and some conclusions drawn. The first question to be resolved is the selection of the base etymon for a morpho-semantic field paralleling that
of French "coup." If the obvious equivalent "blow" is chosen, we can immediately identify a number of striking similarities as well as differences. Like couper, the verbal derivation of English "blow" has changed its meaning and has nothing to do with the notion of "giving a blow to something." However, whereas couper is a valid product of the lexical system and specifies a particular kind of blow, the English verb blow, in spite of the controversy surrounding its etymology, seems to be totally unrelated to the noun. In order to retain the parallelism, therefore, it would appear to be preferable to seek a different base form. Among obvious contenders --hit, slap, stroke, bang, etc.--the word stroke is the most generic and, like its counterpart coup, belongs to an extensive series of derivations and set phrases. Most dictionary definitions describe the word in terms of notions such as "movement" and "impact," although it may also designate "precise movement," "repeated movements," "single blows," and specify the "force" or "lightness" of the blow.

3.1 Even a superficial examination of an analogical dictionary reveals that the notion "stroke" in English is exceedingly prolific and bears a striking resemblance to the same series in French. For example, the morphological derivations include series based on prefixes:
percuss, perforate, incise, rebuff; suffixes: cleavage, buttress, rebuttal, battering; compounds: whiplash, breaststroke, drumbeat, pinch-hit, frogkick. There are also phonetic series based on onomatopeic roots exactly equivalent to those established by Guiraud (1969):


3.2 Semic elements also seem to correspond closely in the two systems: there is the same broad division of the base into notions of "marking," "bashing," "squashing," "dividing," "shattering," "shaking," "punishing" and so on, depending on the particular kind of action, the instrument used, the outcome of the blow and so on.

3.2.1 If we now observe some of these subdivisions a little more closely the remarkable correspondence between the two systems is clearly revealed. For example, just as French possesses a series of words to describe someone who has received a blow on the head and remained, as a result, somewhat unbalanced (touqué, piqué, cinglé, etc.), so English has the series: nut/nutty; crack/cracked/crackpot; strike/moonstruck/stricken; pot/potty; etc. Similarly, large numbers of words in the English series are derived morpho-semantically to designate spotted or striped animals: speckle/spackle; mackle/mottle; dapple;
3.2.2 The word "stroke" itself enters into a number of compounds and nominal locutions which correspond exactly to French forms containing the word "coup":

[RHYTHMIC]  
downstroke  
piston stroke  
coup d'archet  
coup de piston

[SUDDEN/VIGOROUS]  
stroke of lightning  
paralytic stroke  
sunstroke  
coup de foudre  
coup de sang  
coup de soleil

[FIGURATIVE]  
stroke of luck  
stroke of a brush  
coup de veine  
coup de pinceau

[DELIBERATE]  
stroke of a brush  
stroke of a pen  
coup de pinceau  
coup de crayon

PRECISE]  
stroke of midnight  
stroke of a master  
finishing stroke  
coup de minuit  
coup de maître  
coup de grâce

However, this series is heavily weighted in favour of "coup" which occurs in literally hundreds of locutions for notions which English may often render by a single word:

1. INSTRUMENT: SUDDEN BLOW  
coup de sabre (slash)  
coup de poignard (stab)  
coup de bâton (whack)  
coup de massue (club)  
coup de hache (chop)  
coup de griffe (scratch)  
coup de doigt (tap, rap)  
coup de dent (bite, snap, snip)
2. **INSTRUMENT: ANATOMY**
   - coup de bec (peck)
   - coup de corne (butt, gore)
   - coup de pied (kick)
   - coup de poing (punch, cuff, thump)

3. **ONOMATOPEIC**
   - coup de fusil (gunshot)
   - coup de sifflet (whistleblast)
   - coup de sonnette (ring, tinkle)
   - coup de tonnerre (peal, clap)
   - coup sec (snap, crack)

4. **METEOROLOGICAL**
   - coup de vent (gust, blast, squall)
   - coup de foudre (stroke, blast)
   - coup de sonde (sounding)

5. **MEDICAL**
   - coup de froid (chill, cold)
   - coup de soleil (sunstroke)
   - coup de sang (stroke)

6. **WEAPONS**
   - coup de feu (shot)
   - coup de canon (gunfire)
   - fusil à deux coups (double-barrelled gun)

7. **SPORTS**
   - coup de boxe (blow, punch)
   - coup de crosse (shot, hit, slapshot)
   - coup de dés (dice throw)
   - coup de dames (move)

8. **FIGURATIVE**
   - coup de langue (cutting remark)
   - coup de fusil (extorsion)
   - coup d'œil (glance)

9. **SET PHRASES**
   - de premier coup (at the first go)
   - d'un seul coup (all at once)
   - faire les 400 coups (to sow one's wild oats)
   - manquer son coup (to miss the mark)
   - tout d'un coup (suddenly)
   - coup de main (help, hand)

This is just a small sample of an extraordinarily extended series calling upon the French word "coup" to
act as a kind of mediating symbol for the expression of notions involving the idea of a stroking or hitting movement. 18

3.3. The degree of correspondence which this kind of analysis thus reveals is probably not entirely unexpected in view of the close genetic ties between English and French. Yet it is precisely the differences, rather than the similarities, which attract the attention of bilingual lexicography because these are the elements which constitute the difficulties encountered in the translation process. Hence, one byproduct of the investigation reveals a marked contrast in the manner in which English and French designate a wide variety of "hitting" movements; the former by means of an exceptionally extensive range of individual and specific terms; the latter by employing the formula [coup de + NOUN]. The differences relate not to any radical disparity in the way the real world is viewed by members of the two language groups but rather to the analytical as opposed to the synthetic organization of linguistic elements. 19

3.4 Lexicographers are usually familiar with the existence of such differential series, but a more intensive study is likely to reveal tendencies and correspondences which are concealed at the "surface"
level at which most practitioners operate and which should be brought to the attention of dictionary users in the form of synonyms and developmental cross-referencing. Once again, the basic problem in lexicography is that, in order to facilitate rapid consultation, the elements of the lexicon are extracted from their natural structural environment and relocated in the purely arbitrary and artificial milieu of alphabetically-ordered wordlists. In these surroundings, dissociated from the semantic constellations in which they operate in reality, the relational features of lexical units are largely eliminated and with them a considerable proportion of their meaningful elements. One of the tasks of bilingual lexicography, therefore, is the reconstitution of the relational network by furnishing as much information as possible in the form of explanatory labels, definitions, cross-references, synonyms and examples.

4.0 Analysis of "Closed" Fields

Lexical structures like the semantic fields represented by "coup" and "stroke" are open-ended and somewhat difficult to analyse exhaustively with the means presently at our disposal. The study of closed fields such as colour terms, military titles, terms of honorific address, kinship hierarchies and so on is clearly more fruitful and, in spite of the objections
raised above to this kind of investigation (that such series are special cases and not representative of the whole semantic system), it is certainly revealing with respect to what we can learn about the phenomena of equivalence between languages.

4.1 With this goal in mind, we shall now examine a closed semantic field in order to determine the degree of correspondence, or semantic isomorphism, which exists in a relatively restricted area of the lexicon. The field chosen is the measurement and denomination of degrees of temperature which, in theory, should constitute a universal scale.

4.2 Linguistically speaking, the temperature scale is represented by a continuum bounded at one end by the notion "hot" and at the other by "cold" with "warm" as an approximate mid-point. In terms of the lexicon, these terms seem to represent physiological reaction to the effects of heat and cold and the whole continuum may be conceived of as a series of deviations from the norm of "warmth" considered as pleasant, normal body temperature. The relational nature of temperature denomination is such that some terms may appear in more than one position on the scale. For example, the weather may be described as having turned cold after being very hot, but the physical temperature of this kind of "cold" is very different from
the cold used to depict a bleak wintry day with the thermometer registering well below zero celsius. It is also conceptually quite different from the cold used to describe a meal which is no longer hot and yet registers a physical temperature which would constitute almost unbearable heat in terms of the weather. Similarly "hot" in the sense of "feverish" when related to body temperature bears little physical resemblance to the hot used to describe a fire or the sun. Fluctuations in temperature, it would seem, cause parallel fluctuations in the associative lexical field which denotes degrees of warmth.

4.3 It is also necessary to bear in mind that many of the surface manifestations of temperature denominations are the result of sensorial transpositions (synaesthesia), the metaphorical device used in literature to add stylistic import to normal expression. Hence, although warmth is a phenomenon of temperature and would not normally be an appropriate attribute for visual or acoustic impressions, the language abounds in just such combinations: a warm voice, a warm colour, hot music, the cold shoulder, a cool reception, a lukewarm reception, a frosty reception, etc. The mechanisms involved in such sense transfers are quite elementary and are based on the affective analogy existing in each of the sense
impressions conveyed by the image involved. If temperature can be accepted as an attribute of sound in "a warm voice," it is because we are able to perceive a similarity between the pleasant sensations evoked by warmth of temperature and the particular quality of a voice in certain situations.

4.4 However, apart from the vagaries of metaphorical language, which are not the concern of this discussion, the lexical structure for the temperature scale in both French and English appears to be a well-balanced and relatively stable one which reveals a good deal about the way in which these two languages deal with a universal concept. Table III on the next page sets out the equivalent forms in French and English for terms which describe the physical and purely kinesthetic reactions to different temperatures. The scale is divided into four sections representing "hot," "warm," "cool" and "cold" which may be considered as the broad major divisions of the continuum. Within these divisions, both languages appear to operate at a parallel level suggesting that this kind of reality, not altogether unexpectedly, is viewed and depicted in almost identical fashion by the two languages. The main difficulties arise because the pair chaud/hot do not correspond exactly: something which is elevated in temperature may be described as hot
<table>
<thead>
<tr>
<th>HOT</th>
<th>English</th>
<th>French</th>
</tr>
</thead>
<tbody>
<tr>
<td>incandescent</td>
<td>chaufferé à blanc, blanchoyé</td>
<td></td>
</tr>
<tr>
<td>white hot</td>
<td>chauffé au rouge, rougi au feu</td>
<td></td>
</tr>
<tr>
<td>red hot</td>
<td>torride</td>
<td></td>
</tr>
<tr>
<td>torrid</td>
<td>grillant</td>
<td></td>
</tr>
<tr>
<td>sizzling</td>
<td>ardent</td>
<td></td>
</tr>
<tr>
<td>blazing</td>
<td>très chaud</td>
<td></td>
</tr>
<tr>
<td>scorching</td>
<td>torréfiante</td>
<td></td>
</tr>
<tr>
<td>scalding</td>
<td>brûlant</td>
<td></td>
</tr>
<tr>
<td>burning</td>
<td>rôtissant</td>
<td></td>
</tr>
<tr>
<td>broiling</td>
<td>rougeoyant</td>
<td></td>
</tr>
<tr>
<td>flaming</td>
<td>enflammé</td>
<td></td>
</tr>
<tr>
<td>boiling</td>
<td>bouillant</td>
<td></td>
</tr>
<tr>
<td>hot</td>
<td>tout chaud</td>
<td></td>
</tr>
<tr>
<td>piping hot</td>
<td>rougeoyant</td>
<td></td>
</tr>
<tr>
<td>glowing</td>
<td>embrasé</td>
<td></td>
</tr>
<tr>
<td>simmering</td>
<td>qui bouillotte</td>
<td></td>
</tr>
<tr>
<td>steaming</td>
<td>chaud (mais non brûlant)</td>
<td></td>
</tr>
<tr>
<td>warm</td>
<td>à la temperature du sang</td>
<td></td>
</tr>
<tr>
<td>blood-hot</td>
<td>tempéré</td>
<td></td>
</tr>
<tr>
<td>temperate</td>
<td>doux</td>
<td></td>
</tr>
<tr>
<td>mild</td>
<td>dégourdi</td>
<td></td>
</tr>
<tr>
<td>tepid</td>
<td>tiède</td>
<td></td>
</tr>
<tr>
<td>lukewarm</td>
<td>frais</td>
<td></td>
</tr>
<tr>
<td>cool</td>
<td>frisquet</td>
<td></td>
</tr>
<tr>
<td>frosty</td>
<td>froid</td>
<td></td>
</tr>
<tr>
<td>COOL</td>
<td>gelid</td>
<td></td>
</tr>
<tr>
<td>icy</td>
<td>gelé</td>
<td></td>
</tr>
<tr>
<td>frozen</td>
<td>glacié</td>
<td></td>
</tr>
<tr>
<td>frigid</td>
<td>congelé</td>
<td></td>
</tr>
<tr>
<td>COLD</td>
<td>frozen</td>
<td></td>
</tr>
<tr>
<td>cold</td>
<td>glacial</td>
<td></td>
</tr>
<tr>
<td>gelid</td>
<td>surgelé</td>
<td></td>
</tr>
</tbody>
</table>
in English, but must be très chaud in French because chaud alone is not sufficiently specific. Conversely, things which would be described as moderately hot or only warm in English are depicted as chaud in French. What we are confronted with here is just one of the many cases where one of a pair of so-called translational equivalents has a wider semantic extension than the other. This differential situation may be represented according to the following schema which suggests that, in accordance with the facts, the ternary system in French is being related to a dual one in English:

<table>
<thead>
<tr>
<th>TIEDE</th>
<th>CHAUD</th>
<th>AUERD</th>
<th>TRES CHAUD</th>
</tr>
</thead>
<tbody>
<tr>
<td>WARM</td>
<td>HOT</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

4.5 Further analysis of the data emerging from this examination of the temperature scale in the two languages reveals that some of the words used to describe the heat/cold continuum are restricted to certain kinds of physical matter (see Table IV next page). Moreover, we are able to observe clearly the semantic disparity between certain pairs of words, such as piping hot/tout chaud where the English expression is restricted to liquids and solid food whereas the French equivalent may apply to any physical substance or phenomenon. Of
<table>
<thead>
<tr>
<th>ENGLISH</th>
<th>TEMP</th>
<th>SOLID</th>
<th>LIQUID</th>
<th>GAS</th>
<th>WEATHER</th>
<th>FOOD</th>
<th>FRENCH</th>
</tr>
</thead>
<tbody>
<tr>
<td>incandescent</td>
<td></td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>incandescent</td>
</tr>
<tr>
<td>white hot</td>
<td></td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>chauffé à blanc</td>
</tr>
<tr>
<td>red hot</td>
<td></td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>chauffé au rouge</td>
</tr>
<tr>
<td>torrid</td>
<td></td>
<td></td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>torride</td>
</tr>
<tr>
<td>burning</td>
<td></td>
<td></td>
<td>(+)</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>brûlant</td>
</tr>
<tr>
<td>HOT</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>(TRES) CHAUD</td>
</tr>
<tr>
<td>piping hot</td>
<td></td>
<td></td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>tout chaud</td>
</tr>
<tr>
<td>glowing</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>rougeoyant</td>
</tr>
<tr>
<td>WARM</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>TIÈDE</td>
</tr>
<tr>
<td>lukewarm</td>
<td></td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>dégourdi</td>
</tr>
<tr>
<td>temperate</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>tempéré</td>
</tr>
<tr>
<td>mild</td>
<td>+</td>
<td></td>
<td>-</td>
<td>(+)</td>
<td>+</td>
<td>-</td>
<td>doux</td>
</tr>
<tr>
<td>COOL</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>FRAIS</td>
</tr>
<tr>
<td>frosty</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>frisquet</td>
</tr>
<tr>
<td>COLD</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>FROID</td>
</tr>
<tr>
<td>stone cold</td>
<td></td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>tout froid</td>
</tr>
<tr>
<td>frozen</td>
<td></td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>congelé</td>
</tr>
<tr>
<td>frigid</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>glacial</td>
</tr>
<tr>
<td>icy</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>glacé</td>
</tr>
</tbody>
</table>
course, the expression "piping hot" in English has been contrived for a particular purpose and merely replaces the less expressive utterances "very hot," "quite hot," "very warm" and so on which could do service in the same context equally well. Similar remarks apply to the expression stone cold and its equivalent tout froid, whereas the pair lukewarm/dégourdi cover the same conceptual surface except with regard to food to which dégourdi does not apply. Perhaps the most striking feature of the temperature matrix is the fact that the generic terms hot/chaud, warm/tiède, cool/frais, cold/froid are exactly equivalent in terms of their application to physical phenomena and yet exhibit differences in contextual application, while the intervening lexical items tend to be more heterogeneous.

4.6 The information derived from investigations of this kind is intrinsically revealing and may eventually prove to be an important means of adding to our understanding and knowledge of the vast and virtually untouched domain of differential lexicology. It is equally important for providing tangible evidence for the view expressed earlier that one of the fundamental problems of bilingual lexicography is the fact that the translations provided in a dictionary are abstractions of reality, instances of langue which may not be realized in parole. We are now
in a position to determine that one of the main reasons for this situation is that the semantic components embodied in any given pair of equivalent forms in two languages rarely if ever constitute two identical sets. Only in the relatively small zone where near semic identity exists (scientific language, rare words and so on) may we expect true surface equivalence. Thus, for example, if we provide French equivalents for English lukewarm we must clearly indicate the area of application for each term since none is exactly equivalent to any other (as we observe in Figure 3). English lukewarm may refer to things which have cooled down from "hot" or which have become warm from a desired state of "cold" (cool drinks, etc.). French tiède likewise operates in both directions on the continuum but it has a much wider semantic extension than its English counterpart since, unlike the latter which implies "insufficiency of warmth"

---

Figure 3. An example of differential equivalence: "lukewarm."
or "lack of coldness" and has rather pejorative connotations, it denotes "pleasantly warm" situations such as spring weather, a bath, a bed and so on, none of which could be described in English as "lukewarm." All of these characteristics, related to the semic structure of the individual lexical items, should ideally be reflected in lexicographical methodology. However, as the following examples taken from a variety of bilingual dictionaries illustrate, the treatment is not always in the best interests of the user:

HAP1 lukewarm a. tiède. To become 1., s'attiédir.

HAP2 lukewarm a. Tiède. To become 1., s'attiédir, tiédir.

HAP3 lukewarm a. (of water, friendship, etc.) Tiède; (of water) dégourdi. To become 1., s'attiédir. To make 1., (at)tiédir (l’eau, etc.).

DPL lukewarm tiède (a. fig).

OXF lukewarm, (lit. and fig.) tiède.

BELL Lukewarm a. tiède.

CASS1 lukewarm a. Tiède; (fig.) peu zélé. To get lukewarm, s'attiédir.

CASS2 lukewarm a. tiède; (fig.) peu zélé.

GASC lukewarm adj. tiède. To become or get —, s'attiédir. To make —, attiédir.

CD1 lukewarm adj. [proper and Fig.] tiède; [Fig.] froid, indifférent, sans enthousiasme.
4.7 For most of these dictionaries the main goal seems to have been the requirement that the word's figurative connotations should be stressed or that contextual examples should be provided. The important differential features are ignored in all but the largest Harraps dictionary (HAP3), although even this does not furnish sufficient information on the true equivalence relation. One way in which the nature of the equivalence relation might be more carefully and scientifically expressed in a bilingual dictionary is as follows:

\textbf{lukewarm} adj 1. [food, drink, water: désagréable] tiède, refroidi; peu chaud; [water] dégourdi. 2. [Fig.] [friendship, reception, etc.] tiède, assez froid.

Although this article might be criticized as being excessively long and the explanatory material cumbersome, it does capture the essential semic features and specify the differential nature of the equivalence relation. Without going so far as to provide complete definitions for the word and its equivalents, the bracketed details at least enable the user to make the correct selection and, as a pedagogical adjunct, reinforce his conception of the relationship between the equivalents. These, then, are the phenomena of lexical structure with which lexicographers must be familiar and which may well become available to them as a consequence of an extensive
investigation of the kind suggested by field theory and componential analysis.
Notes to Chapter 6

1Cf. P. Guiraud (1975): "ils constatent la complexité du lexique et renoncent à le réduire à une structure systématique; en fait ils nient l'existence d'une telle structure au niveau lexical. C'est, en gros, l'opinion des linguistes structuralistes."

2"Les champs apparaissent comme des petits îlots de structures isolées sur l'immense surface inorganique du lexique. Le problème demeure de les réduire à un système général qui en intégrerait l'ensemble."


4As A. Rey (1970: 130) puts it, "le fait de partir du concept sans aucun contrôle formel donne selon nous à cette méthode un caractère plus métaphorique que scientifique."


6Ibid., pp. 73-4.

7Mounin (1963), p. 78.


15. "... elles (reposent) sur des relations à la fois sémiques et morphologiques d'une très grande généralité dont on pourrait penser que, à un certain niveau d'abstraction, elles étaient universelles"; P. Guiraud, "Distribution et transformation de la notion de coup," Langue Française 4, 1969, p. 67.


17. Ibid., pp. 67-74.

18. For a complete inventory of expressions involving the word "coup" see Appendix No. II.


20. J. Rey-Debove (1970: 28) considers this to be essential: "Nous pensons personnellement qu'un dictionnaire bilingue devrait présenter au moins un double système d'équivalents: pour un dictionnaire A/F: une définition de l'entrée A en langue F et une équivalence lexicale de A en langue F. L'idéal serait un système quadruple, c'est-à-dire en fait, la confrontation de deux monolingues."
References for Chapter 6


CHAPTER 7

TRANSFORMATIONAL-GENERATIVE GRAMMAR IN
BILINGUAL LEXICOGRAPHY

1.0 Preliminary Notions

The impact of generative grammar on every aspect of linguistic theory over the past two decades is such that it cannot be ignored in a discussion of the theory of bilingual lexicography. In the initial model presented by Chomsky (1957) semantic considerations were held to be of marginal concern to syntactic description; while admitting that there are "striking correspondences between the structures and elements that are discovered in formal, grammatical analysis and specific semantic functions" (p. 101), Chomsky argued in favour of treating syntax as the main object of study which would eventually shed light on semantic problems. In the first model meaning was determined in part by a level of underlying structure. Important generalizations could be captured by the theory, but no explicit mechanisms were provided since the theory was basically concerned with the formal syntactic devices of language.

1.1 The view that there should be a clearly formulated
and separate semantic component in the grammar was first expressed in emphatic terms by J. Katz and J. Fodor (1963) who attempted to formulate a semantic theory which would account for the ability of native speakers "to produce and understand sentences never before encountered" (pp. 481-2). A speaker's knowledge of his language, according to the theory, takes the form of "rules which project the finite set of sentences he has fortuitously encountered to the infinite set of sentences of the language" (p. 482). The problem of formulating these rules is referred to as the projection problem.

1.2 More specifically, Katz and Fodor wished to account for the fact that the difference between a human subject and a mechanical device, both of which are required to interpret various sentences, is that the machine cannot distinguish between sentences which are syntactically correct but semantically ambiguous. The native speaker has at his disposal rules for applying "dictionary information" in such a way that he can interpret semantic relations between the elements of a sentence and the interaction between meaning and syntactic structure in order to arrive at a correct reading of the sentence. Katz and Fodor endeavoured to typify the speaker's control of the semantic aspects of language by assuming a dictionary-style structure consisting of two discrete
parts—a grammatical and a semantic.

The function of the grammatical markers in their theory is to indicate part of speech, the first element in their ramified structure, which enables an initial interpretation to be made. The semantic markers, comparable to the "classèmes" of A. J. Greimas, further categorize the sense paths which may be followed in the interpretation of meaning until the terminal distinguishers are reached. These elements decompose meaning into its final components, unless further categorization is rendered necessary by selectional restrictions, such as the special meaning of "honest" when referring to women. The Katz and Fodor "dictionary" is characterized by a "list (ordered or not) of the lexical items of the language, each item being associated with an entry in our normal form."

1.3 Although it is not made clear by their choice of terminology, the Katz and Fodor "dictionary" is not to be confused with a conventional, physical dictionary. In order to emphasize the unsatisfactory structure associated with regular lexicographical articles, Katz and Fodor demonstrated their own conception of the structure by transforming the monolingual dictionary entry for the word "bachelor" so that the senses are formalized according to semantic markers or components.
Although their goal was to postulate a structure which would operate within the semantic component of a general theory of language, the reorganized dictionary entry bears a striking resemblance to traditional bilingual dictionary entries with their disambiguating labels:

A bilingual dictionary entry for the same word might look thus:

```
+--------------------------+
| bachelor                |
+--------------------------+
| noun                    |
| (Human)                 |
| (Animal)                |
| (Male)                  |
| (Male)                  |
| (Young)                 |
| (Young)                 |
| who has never married   |
| knight serving          |
| Fur seal when without a mate during breeding time |
|                      |
| who has the first or lowest academic degree |
| under the standard of another knight |
|                       |
+--------------------------+
```

```
bachelor

n

1. [Soc]
   [unmarried]
   célibat m

2. [Hist]
   [degree]
   licencié m
   licenciée f

3. [Univ]
   [seal]
   jeune phoque m

4. [Zool]
   chevalier m
```
The material presented in the bilingual article is characterized by the same general semantic markers as the Katz and Fodor model in which, for example, the extreme left "sense path" is + (Human) + (male) + [who has never married]; the markers "human" and "male" thus terminate with the definition "never married." In the bilingual entry semantic components are not explicitly stated but are subsumed under various labels: [Soc] (= society, sociology) implies human; [unmarried] provides further specification; the gender of the translation, although not an invariably reliable indicator, suggests the feature "male."

2.0 Dictionaries as Models of Verbal Behaviour

The Katz and Fodor formulation of semantic structure should be viewed as a purely schematic conception of the manner in which verbal information may be internalized since, even though monolingual dictionary entries (the source of their model) are based on an intuitive and probably mentalistic supposition about the storage of information in human memory, it is unlikely that this structure even remotely resembles that inaccessible facility.

2.1 In order to investigate the directly unobservable neurolinguistic system of the human brain it is necessary
to infer its properties by an examination of its output, both in normal subjects as well as in aphasics whose brain malfunctions may provide clues to the manner in which information is stored. The monolingual dictionary, in fact, is basically an unsatisfactory model because it is a unilateral device (going only from word to definition), which might be an appropriate way of characterizing the reception of speech—the decoding of a message by a stochastic process which incrementally assigns meanings to the elements of a string. It is, however, inappropriate for the reverse process, the production of speech, which clearly proceeds from concepts and meanings towards the linguistic signs, appropriately organized, which represent them. This process would be better served by a model based on the so-called "reverse" dictionaries which offer lexical units as the reciprocal elements of entries which are, to all intents and purposes, phrases or conceptual units.

2.2 A number of psycholinguistic studies are of significance in this regard. One of them seeks solutions to the problem of information storage and retrieval by postulating a model based on an analogy with a punched card system in which a variety of different semantic markers corresponding to different associations are punched for each target word. Cards can be stored
randomly but retrieved speedily because of the broad associative network which links them. It has been proposed that most experimental work with human subjects points clearly to the fact that randomly presented material is reorganized for storage and subsequently recalled in a different sequential form involving "associative clustering" or "category clustering." 7

This would seem to suggest that the structure of the semantic portion of a theory of language should use the analogical dictionary, with its associative matrix, as the model for overall internal storage. Discrete meaning elements (words and their meanings) would then be characterized as bipartite units, similar in structure to the Katz and Fodor model, but composed of "sense paths" oriented in both directions in order to reflect reception as well as production.

2.3 If conventional monolingual dictionaries are an inadequate model in the theory, then bilingual dictionaries are doubly so since they imply that the theoretically perfect bilingual subject stores and retrieves information in two independent components related only when a pair of signs (one for each component) is attributed to a single concept. This typifies, rather, a state of artificial bilingualism, the most common variety, which may well be induced by the imperfect
nature of dictionaries for, as the discussion on the neurological basis for bilingualism (Chapter 8) reveals, natural bilingualism implies the interdependence of the two codes.

2.4 J. D. McCawley (1968) contends that the Katz and Fodor adaptation of dictionary entries like bachelor, which are polysemous, merely perpetuates traditional lexicographical practice by grouping together under a single entry "all the readings that can be associated with a given phonological shape and belong to a single syntactic class." There are a number of compelling reasons, suggests McCawley, for supposing that language operates, instead, in terms of the "lexical items" suggested by U. Weinreich (1966) in which the four polysemous senses of bachelor would instead be represented as four independent lexical items, each pronounced and spelt "bachelor."

3.0 The Revised Semantic Component

Influenced largely by the proposals made by Katz and Fodor (1963) and subsequently by Katz and Postal (1964), Chomsky reformulated the theory in order to give more prominence to a "semantic component" which is nourished directly by the "base component" and enables the meaning of sentences to be analysed in the same precise way as
their syntax was in the earlier model. Elaborating on the notion of the Katz and Fodor "dictionary entry," Chomsky envisaged a complex which he described thus: "A lexical entry may be considered a triplet \((P, G, \mu)\) in which \(P\) is a set of phonological features, \(G\) a set of syntactic features and \(\mu\) a set of semantic features."\(^{11}\) More specifically, these elements are classified as (a) **strict subcategorization features**, which prescribe the categorial contexts in which the lexical item may appear, and (b) **selectional features**, which determine the lexical features with which a given item can be related in combination with other items. These lexical features are further characterized by Chomsky as "higher-level lexical features," such as [Count], and "lower-level lexical features," such as [Human].\(^{12}\) The higher-level features are those which are encountered closest to the base component, that is, in deep structure. Lower-level features are more specific components of meaning whose impact is felt closer to the surface and which, Chomsky argues, are less crucial to the "correctness" of the final reading of the string. This is because it appears that "deviation from selectional rules involving higher-level features is apparently more serious than deviation from selectional rules involving lower-level features."\(^{13}\)

3.1 This conception of the lexical unit as a complex of features\(^{14}\) is not a significant departure from tradi-
tional lexicography which invariably describes the purely physical and formal features of each entry (graphic and phonological), followed by syntactic and semantic information. Nor is the hierarchical structure proposed by Chomsky for the order in which the features apply especially novel, since the first content element provided in all dictionaries is grammatical category and, in bilingual dictionaries at least, lower-level features such as [animate], [male], are subsumed in the various explanatory labels which precede the terminal elements (translations) and are in the form of explanations or definitions. The novelty lies, rather, in the conception of the "projection" of semantic information onto phrase markers which constitute the "deep structure" from which meaning is derived by the rules of semantic interpretation.

4.0 Precision in Lexicographical Description

What bilingual lexicography has to learn from the outcome of generative grammar and generative semantics is that a more rigorous analysis of the semantic components of meaning is called for, because these elements, which are transparent for the native speaker, may be elusive for the non-native learner of the language and should be spelled out in the interests of clarity and scientific attention to important linguistic detail. Thus, in
addition to supplying explanatory and disambiguating material, it is appropriate to provide for the non-native user information on the selectional restrictions which apply to certain words. This does not mean that all such restrictions should be tabulated, but simply that there should be some elaboration on features which may be peculiar to one of the languages and which are unpredictable in the normal course of events. For example, the word buxom is marked as being restricted to female subjects and this feature must be indicated for non-English speakers:

buxom adj [femme] (jolie et) rondelette; accorte.

Similar considerations would govern the treatment of the following:

kin n(inv) [Ethn] les parents mpl.
célibataire nm/f [Soc] [homme] bachelor; [femme] spinster.

4.1 Certain lexical items display selectional restrictions which are important enough to warrant their specific mention in dictionary articles. For example, the adjective mere, unlike its French equivalents simple, pur, seul, which are fully versatile adjectives, can only be used in certain contexts. According to Carlotta Smith
(1969), modifiers like mere, utter, real, are different enough from regular adjectives to require special treatment. She classes such words under the heading of intensifiers and claims that, since they cannot occur in kernel structures of the configuration [NP is Adj] and do not behave as noun constituents under transformations, they should be considered as a special category of determiners. Hence, although one can say:

(1) He was a mere boy
(2) C'était un simple garçon,

it is not possible to transform (1) in the following ways:

(3) *The boy was mere
(4) *The boy was very mere.

The French equivalent (2), however, may undergo these syntactic alterations, although not without the possibility of semantic deviation:

(5) Le garçon était simple (The boy was simple)
(6) Le garçon était très simple (The boy was very simple).

4.2 The manner in which selectional restrictions should be recorded in a dictionary is not altogether clear, since it would be necessary to adopt conventions which
are both succinct and fully comprehensible to the non-native user for whom they are contrived. Probably the most convenient arrangement would be to discuss questions of this kind in the preface, to identify them in the body of the work by a series of conventional symbols \((\text{adj det, adj inv, adj [intense]}\text{ and so on})\) and to employ as much exemplifying material as is necessary to remove all possible confusion. Although it is not the task of a bilingual dictionary to provide a complete grammar of each of the languages, there is certainly an obligation to furnish non-native users with as much support as possible, particularly where this material is an essential element of total lexical structure as posited in transformational generative grammar.

5.0 Transformational Solutions for Lexicographical Problems - Verbs

Generative grammars have addressed themselves to many of the problems which have long plagued lexicography and for which universally acceptable solutions have generally not been found. Many modern monolingual dictionaries, for example, employ a single category label for all verbal entries \((v\text{ or vb})\) thus avoiding the issue whether there should be subdivisions for transitive, intransitive or pronominal use; and these divisions are by no means invariably transparent even for the
sophisticated native speaker of the language. The argument in favour of this approach, which is nowhere openly expressed, is probably that the definitions given for verbal forms are sufficient to indicate those features which are generally of little significance to the native speaker. For the non-native user of the dictionary, however, such information is vital and this is why some bilingual dictionaries have attempted to describe verbal entries in more detail, indicating instances where a transitive use in one language may be rendered by an intransitive or pronominal in the other. On purely theoretical grounds, moreover, there is sufficient evidence to justify the treatment of verbs in independent categories where they effectively function as independent lexical units demanding separate entry status. This is precisely the view expressed by Chomsky (1967: 214):

If we regard a lexical entry as a set of features, then items that are similar in sound, meaning, or syntactic function will not be related to one another in the lexicon. For example, the Intransitive "grow" of "the boy grew" or "corn grows," and the Transitive "grow" of "he grows corn" would have to constitute two separate lexical entries, despite the meaning relation that holds between them, since there is apparently no way to derive the Intransitive structure from the Transitive ones, as can be done in the case of "the window broke," "someone broke the window."
5.1 The basis for Chomsky's contention that separate lexical units must be established for such verb groups is purely syntactic; he does not suggest that every sense change should be accompanied by the establishment of independent lexical entries, as has been proposed by some lexicographers and more recently by McCawley (1968). Further evidence in favour of Chomsky's view is provided by the bilingual situation in which the respective TL equivalents are not-necessarily categorized syntactically in the same manner as in SL. French, for example, has what amounts to a double series of intransitives in the form of its abundant reflexive verbs; sometimes the reflexive does duty for a non-existent intransitive form, in other cases it is merely a synonymous form:

\begin{verbatim}
stop vtr arrêter
stop vi arrêter, s'arrêter
abandon vtr renoncer à (qqch).
\end{verbatim}

5.2 Another lexicographical problem which has been analysed in generative terms is the question of phrasal verbs like "bring in," "call up," "put up with," which Chomsky has characterized as "verb + particle (V + Prt)." The original phrase structure rules for simple verbs were set out thus:

\begin{verbatim}
(i) Verb + Aux + V
\end{verbatim}
(ii) \( V \rightarrow \text{hit, take, walk, read, etc.} \)

(iii) \( \text{Aux} \rightarrow C(M) \) \( \text{(have + en)} \) \( \text{(be + ing)} \) 
\( \text{(be + en)} \)

(iv) \( M \rightarrow \text{will, can, may, shall, must} \) \( ^{18} \)

But this description will not account for discontinuous elements as in the following sentences:

(7) The police brought in the criminal

(8) The police brought the criminal in

(9) The police brought him in

(10) *The police brought in him. \( ^{19} \)

Recognizing the importance of such verbal groups and wishing to be able to analyse them within the framework of his theory, Chomsky then added the following rule to rules (i) to (iv) above:

(v) \( V \rightarrow V_1 + \text{Prt} \)

which has the effect of factorizing the verbal group thus:

\[
\begin{array}{c}
V \\
V_1 \quad \text{Prt} \\
| \quad | \\
\text{bring} \quad \text{in}
\end{array}
\]

A set of supplementary rules is posited to indicate what combinations of \( V_1 \) and \text{Prt} are permissible and an
optional transformation enables strings of the configuration \([V_1 - \text{Prt} - NP]\) to be transformed to \([V_1 - NP - \text{Prt}]\) so that sentences (7) and (8) above are satisfactorily accounted for. However, in order to be able to capture the possibility of (9) without generating (10), it is necessary to add an obligatory transformation which operates on strings with the structural analysis \(V_1 - \text{Prt} - \text{Pron}\). \(^{20}\) We have then:

```
S
\-----------------------
| NP             VP |
\-----------------------
    \-----------------------
    | V         NP         |
    \-----------------------
        \-----------------------
        | V_1       Prt       Pron |
        \-----------------------
            \-----------------------
            | bring   in   him |
            \-----------------------
```

which by the obligatory rule just quoted becomes:

```
VP
\-----------------------
| V         NP       Prt |
\-----------------------
    \-----------------------
    | V_1       Pron     |
    \-----------------------
        \-----------------------
        | bring   him   in |
        \-----------------------
```
In order to ensure that the correct passive constructions are generated:

(11) The criminal was brought in by the police
(12) He was brought in by the police

but not:

(13) *He in was brought by the police

the transformational rules are so ordered that the passive rule applies before the obligatory factoring transformation.

5.3 The same question is again taken up by Chomsky (1965) where the Verb-Particle constructions are likened to Verb-with-Complement constructions such as "take for granted," "take offense at" whose syntactic behaviour is similar.\(^{21}\) The former group can be divided into those where the "adverbial" element is fairly free (as in bring in, bring out, bring up, bring down, bring together, etc.) and may be considered, to all intents and purposes, as a totally independent distributional variable, and those where the Verb-Particle relationship is so close that, distributionally as well as semantically, they constitute separate lexical units.\(^{22}\) This would imply that dictionaries should create separate entries for such phrasal verbs as see to, bring off, carry out
(in the respective senses "ensure," "accomplish" and "complete") which more often than not are simply cited as examples under see, bring and carry respectively, or else ignored altogether. Without coming to any conclusive results, Chomsky discusses various ways in which Verb-Particle constructions, all of which seem to obey various familiar transformational rules in a regular manner, may be accounted for in the grammar. 23

5.4 Thus, the intuitively-based practice of a few innovative lexicographers finds theoretical support in the elementary workings of generative grammar and the treatment of phrasal verbs as autonomous lexical units requiring independent entry status in the dictionary is justified. In monolingual lexicography, which is aimed almost exclusively at native speakers, it is considered necessary only to identify the phrasal verbs and to provide their definition. Bilingual lexicography, on the other hand, must furnish full details, including grammatical information on special conditions governing their use in context. Since such elements tend to be consistent in their syntactic behaviour, it would probably be sufficient to include a preface note outlining various distributional realizations which might be encountered. 24
6.0  **Transformational Solutions for Lexicographical Problems - Adjectives**

Another aspect of grammar of concern to lexicographers and which has claimed the attention of generative grammarians is the status and origin of adjectives.\(^{25}\) Chomsky (1957) demonstrated that it would be simpler if the attributive position of adjectives were derived from a transformation which converts strings of the configuration \([N - is - Adj]\) to \([Adj + N]\), thereby enabling all adjective-noun combinations to be eliminated from the kernel.\(^{26}\) A question which is of equal interest to lexicography and grammatical theory, however, is the distinction which must be drawn between true adjectives, possessing a set of inflexional affixes (-er, -est) and capable of being modified (very, most), and words which superficially appear to function as adjectives but which are derived from other categories (noun and verb):

(14) The child is sleeping
(15) The sleeping child.

Whereas *sleeping* in (15) is a verbal form functioning like an adjective, words like *interesting*, as in:

(16) The book is interesting,

are taken to be true adjectives, both because they do not have the same intransitive verbal relationship as
sleeping (the child will sleep, but not *the book will interest), and also because they can occur with modifiers (or intensifiers) such as very, quite, extremely, whereas their verbally-derived counterparts may not:

(17) The book is quite interesting
(18) *The child is quite sleeping.

6.1 Chomsky's formulation bestows special status on verbally-derived adjectives like sleeping in order that they may undergo the attributive adjective transformation which converts (14) to (15); the rules preclude their occurrence in contexts such as [very + Adj] or [seem + Adj]. Words like sleeping, jumping, hanging, flying, etc., are traditionally described as participles, that is, they are derived from verbs but behave like adjectives with certain limitations. There is a basic difference between these words and the other class of lexical items which behave in every respect as true adjectives: interesting, frightening, amusing, charming. The latter share the common semantic feature that they are connected with a specific human reaction; they also have in common a number of important syntactic features not shared by the former group, such as transitivity and the ability to occur with modifiers and be dependent on verbs like seen. 27
A further distinction has to be drawn between participial adjectives used in expressions like (15) as well as:

(19) The hanging gardens of Babylon
(20) Flying saucers (are common)
(21) A going concern

all of which are derived from intransitive verbs:

(22) The gardens are hanging
(23) The saucers are flying
(24) The concern is going (well)

and the following group:

(25) Eating apples (are delicious)
(26) Drinking water (is not poisonous)
(27) Typing school

which clearly are not so derived. Both types have been the object of fairly detailed analysis in generative grammar with the goal of attempting to posit identical deep structures for both kinds. On the other hand, lexicographers have tended to overlook research in the field and persist in using traditional and inaccurate grammar labels.

From a lexicographical point of view, there is a
fundamental difference between the three kinds of lexical unit represented in (16), (19) to (21) and (25) to (27). Dictionaries almost invariably list words like interesting (16) as adjectives, although there is a good deal of inconsistency with respect to which are included and which are omitted, perhaps on statistical grounds or else because it is assumed the reader can generate such forms for himself. For example, most dictionaries give charming, overbearing, exacting, frightening, but omit mystifying, electrifying, scintillating, presumably because their existence is transparent for native speakers.

6.3.1 To illustrate the confusion and inaccuracy which these lexical units may cause, one dictionary gives going as an adjective and illustrates it with the sentence the clock is going\(^{29}\) in which it cannot be construed as anything other than a purely verbal use. Sentences (19) to (21) pose a different problem in bilingual lexicography. Most monolingual dictionaries enter these forms as adjectives with no further comment, a situation which probably poses little problem for most native speakers (although young users might well find themselves bewildered by it) but which does nothing to inform the non-native speakers about the special conditions which distinguish these words from those like interesting which,
to all intents and purposes, are unrestricted adjectives. But bilingual dictionaries likewise tend to label such words as adjectives:

Flying adj 1. volant, flottant, léger.
2. rapide, court: to pay a flying visit to Montréal, faire un voyage éclair a M. 30

In this entry, volant and flottant are exact morphological equivalents of flying in that they are really present participles and obey the same restrictions as the English word:

(28) Une machine volante
(29) Des machines volantes

but not:

(30) *Cette machine semble volante
(31) *Cette machine est extremement volante.

6.3.2 All other translations given here (léger, rapide, court) are true adjectives, although nothing indicates that this is the case. Speakers of the other language will no doubt be able to interpolate, from the information given, the special conditions which apply, yet it would probably make more sense, linguistically as well as pedagogically, to modify the category label in some way, for example, by the addition of an indicator that it is used attributively or that it is a participle. 31
6.4 The third group of examples, (25) to (27), illustrate what modern lexicography has come to regard as lexical units or "multiword lexical units" which appear as separate entries in an increasingly large proportion of monolingual dictionaries. This practice, which is largely based on the intuitive feelings of lexicographers as well as on the notion that "word = sense unit," probably has something to offer grammatical theory which generally still persists in trying to derive such forms from deep structures involving one element in NP and the other in VP. Phrases of the form [V + ing + N] are derived by transformations which depend on notions like the "subjectivity" or "objectivity" of the relationship between the noun and the verb which requires additional rules to block the generation of:

(32) *The apples are eating

but permit:

(33) The eating of the apples.  

6.5 Language is in a state of continuous evolution and presumably theories which are concerned with describing the lexicon must conclude, at some point, that a "multiword" expression has become lexicalized and behaves as though it were a single "word." Although the separation by blanks of drinking water, typing school, fishing
boat and even dining room tends to keep us aware of the individual components of meaning of which they are composed, there is no logical reason why they should be considered any differently than blackbird, soundproof, peroxide, bandy-legged or even graceful and prevent all of which, in varying degrees of cohesion, are compounds of morphemes or lexemes. Lexical units of more than one element like flying fish, kindling wood and driving school behave exactly like their single element counterparts with respect to syntactic function, morphological variation (two driving schools, several flying fish, some kindling wood) and semantic content. Moreover, the existence of single element equivalents for the identical concept in other languages lends weight to the contention that they should be considered as autonomous lexical units: driving school/auto-école, skating rink/patinoire, spinning wheel/rouet.

6.6 Finally, there is a large group of English nominals which are used attributively and which a generative grammar would derive on an analogy with those in (25) to (27):

(34) sports clothes
(35) home plate
(36) mother country.
There are three ways of treating the attributive portion of these lexical units in dictionaries: they may be labelled as adjectives which is what most dictionaries do; they may be labelled as attributives which a few dictionaries do, usually as a sub-category of a nominal entry; or they may be treated as indivisible lexical units in those cases where the attributive occurs only in such units and not in free distribution. The first alternative, as we have seen is inappropriate in bilingual dictionaries and linguistically unscientific. The second choice would be valid in situations where the word is not restricted in its distributional behaviour; if this were the case, it would be better to assign the word a separate entry from the noun entry. The final alternative depends to some extent on the whole question of compounds which we shall now examine in some detail.

6.7 R. B. Lees (1960) proposed that what is generally referred to as "deep structure" may in fact be only an intermediate level of representation and that there is a much deeper level which is "so close to what we might call the meaning of a sentence that there may be no validity to maintaining a distinction between these." In his study of English nominal compounds, Lees observed that native speakers understand the relationship between their constituent elements in much the same way as the
elements of a whole sentence. He formulated a set of transformational rules which derive compounds from their underlying sentence representations. For example, drawbridge is supposed to undergo the following transformations:

(i) a bridge which is for someone to draw the bridge
(ii) a bridge for someone to draw the bridge
(iii) a bridge for someone to draw
(iv) a bridge to draw
(v) a drawbridge. 36

A general inversion rule, \([N_1 + P + N_2 \rightarrow N_2 + N_1]\), was shown by Lees to apply to the vast majority of cases in order to derive the final compound, 'complete with primary stress on the first element and secondary stress on the second element.' 37 Compounds involving two nouns: gunpowder, briefcase, dog food, were somewhat differently derived:

(i) the powder is for guns
(ii) ... powder for guns
(iii) ... gunpowder. 38

The various compounds are categorized according to an agent-instrument criterion; for example, in words like blood donor, an animate \(N_2\) does something to or with an
inanimate $N_1$; in compounds like *water pistol*, $N_2$ is inanimate and can be used by an animate subject to do something with $N_1$, and so on. Lees considers that these phenomena tend to confirm C. J. Fillmore's hypothesis that the base component of universal grammar consists of primitive case relationships in which concepts such as "subject" and "direct object" are absent.

6.8 Modern grammatical theory thus suggests that lexicographers should rethink their traditional notions about part of speech categories such as "adjective" which tend to be used far too freely in dictionaries. To suggest that the words *dress*, *spring* and *crash* in the expressions *dress suit*, *spring salmon* and *crash programme* are adjectives merely because they are distributively equivalent, in this position, to *red* (suit), *fat* (salmon) and *stupid* (programme) is to take an excessively superficial view of their true functional role. While it may be simpler to assign such words to the conventional categories which seem to describe them, it appears that their true function would be better denoted by a more neutral term such as "attributive." As we have seen, in bilingual dictionaries it is essential to use great caution with regard to part of speech labels which are sometimes chosen on the basis that the equivalents in the
other language have a clearer function. Thus, many
dictionaries give English sport as an adjective trans-
lated in French by sportif; in the absence of any other
indication, this suggests to the French-speaking user
that the two words behave similarly in equivalent
syntactic situations, and that therefore it should be
possible to modify sport with words like very and
extraordinarily or even to endow it with a set of
comparative and superlative affixes (-er, -est). The
French user, moreover, would be no further enlightened if
he took the trouble to check his word sportif in the
French-English section where he would discover that the
English equivalent is sporting.

7.0 The Bilingual Dictionary
as a Model of Lexical
Competence

The dichotomy langue/parole, first distinguished by
de Saussure and subsequently analysed by Chomsky as a
generative process known as competence/performance, is of
direct concern to lexicography because it enables us to
form a clearer notion of the nature of lexicographical
information. For Chomsky (1965), the competence model of
language postulates "an ideal speaker-listener in a
completely homogeneous speech-community, who knows its
language perfectly"; performance is characterized by "the
actual use of language in concrete situations." But
Chomsky diverges from Saussure's conception of langue as "basically a store of signs with their grammatical properties" which lacks a scheme for "rule-governed creativity" and which cannot account for the "recursive processes underlying sentence formation." In fact, Saussure regarded sentence formation as a matter of parole rather than langue, and the distinction between the two is not always clear although the latter is intended to depict the system of language and the former the act. Thus, for Saussure the characteristics of langue exist as a reality in the brain of each speaker, not as an abstraction which permits him to "create" language as Chomsky's competence model purports to do.

7.1 The Saussurean view of langue is an inventory of signs and their grammatical properties, a tangible reality and not an abstraction; if this position is acceptable it would seem to apply directly and precisely to dictionaries. However, a number of factors tend to obscure what might otherwise seem to be a fairly clear issue. The subject-matter of a bilingual dictionary, for example, considered as the expression of the combined lexical skills of a representative sample of bilingual individuals (as well as monolingual native informants for each language) would seem rather to be best described by parole, actual language acts or performance. On the
other hand, since performance is perceived as an individual affair and dictionaries generally far exceed the lexical capacities of any given individual speaker of the language, it might also be argued that we are dealing instead with a theoretical competence. Moreover, since no bilingual dictionary, no matter how extensive its coverage, attempts to present every possible lexical realization, the materials it contains must be either a representative sample of the total lexicon or else some kind of abstraction of a theoretical (and as yet unspecified) total. Also, if the inventory in a dictionary is a sampling which nevertheless exceeds individual performance capabilities, then it must in some way reflect the manner in which lexical competence operates; if it is an abstraction of some total lexicon, then it must represent the system of the language rather than its actual functioning.

7.2 For a number of reasons, dictionaries cannot be considered as an inventory of linguistic signs in the manner of Saussure's conception of langue. In the first place, even at the level of the discrete lexical units which constitute the entry words of a dictionary, we are dealing with an abstraction because the elements selected for the canonical form are a representation, based on conventional graphic structures, of a base element from
which the user is expected to be able to generate the configurations which make up the total paradigm and which are appropriate to the context of an actual utterance. The addition of other elements of information, such as the principal parts of verbs, irregular plurals, gender flexions and so on, may be regarded as a pedagogical adjunct constituting the minimal requirements for depicting the system of the language and not as an attempt to characterize the totality of the practical realizations of each word in the dictionary.

7.2.1 Another aspect of lexicographical description, phonology, is also a feature which can usually be considered an abstract representation. Those dictionaries which choose to employ a phonemic transcription of pronunciation, rather than a phonetic respelling, are in fact presenting a potential scheme from which allophonic variations may be generated. Even a narrow phonetic representation is an abstraction in the sense that the pronunciation of separate words rarely, if ever, corresponds to the realizations in context. Even more significant in bilingual lexicography is the fact that the equivalent lexical units provided, rather than possessing a functional contextual value, have been isolated on the basis of semantic correspondences which are not invariably precise and unequivocal. They are a
guide to probable equivalence in some contexts but
certainly not in all; for example, we may propose that
the equivalent of French jaune is yellow, but only
provision of the actual example will enable the non-
French speaker to extrapolate the meaning of rire jaune
(to laugh on the wrong side of one's face).

7.3 Our characterization of bilingual lexicography thus
seems to conform more closely to the generative notion of
competence than to Saussure's langue since, like Chomsky,
we cannot conceive of the material of a dictionary as
being a simple storehouse of tangible elements (to use
Saussure's description) which constitute langue. The two
lexicons of a bilingual dictionary may be idealized as
the lexical system which is at the disposal of speakers
of either language and which constitutes a part of an
individual's total competence. It is generally conceded
that the average native speaker commands a combined
active and passive vocabulary of about 20,000 items from
which he is capable of producing a theoretically infinite
number of different sentences (where different means
structurally different because there are diverse ways of
saying essentially the same thing). But a dictionary,
even of comparatively modest proportions, may contain
anything from 40,000 to 200,000 entries (varying
according to individual judgements on what constitutes an
entry and therefore a "word" in the sense usually intended), and even these figures fall far short of the probable total lexical inventory of languages like French and English which may reach ten times this quantity if all technical terminology is included and if a consensus could be reached on the definition of the term "lexical unit." Furthermore, among the theoretical 20,000 items known by an average speaker there may be many hundreds which contain senses completely unknown to most speakers. For example, the word *glide* as used in phonology will be comprehended only by linguists or those capable of seeing the analogy with the same usage in music and other technical activities, and even fewer speakers will be aware of the metallurgical sense: "plastic deformation of one part of a metallic crystal relative to the other part due to shearing action" (RHD). For some linguists, the separate senses of polysemous words constitute discrete lexical units, a view which would augment the total lexicon by a factor of about 10 and further widen the gap between individual performance and theoretical lexical competence.

7.3.1 But just as no grammar attempts to describe all the possible strings of the language, depicting instead the system which will permit them to be generated, so also dictionaries illustrate the manner in which lexical
competence operates by providing abstract representations of a sample which reflects the probable combined knowledge of an entire speech community. Similarly, just as the totality of well-formed sentences of a language is a theoretical possibility, so also the total lexical inventory is a potential reality which is never manifested in any individual.

7.3.2 Individual lexical competence is a relative phenomenon. While the vast majority of the speakers of a given language can operate very efficiently with a comparatively restricted vocabulary, some individuals, because of superior intelligence or the nature of their occupation and education or training, may control quite extensive areas of the theoretical general lexicon which are entirely ignored by their peers. They may carry on a conversation, employing the jargon of some technology, which is phonetically and syntactically recognizable by other speakers but semantically meaningless to them. Still other individuals may control only a few hundred lexical items and yet their linguistic competence permits them to express their thoughts and comprehend the utterances of others without difficulty. Conversely, the learner of a second language usually discovers that mastering large numbers of vocabulary items is no guarantee to complete control of the language; this is
because those elements are usually solicited in dictionaries and word lists which are themselves abstract representations of the total lexicon.

7.3.3 Because of his general linguistic competence, the native speaker is able to generate an indefinitely large number of the possible strings of his language. But certain elements of language are not predictable simply on the basis of a knowledge of the surface meanings of the individual units which make up a sentence or phrase. For example, returning to the expression mentioned above, rire jaune is a "locution figée" based on a particular metaphor whose sense is unpredictable and must be learned independently of the elements of which it is composed, in much the same way that new and individual lexical units must be internalized as they are encountered. Whereas the total grammar of a language is mastered at a very early stage in intellectual development, its lexicon is acquired by a process of continuous augmentation through initial education and as a result of experience of the real world associated with linguistic symbols. Once the "common core" of vocabulary has been acquired, the individual has reached a sort of plateau beyond which he may or may not proceed, depending on personal motivations, intelligence, experience, opportunity and so on.

7.4 Expressions like rire jaune are entered in a
dictionary not just because they are non-predictable elements of the language, but also because the native speaker must be prevented from attempting a literal translation on the mistaken assumption that the other language operates in the same way with respect to this kind of item. The non-native speaker clearly needs the information because it introduces him to the solution of the problem which this unusual combination of otherwise fully comprehensible elements poses for him. In such circumstances the dictionary might be described as a minimal performance model, providing instances of the actual functioning of lexical units in real situations.

7.5. However, there are other even more compelling reasons which seem to support the proposition that dictionaries in general, and bilingual dictionaries in particular, should be considered a reflection of lexical competence rather than performance. We have already observed in earlier discussions (Chapters 4, 5 and 6 of this section) that the overwhelming weight of evidence suggests that lexicographical material is an abstraction of real language and that in realized situations the information undergoes various transformations and yields to the overall meaning of the broader linguistic elements—syntagm, phrase, sentence and text. V. Gak (1970) summarizes the reasons for this situation, noting
that dictionaries describe the system of language rather than its actual operation:

En détachant artificiellement les mots de leur contexte, tout dictionnaire, par définition, ne présente le lexique qu'au niveau du système, de la langue (p. 103).

In monolingual lexicography the definitions and synonyms provided are generally not capable of replacing the headword in every situation. Their role, in fact, is not to achieve this but rather to perfect the "semic equation" by the accuracy of definitions based on extensive analysis of words in a variety of contexts and in all their lexical relationships with other words.

7.6 The equivalents provided by bilingual dictionaries, on the other hand, purport to do exactly this, that is, to replace the source language word by a target language word in contexts which, ideally, are clearly specified. This is the goal of the dictionary designed for composition (or production, French "thème") serving the user who wishes to compose a text in the non-native language. A bilingual dictionary designed for interpretation (translation from the second language, French "version") is consulted to seek the meaning of words in the non-native language text, in which case the mother-tongue equivalents function as definitions on a similar basis to those of a monolingual dictionary. The
two kinds of dictionary may be schematized thus:

(1) \[ \text{TRANSLATION/COMPOSITION} \rightarrow B \]

(2) \[ B \rightarrow \text{TRANSLATION/DEFINITION} \rightarrow A \]

(\text{where A is the mother tongue})

7.6.1 But the abstract characteristics of a bilingual dictionary become clearly apparent when the user attempts to apply the information gleaned in the dictionary to real situations. The success of the application is related to the complexity of the semantic meanderings of the lexical unit concerned—in fact, success stands in inverse relationship to the relative frequency of the word. In statistical terms, it is a well-known fact that the words of lowest frequency develop the fewest number of senses, and therefore, since the phenomena of statistical frequency constitute a language universal, the equivalent form in the other language will tend to exhibit similar quantitative features. Hence, a low-frequency word should stand in a monovalent relationship to an equivalent term of another language and augment the possibility of the success of the pair as translation equivalents in real contexts. Conversely, lexical units of high relative frequency will enter into polyvalent combinations with units from the other language, rendered more complex by the nature of polysemy, and the chances
of a successful realization in context will be proportionately reduced.

7.6.2 The bilingual dictionary offers multiple equivalents and attempts to reduce the consequent ambiguity by furnishing context labels or exemplifying strings as an aid to the transfer from langue to discours. However, unless large quantities of such material are included, there is little chance that anything like the totality of the realignations of a lexical unit will be attained (except in the case of the low-frequency words just mentioned). Thus, for example, the English word lump may have the following set of equivalents in French:

\[\text{lump} = \text{morceau, masse, motte, bloc, bosse, grumeau, etc.}\]

A first attempt at contextual disambiguation of the alternative translations would indicate that morceau relates to "contrived pieces" such as sugar; masse refers to comparatively large quantities; motte is used for soil or mud; bloc is reserved for ice; bosse is a raised portion on skin, metal, etc.; grumeau refers to congealed particles in liquid food. Even when these definitions are supplied in the form of metalinguistic labels attached to the equivalents, they must still be considered abstract representations. If we wish to say: "thinking of that sad event raised a lump in my throat,"
none of the equivalents will serve the requirements of this particular context and the translation would be something like: "le souvenir de ce malheureux événement me rendait triste." Similarly, a "lump sum" is somme globale or somme payée en un seul versement; "in a lump" is en gros, tous ensemble, etc.

7.7 When the translator attempts to apply to real situations the equivalents he has solicited in a dictionary he usually finds that the transfer from langue to discours is not as simple and uninhibited as the lexicographical matching of lexical pairs seems to imply. Language teachers have long been aware of this situation in exhorting second-language learners to avoid using composition dictionaries altogether in the early stages and to employ them cautiously at a more advanced level, and always in close collaboration with the monolingual dictionary for the non-native language. The so-called direct method of language instruction also advocates this approach by implication. The problem, as we have seen, is that the translations provided in a dictionary are neither functional nor situational equivalents but are basically semantic correspondences at the level of concepts.

7.7.1 If we attempt a componential analysis of various equivalents from a bilingual dictionary, several facts
become immediately more visible. In the first place, in accordance with the remarks made above with regard to the quantitative features of lexical units, the components of a given pair coincide precisely only in the case of rarely used words or technical terminology; for the majority of word pairs, there is only approximate parallelism. Secondly, the discrepancy in semic structure revealed by the analysis suggests that in contextual situations, where the individual semic elements enter into various relationships with the individual semic elements of other words, equivalence is likely to be even less clear. The feature matrix in Table V compares the semic elements of an English word, "break," with the semic elements of the French equivalents taken from a bilingual dictionary. 47

7.7.2 The matrix reveals that one of the equivalents, *casser*, is very close to "break" in terms of semic structure: it lacks the notion of "putting an end to" (a habit, etc.) and that of "revelation" (of news, etc.), but has in addition the sense of "annulment" (of a marriage, etc.), "demotion" (of an officer) and "reduction" (of values, prices, etc.). Of the other equivalents, *rompre* is the next closest, being marked for six of the features, followed by *bris*er, which has four, and the remainder which have only one or two features in
<table>
<thead>
<tr>
<th>Feature Matrix for &quot;Break&quot;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>---------------------------</td>
</tr>
<tr>
<td>BREAK</td>
</tr>
<tr>
<td>casser</td>
</tr>
<tr>
<td>briser</td>
</tr>
<tr>
<td>rompre</td>
</tr>
<tr>
<td>interrompre</td>
</tr>
<tr>
<td>amortir</td>
</tr>
<tr>
<td>troubler</td>
</tr>
<tr>
<td>couper</td>
</tr>
<tr>
<td>disloquer</td>
</tr>
<tr>
<td>enfreindre</td>
</tr>
<tr>
<td>révéler</td>
</tr>
<tr>
<td>dompter</td>
</tr>
</tbody>
</table>
common with "break." It is apparent that, as the senses of "break" become more specialized or denote figurative and technical meanings, the degree of semantic overlap with equivalents becomes proportionately smaller. These are precisely the areas in translation which cause the greatest difficulties and which therefore require the most careful treatment in dictionaries so that, for example, an English speaker involved in composition is not left in any doubt that the expression "break in a horse" or "break the news" cannot be rendered by casser.

7.8 This kind of analysis, admittedly easier for content words than for more abstract lexical units, sheds light on one of the essential problems in translation—the notion that the material of a bilingual dictionary is an abstraction due to the fact that dictionaries are characteristically models of "lexical competence" and that therefore the proposed equivalents may not be realized in performance exactly as implied by the lexicographical equivalence relation. Exact equivalence, this analysis tells us, is virtually non-existent, even between pairs of closely related languages. Although some word pairs may appear to be very close in semic structure, they almost invariably differ when taken from the artificial setting of the dictionary and returned to the context of situations. Still others reveal only a
vague correspondence when measured in terms of semic features. This is therefore a theoretical explanation of why the lexicographical practice of using explanatory labels is really an indispensable procedure and one which probably requires to be much more fully exploited than it is in traditional methodology. It also makes clear why the equivalents provided in a bilingual dictionary will not invariably function in real contexts in the manner which their association implies, even in those cases where semic equivalence is close. The word casser, for example, possesses nine of the twelve semic features isolated in the matrix for "break" and three additional features not accounted for in the matrix. Its use as an equivalent will therefore be restricted to that extent as well as by the fact that certain English expressions involving specialized uses of "break," because of the overall meaning of the phrase, cannot be rendered by casser even though the implied semic features appear to be common: break wind (péter, lâcher des vents); break news (révéler une nouvelle); break a five-dollar bill (changer un billet de cinq dollars); break a horse (dompter un cheval); to be breaking one's neck to do something (s'acharner à faire quelque chose).

8.0 Conclusion

Modern grammatical theory has analysed, as we have seen, many of the semantic and syntactic difficulties
which lexicography has so far been unable to resolve. The examples provided in this chapter for verbal and adjectival elements could be increased to cover most of the problem areas of dictionaries and lexicographers would do well to heed the suggestions of generative grammarians in this respect. Further evidence has also been derived from transformational-generative grammar to support the notion that bilingual dictionaries are an abstraction representing "lexical competence" and that the equivalence relation expressed in them should be viewed in this light. We will now enter into a brief discussion on various features of bilingualism which are of significance to bilingual lexicography and then illustrate by practical examples in Part II the methodology which the theoretical analysis demands.
Notes to Chapter 7


2 Greimas (1966) also uses equivalent terminology: "sèmes nucléaires."


9 Ibid., p. 126.

10 See the discussion in Chapter 5 (above), section 4.0 as well as Part II, Chapter 11.

11 Chomsky (1965).

12 Ibid., p. 150.


14 J. Katz and P. M. Postal (1964) also typify the lexical unit in these terms.


17 Chomsky (1957), p. 75.

18 Ibid., p. 39.

19 Chomsky's examples, ibid., p. 75.
23 Chomsky (1965): "I see no way, for the present, to give a thoroughly satisfactory treatment of this general question" (p. 190). "I see little basis for choosing among these alternatives at the moment; until sharper criteria are discovered, these can only be regarded as essentially notational variants" (pp. 191-2).

24 Another discussion on the question of the status of lexical units and phrasal verbs will be found in Part II, Chapter 11.


26 Ibid., p. 72.


28 Ibid., pp. 76-8.

29 Dictionary of Canadian English, Gage, Toronto, 1966; Senior Dictionary.

30 Example from CD1.

31 The Gage series of school dictionaries (DCE) has attempted to resolve some of the problems posed here by using more specific labelling procedures and distinguishing between true adjectives and nouns used attributively; the latter will be called "adjectival."


34 See the discussion in Part II, Chapter 11, on how practical lexicography deals with the problem.


37 Ibid., p. 139.
38 Ibid., p. 140.

39 Ibid., pp. 140-141.


41 Chomsky (1965), p. 4.


44 This question is discussed in more detail in Part II, Chapter 11, section 4.0.


47 This is a development of a procedure suggested by Gak (1970), p. 107.
References for Chapter 7


A bilingual dictionary addresses itself (in addition to teachers and specialists) to bilingual individuals, or else to those who wish to become bilingual. In a theory of bilingual lexicography, therefore, it will be necessary to account for bilingualism, to study its mechanisms and to determine the degree to which lexicography can learn from the findings of psycholinguistic research.

1.0 The Problem of Definition

As with many linguistic concepts, the definition of bilingualism is a controversial subject and there is a wide divergence of opinion on its true nature. For some scholars, bilingualism is essentially the equal and advanced mastery of two languages—a somewhat rare phenomenon; for others, it can be applied to any situation where two distinct patterns of language behaviour are observable in the same individual. The following samples of opinion serve to illustrate the point:
L. Bloomfield (1936): "the native-like control of two languages" (Language, p. 55).

H. Hormann (1970): "a bilingual person is one who expresses himself as efficiently and fluently in two languages as others usually express themselves in their native tongue" (Psycholinguistics, p. 197).

E. Blöcher (1910): "By bilingualism we must understand the membership of an individual in two language communities, to the extent that it is difficult to determine ( . . . ) which is his native tongue" (Zweisprachigkeit; Vorteile und Nachteile, p. 665).

J. Marouzeau (1961): "Bilingualism: the fluent use of two languages by an individual or community, without a marked aptitude for one rather than the other" (Lexique de la terminologie linguistique).

P. J. Meyer (1945): "Bilingualism implies by its very nature two views of the world; it constitutes a dual method of observing and understanding reality, which inevitably leads to two different images of the world corresponding to two respective linguistic forms" (Moedertaal en Tweetaligheid, p. 38).

M. Halliday (1964): "Bilingualism is recognized whenever a native speaker of one language makes use of a second language, however partially or imperfectly" (The Linguistic Sciences and Language Teaching, p. 77-8.)

W. F. Leopold (1939-50): "Bilingual: speaking two languages interchangeably. The ideal form is when both languages are spoken equally well for all purposes of life. In practice, only approximation to this ideal can be expected" (Speech Development of a Bilingual Child, Vol. I, p. 5).

E. Haugen (1969): Bilingualism is understood to begin at the point where the speaker of one language can produce complete meaningful utterances in the other language" (The Norwegian Language in America; a Study in Bilingual Behavior, p. 7).

R. Hall (1952): "We might perhaps say that bilingualism does not begin until the speaker has at least some knowledge and control of the grammatical structure of the second language, as opposed to acquaintance with nothing but the individual items of vocabulary" (Bilingualism and Applied Linguistics, p. 16).
W. F. Mackey (1962): "It seems obvious that if we are to study the phenomenon of bilingualism, we are forced to consider it as something entirely relative" ("The Description of Bilingualism," JCLA, 1962, p. 51).

U. Weinreich (1963): "The practice of alternately using two languages will be called bilingualism and the person involved bilingual" (Languages in Contact, p. 1.

L. Grootaers (1948): "Bilingualism can be classified in three ways: individual bilingualism, social bilingualism, stylistic bilingualism" (Tweetaligheid, p. 293).

N. Brooks (1960): "Bilingualism cannot be defined as the ability to speak two languages, nor can it be considered as a stage eventually reached after prolonged experience with a second language. Rather it implies the presence in the same nervous system of two parallel but distinct patterns of verbal behaviour" (Language and Language Teaching, p. 40).

1.1 Most recent attempts at describing the phenomena of bilingualism have adopted a much less restricted definition than the "balanced-and-equal-mastery" idea, although few investigators have gone far into the question of the different kinds of bilingual behaviour, the manner in which bilingualism is acquired, the distinction which must be drawn between language and dialect, the measurability of bilingualism or the linguistic changes undergone when one language is in contact with another. M. Beziers and M. Van Overbeke (1968) distinguish between three different kinds of definition—descriptive, normative and methodological and offer their own description which, they claim, takes into account the psychological, sociological and linguistic mechanisms which come into play: "Bilingualism is a dual
method of effective communication, either necessary or voluntary, between two or more different communities using two distinct linguistic systems" (p. 133).

W. F. Mackey (1956), on the other hand, maintains that the solution to the problem is to consider bilingualism not as an absolute but as a relative concept—"not "is a person bilingual?", but rather "how bilingual is he?". Since there is no adequate means of measuring how well the second language is known, he argues, bilingualism should not be limited to a question of degree but should include the kind of activity. Such a description "would put the subject on a more stable theoretical basis and would open the way to a systematic measurement of the degree of bilingualism. It would lead to classifications including (. . .) number of languages involved (. . .) type of languages used (. . .) influence of one language upon another (. . .) degree of proficiency (. . .) vacillation between the languages (. . .) social function" (p. 8).

1.2 The question remains open, although it may fairly be claimed that recent research has produced a more scientific approach to the description of bilingualism and even its measurement, since Mackey's work indicates that it is a quantifiable concept. From the point of view of bilingual lexicography, bilingualism is a
theoretical ideal which the dictionary attempts to
describe in terms of an abstraction of the performance
expected of bilingual individuals. The matching lexical
units, set phrases or exemplifying strings are intended
to depict the materials which the bilingual individual
has at his disposal and which may be employed in a vast
array of utterances and with varying degrees of success,
depending on the competence of the individual speaker.
Bilingual lexicography therefore postulates an ideal
speaker/listener situation in which mastery of the two
languages concerned is balanced and equal; at the same
time, this does not preclude the possibility of individ-
ual differences since all dictionaries attempt to portray
a theoretical ideal of language while recognizing that
each individual will control only small and varying
portions of it.

2.0 First Language Acquisition

The processes of bilingual behaviour will also be of
interest to bilingual lexicography because they should
indicate the kind of methodology to be adopted in
dictionaries. Before studying this question, however,
let us first turn our attention to the processes of
language acquisition. Early theories of behavioural
psychologists like Thorndike and Pavlov concentrated on
the investigation of stimulus-response situations and the
observable conditions for their association. B. F.
Skinner (1957) rejected insight altogether and concentrated on the relationship between response and reinforcement to develop his theory of "operant conditioning" which postulates that an increase in "operant strength" (learning) depends on the breaking up of the information or behaviour to be taught at each step into the smallest possible size, the involvement of the learner and the immediate confirmation of the learner's success. These notions exercised an important influence on the design of pattern-drill methods of language learning and on the construction of linear teaching programmes. In practical lexicography the concept of reinforcement has long been considered an important pedagogical adjunct manifested in the repetition of grammatical features (such as gender markings), the provision of exemplifying material, usage notes and so on.

More recently, "neobehaviourists" have attempted to isolate processes occurring between the stimulus-response-reinforcement stages, underlining the role of groups of reactions and their association with the complexes of environmental stimuli. By contrast, Gestalt psychologists recognize the importance of the learner's insight, reinforced by his intelligence and experience, and emphasize the total environment as central to the learning process. They also stress the need for the
learner to recognize a complex goal to which he has favourable attitudes.

2.1 J. B. Carroll (1964) postulates three interrelated sequences of development in child language learning—
cognitive development (the capacity to recognize, identify, discriminate and manipulate items of the real world environment); discrimination and comprehension of speech; production of speech sounds and sequences conforming more and more closely to the patterns of adult speech utterances. Although the manner in which grammar is finally acquired is still not clearly understood, studies indicate that the child is continually making a sort of unconscious linguistic analysis, trying various patterns until he discovers those which are accepted and which achieve the desired response. Carroll contends that, whereas animal communication systems are passed on by heredity, unchanged from generation to generation, human language is always learned from an initial "tabula rasa" state; the only hereditary element is that a normal child is born equipped with the complex neural physiology necessary to learn and produce language. This is the empiricist view of language acquisition.

2.2 Conversely, Chomsky (1967) maintains that the facility with which children are able to infer the grammatical rules underlying the speech they hear and
then to apply these rules for the generation of
indefinitely many novel utterances suggests that humans
are born already equipped with a knowledge of universal,
formal principles which determine the grammatical
structure of the language they eventually acquire. This
restatement of the doctrine of "innate ideas" is the so-called rationalist hypothesis. Studies on the neuro-
physiological basis for language have attempted to
provide tangible evidence that there are valid anatomical explanations for this innate human ability to master
language. Ch. P. Bouton (1974) suggests that if an
individual indulges in introspection about his speech he
usually explains it as a result of having acquired a
knowledge of the grammar of his language through educa-
tion and that this knowledge is precisely what conditions
his use of the language. But self-analysis of this kind
overlooks the neurophysiological contribution to language acquisition:

Cette réflexion intellectuelle sur le langage masque le fait que le comportement verbal dépend de mécanismes, plus ou moins profondément automatisés et relevant de l'activité du cerveau, en particulier du cortex. Ces mécanismes seuls conditionnent l'élaboration et la production du discours, au moment de l'émission, aussi bien que son interpré-
tation, au moment de la réception. La compétence verbale de chaque sujet se ramène donc à un comportement dynamique intégré au plus profond de l'activité nerveuse supérieure (p. 101).
2.4 Some linguists have taken issue with the "innate language acquisition system" propounded by E. Lenneberg (1967) and claim that recent studies of language development, largely influenced by Chomsky's views, start at a period of life when syntactically structured speech is just beginning, thereby ignoring the important early stages from birth to this point investigated by Leopold (1939-50) and others. They would devote more attention to the communicative environment in which the child acquires his language, without reverting entirely to the stimulus-response formulation of behaviourists, and place more emphasis on sources other than the utterances to which the young child is exposed—such as the feedback he receives from his own everyday communication acts.

3.0 Second Language Acquisition

If the acquisition of the mother tongue proceeds according to an almost universal pattern (whether or not there is agreement on the psychological processes involved), the same cannot be said of second language learning where the patterns vary considerably.

The bilingualism that results when a child acquires two languages simultaneously during infancy is usually referred to as "natural" bilingualism and has been studied in depth by Leopold (1939-50) and by J. Ronjat (1913). More recent investigations have subdivided natural bilingualism into three levels: (1) where the
child understands but does not speak the second language; (2) where the first language predominates and the second is less well-known or only used in certain situations; (3) where the child has perfect control over two languages and uses both constantly.

3.1 A second kind of bilingualism is known as "childhood" bilingualism which results when a second language is acquired, during early childhood, after practical efficiency has already been well established in the first language. This is generally agreed to be a favourable time for learning a second language since it does not compete directly with the mother tongue, as it may during earlier periods, and the child is still at a mentally "plastic" stage of development.

Finally, there is "adolescent" or "adult" bilingualism, the most artificial variety, which is acquired much later in life and is characterized by neglect of the formal system of language and overemphasis of content. Adults can more readily solve the intellectual problems in language learning than young children but have less aptitude for the repetition and rote memorization that is inevitable and often enjoyed by the young.

3.2 Objective methods for studying second language acquisition are difficult, if not impossible, to determine and introspection on the part of bilingual
individuals is probably the only way in which the psycholinguistic processes can be exposed to analysis. One fact which has always been obvious to language teachers is that certain individuals have more aptitude for learning a second language than others, and that this aptitude is not necessarily related to intelligence. Some have determined a correlation between musical ability and language aptitude but, as Carroll (1964) points out, there is no scientific proof for this popular notion; language aptitude tests should rather be based on language-specific parameters such as measures of mimicry, cohesiveness of speaking, rote memory, phonetic discrimination, ability to develop meanings inductively, and so on.

3.3 Other factors which are significant in determining the effective acquisition of a second language are:

(1) the age at which it is begun—a crucial matter, since aptitudes, opportunities and motivations for learning vary enormously at different ages; (2) motivations for its use, especially at the post-childhood stage when there must be sound justification for the considerable effort involved; (3) the attitude of the learner towards his two languages, which in large measure determines the permanency of the bilingual state.

3.4 Second language acquisition can also be related to
neurophysiological development. As Ch. P. Bouton (1974) notes, such a study explains why adolescents or adults attempting to learn a foreign language have much greater difficulty than a young child who can acquire two languages with the same ease with which everyone learns their first one, because he is doing so at a stage in his mental development which proceeds parallel to the mechanisms of language acquisition:

Pour acquérir la langue de son milieu, l'enfant bénéficie d'étapes ontogéniques privilégiées si bien que, dès les premières manifestations de son activité verbale, il dispose d'automatismes préalablement élaborées. Toutes les étapes de l'acquisition sont ainsi étroitement associées aux étapes du développement neurophysiologique de l'enfant et marquent, en même temps, les phases de son développement mental (p. 167).

3.5 Probably the most important factor in second language learning at the post-childhood stage is the intervention of the mother tongue which predisposes the learner to restructure the new language in terms of the existing one—a problem which does not exist for the young child. There is thus a progression from language to experience, rather than the reverse which occurs in first language acquisition.5

4.0 Psycholinguistic Description of Bilingualism

The processes of bilingual behaviour have been the
subject of considerable psycholinguistic research aimed at establishing a model to depict various kinds of bilingualism. In this regard the work of C. E. Osgood and J. J. Jenkins (1965) on learning theory is usually taken as the point of departure to account for the activity of human communication in more than one code. They describe language decoding as the process whereby patterns of stimulation (auditory or visual) elicit certain representational mechanisms (ideas or meanings) through the operation of a complex nervous system. Language encoding, on the other hand, is the process by which a speaker's intentions are coded into vocal sounds for a given language (expression of ideas) by a gradually changing series of stages.

4.1 It can be assumed that the processes described by Osgood and Jenkins' model for decoding and encoding operations in human communication, based on a "mediational" learning theory, will apply equally well to a monolingual as to a bilingual child exposed to more than one language from birth. However, for those who begin to learn a new language after practical efficiency has already been well established in the first language, new sets of decoding and encoding habits will be formed and act in direct conflict with the original ones. A model to depict the two kinds of bilingualism has been
developed by C. E. Osgood and S. M. Ervin (1965). The situation where a second language is learned at a later stage in life typically involves a constant perceptually experienced object which merely requires the addition of a new sign in the other language. Thus, for example, the object HORSE is associated with two linguistic signs—horse/cheval for French or English speakers—and they come to believe that the two signs must have the same meaning. This learning method is what Osgood and Ervin refer to as a compound language system, where two linguistic signs (or codes) are associated with the same set of representational mediational processes (or meanings) and which they depict schematically thus:

This is the typical mature foreign language learning situation where traditional instruction encourages the learning of pairs of vocabulary items by associating a sign from language B with a sign and its meaning from language A. It is also the procedure exemplified by many
bilingual dictionaries which encourage the naive belief that pairs of signs and their meanings are more or less parallel in any pair of languages. The bilingualism acquired by a child who is reared in an environment where the two languages are constantly interchanged by the same people in the same situations would also be described as a compound system.

4.2 Natural bilingualism, conversely, is presented by Osgood and Ervin as a coordinate language system in which two sets of linguistic signs are associated with different sets of representational mediating processes:

\[
\begin{align*}
\text{S}_A & \rightarrow \text{rm}_1 \rightarrow \text{sm}_1 \\
\text{S}_B & \rightarrow \text{rm}_2 \rightarrow \text{sm}_2 \\
\text{R}_A & \rightarrow \text{rm}_1 \\
\text{R}_B & \rightarrow \text{rm}_2 \\
\end{align*}
\]

The coordinate system describes the separate but parallel development of two languages in the same individual. Since the total situations and the total behaviour when one language is being employed will differ from those occurring when the other is being used, the kinds of representational processes will be discrete, as also the meanings of the respective signs. This is the ideal
which bilingual lexicography should attempt to reflect by whatever means are at its disposal.

The effectiveness and durability of a coordinate bilingual system depends on a lack of interference, the main flaw in the compound system. It can and does occur in the coordinate bilingual when, for example, cultural similarities cause confusions in the representational processes elicited by the translation-equivalents of the two languages. Both in decoding and encoding, the coordinate bilingual is under constant pressure to confuse meanings and to interpret a sign in one language as its translation-equivalent in the other, and so on. Empirical evidence has been put forward to support these views. 6

4.3 Although the neural areas of the brain which govern various kinds of language behaviour are not readily accessible to investigation, recent studies of aphasia victims have enabled some progress to be made. W. E. Lambert and S. Fillenbaum (1959), for example examined the course taken by aphasia in bilingual patients and were able to theorize that it would depend on factors such as the order in which the languages were acquired, comparative skill attained in each and the affective value each language had for the individual. They speculated that the functional relations of the
bilingual's two codes have some kind of neural representation in those areas of the brain necessary for the functioning of language. Hence, aphasic damage would be likely to affect all language of a compound bilingual but produce a more selective kind of disturbance in coordinate bilinguals. Their analysis of case studies of Montreal polyglots tended to confirm their postulate. W. E. Lambert (1968) concludes that research tends to support the view that coordinate bilinguals have a built-in resistance to interference, whereas compound bilinguals have to rely on cues emanating from language usage contexts in order to minimize potential interference.

5.0 Bilingualism in Lexicography

The findings of psycholinguistic and neurolinguistic research have a clear message for bilingual lexicography. In over 400 years of practice lexicographers have made judgements on the manner in which their materials should be presented based almost exclusively on their own experience and intuitions. Thus, it has become obvious to most editors that a bilingual dictionary must be much more than the simplistic matching of pairs of signs with no attempt to eliminate some of the causes of interference. In fact, bilingual dictionaries which are prepared with less than scrupulously scientific attention
to detail may well be one of the greatest contributors to language interference, at least at the lexical level.

5.1 Psycholinguistic studies of the phenomena of bilingualism thus provide us with tangible evidence to account for the operation of dual language acquisition and use and seem to support the view proposed in the earlier chapters of this section that bilingual lexicography must be based on a theory which typifies dictionaries as models of *lexical competence*. We have observed that the content of bilingual dictionaries is not a simple matching of all the symbols of one language with all the symbols of another.\(^8\) It is not a question of a unique referent which is invariably shared by a pair of symbols, one from each language, because reality tends to be construed in different ways by different language groups. Instead, as the discussion on lexical structure suggested, there appears to be a referent for each symbol and the referent pair is mediated through a common referent which depicts features of each element of the pair; this common referent might be typified as the description that would be given by an impartial judge from neither language group.\(^9\)

5.2 As well as being related to a model of "lexical competence," our conception of the independent and abstracted lexicons of a bilingual dictionary seems to be
very close to the ideal of coordinate bilingualism discussed above (4.2) in which the individual possesses two parallel but separate language systems. He can switch from one code to the other with facility, but finds direct word-for-word translation difficult and usually reverts to paraphrase as the expedient which will enable him to render entire sense groups rather than individual lexical units. Dictionaries, likewise, must be composed of two parallel systems, each integrated within themselves, and in which the attempt to match lexical units from each is an artificially-contrived arrangement—a fact which is emphasized by the failure of word-for-word translation based on the material provided in a dictionary.

5.3 Second language teaching studies have confirmed the view that coordinate bilingualism is the ideal, leading to the promulgation of the so-called "direct method" in which the symbols of the second language are related directly to the referential environment in which they occur in that language rather than being indirectly related through the association of a (roughly) equivalent symbol of the first language. Lambert (1963) carried out practical experiments from which he concluded that the coordinate approach to the acquisition of vocabulary was relatively more efficient than the compound method
because the task of associating new words directly with their own referents afforded greater distinctiveness of the elements to be learned than by associating pairs of symbols from each language with a common referent.\textsuperscript{10}

On the other hand, Lambert also conducted tests which revealed that advanced students who kept their two languages functionally separated during an intensive course of instruction performed more poorly on vocabulary tests than those who permitted interaction of the semantic features of the two languages.\textsuperscript{11} In other words, a tendency towards linguistic interdependency of semantic features seems to assist learners in acquiring the vocabulary of a second language. However, in the artificial context of bilingualism acquired at the post-adolescence level it is not surprising that the "coordinate" group achieved less than the expected results since they were probably unable to dissociate the newly acquired vocabulary from the parallel symbols of the first language (in accordance with the neurological phenomena explained by Bouton (1974) in section 3.4 above). For bilingual dictionaries this may well imply, therefore, that the pedagogical aspects of second language acquisition are not necessarily impaired by the artificial coordination of the lexical units of two languages.
5.4 Natural bilinguals seem to possess a set of distinct connotative meanings for the translational equivalents of their two languages because the lexical elements have been acquired in the independent structural environment of the language to which they belong. Evidence for the discrete nature of the respective vocabularies of natural bilinguals is provided by the fact that "saturation tests" (in which the intensity of the connotative meaning of a lexical item is reduced by systematic repetition) show that the semantic import of equivalent words of the other language is not proportionately weakened, whereas in compound bilinguals this is precisely what occurs.12

5.5 Weinreich (1964) discussed the coexistence in coordinate bilinguals of parallel linguistic systems, using Saussurean semantic theory as the basis for his arguments. Thus, if book and livre are interpreted by a bilingual speaker as possessing independent meanings, they must therefore be two distinct signs belonging to coexistent systems and not just alternative signs for the same reality. If, on the other hand, they are interpreted as having the same meaning they therefore constitute a compound sign and must belong to partially merged systems. Moreover, if one of the signs is interpreted as denoting, not an independent referent at
all, but another sign in the other language, then there is total merging of the two systems with one being subordinated to the other. 13

5.6 We may therefore return to the notion of the translation equivalents in a dictionary being depicted as pairs in which the relationship is mediated through a "generic" referent which combines the semic features inherent in the respective referents of the word pairs, without in any way reducing their autonomy within the independent language system from which they have been extracted. As we have observed in earlier chapters of Part I, it is this autonomy, unimpaired by the contrived methods of bilingual lexicography, that determines the different realizations which occur in contexts beyond the scope of the dictionary. But for practical lexicography, yet again, these phenomena only further emphasize the need to provide adequate explanatory or definitive material in order to avoid the belief which is usually fostered in the naive dictionary user that the equivalents provided are identical in every respect.

5.7 Related to this question is the important matter of language interference, which is demonstrably more prevalent in compound bilingualism than in coordinate bilingualism. Research suggests that natural bilinguals have a built-in defense against interference whereas
compound bilinguals are obliged to rely on cues emanating from contexts in order to minimize potential interference; but since the contexts themselves frequently provide conflicting cues, interlingual confusion is usually increased. Interference is not entirely eliminated from natural bilinguals, however, because cultural similarities may cause similarities in the representational processes elicited by the signs of translational equivalents in the two languages. Thus, in the decoding process, the bilingual is under constant pressure to confuse meanings and to interpret signs in one language as the translational equivalent of signs in the other; the more genetically related the two languages, the more this is likely to occur, so that pairs of cognate words may not necessarily be equivalent and yet be perceived as such.

Bilingual dictionaries are veritable storehouses of language interference. In fact, it may be no exaggeration to suggest that dictionaries prepared with less than scrupulous attention to detail are the greatest contributors to the phenomenon at the lexical level. As we shall see in Part II, there are various methods which may be employed in practical lexicography in order to reduce interference: careful labelling and explanation, indication of deceptive cognates, diversified exemplifying material and so on. These expedients may reduce
interference, but they will never eliminate it; this is the task of the language learner himself as well as the responsibility of his instructors, a responsibility which cannot be entirely shouldered by the dictionary.

The very act of drawing together pairs of lexical items which have been artificially removed from their natural associative context is sufficient to engender confusion of the kind observed in compound bilinguals. Thus, in order to retain the ideal characteristics of coordinate bilingualism, a considerable amount of painstaking attention must be devoted to potential areas of interference in the reconstituted context of lexico-graphically-oriented material.
Notes to Chapter 8

1 Cf. W. F. Mackey (1956).
3 Cf. bibliography in Beziers and Van Overbeke (1968).
6 W. E. Lambert, J. Havelka and C. Crosby (1961), for example, tested the prediction that bilinguals with experience in separated contexts would show comparatively greater semantic differences between symbols in one language and translation-equivalents in the other, more associative independence of translation-equivalents in their two languages and less facility in switching from one code to the other. As well as confirming their hypothesis, their results indicated that bicultural experience accounted for the overall difference between natural and artificial bilinguals. W. E. Lambert (1968) has also tested the distinction between the two kinds of bilingualism using the concept of semantic satiation in which the meaning of a word is diminished in value by constant repetition. Theoretically, in compound bilinguals, this meaning impairment should also extend to the translation-equivalent, but much less so, if at all, in coordinate bilinguals. Lambert's results confirmed the case. Similar kinds of experiment conducted by P. A. Kolers (1968), using colour coding lists, similarly highlighted the posited differences between the two kinds of bilingualism.
7 The thorough studies of Ch. P. Bouton (1974, 1976) may be consulted for exact details of our current knowledge of the anatomical and neurophysiological aspects of language acquisition and functioning.
8 Cf. Chapters 3-6 of Part I.
9 Cf. Chapter 4 of Part I.
11 Ibid., p. 120.
12 Ibid., p. 119.
14 W. E. Lambert (1963), p. 120.
15 Osgood and Ervin (1965), pp. 139-140.
References for Chapter 8


CONCLUSIONS TO PART I

In the introduction to Part I we expressed concern at the lack of a sound theoretical basis for the practice of bilingual lexicography and emphasized the need to examine various themes in general linguistics in order to determine those aspects which might fruitfully be applied to dictionary-making. The view that bilingual lexicography is merely a subset of general lexicography or of translation was shown to be inaccurate and inappropriate and the discussions which followed attempted to reinforce the autonomy of the former as an independent discipline within linguistics.

By locating bilingual lexicography in its historical context, we wished to demonstrate the unique nature of its methodology, developed over several centuries and preceding monolingual lexicography itself. The descriptive analysis which followed aimed at establishing a clear and unequivocal frame of reference for the study of the theoretical and methodological features of existing bilingual dictionaries; it also provided a typological classification which would serve lexicographers as a basis for determining the orientation of their work.

278
Given that lexicography, a branch of the vast science of lexicology, is concerned with word meanings in isolation as well as in context, it is essential for a clear appraisal of current theory to precede a detailed analysis of the semantic aspects of bilingual lexicography. From an analysis of the phenomena of semantic isomorphism we conclude that the existence of complex multiple correspondence between the lexical units of a pair of languages is the most significant element of bilingual lexicography and the one which engenders the greatest difficulties in its methodology. We also determined that the units of manipulation in lexicography, the "words", are complex elements composed of a bewildering array of independent and interdependent components whose various combinations dictate their basic meanings, themselves subject to variations imposed by different contexts. In their dictionary situation lexical units thus constitute a simplified representation of this complex and often intangible composition of semic elements and if this is so for monolingual dictionaries, we are left in little doubt as to the intricacy of the equivalence relation which links the meaningful elements of two languages and which is the object of concern for bilingual lexicographers.

A theoretical foundation for bilingual lexicography, according to the findings of Part I, demands that a bilingual dictionary be construed as an abstract representation of the lexicons of a pair of languages. Central to the whole notion of bilingual lexicography is the requirement that an
equivalence relation be established between word pairs—not an equation of semic elements, as in monolingual lexicography because exact equivalence in a formally balanced equilibrium is virtually unknown. This equivalence relation takes into account the fact that the dictionary context in which the word pairs find themselves is an abstraction of the reality from which they have been artificially isolated and requires that the practical methodology which implements the relation employ all the means at its disposal in order to render this state of affairs transparent to all users of the dictionary. The abstracted elements of a bilingual dictionary are thus a reconstituted arrangement of the semic elements of two lexicons, brought together on the basis of a somewhat imprecise correlation of their respective meanings as derived in contexts which permit a convergence of meaning. In this respect the bilingual dictionary may be said to represent langue rather than parole, and the equivalence relation to typify potential equivalence between word pairs while lacking the capability of depicting all the possible realizations which the innumerable contexts available to a given pair of lexical items may engender. The precision and completeness of the equivalence relation, in effect, is a function of the genetic proximity of the languages under consideration but is complicated by questions such as cultural diversity and physical isolation.

By definition, large lexicographical enterprises such as the Trésor de la Langue Française are based on
distributional analysis in order to constitute the "semic equation" of monolingual dictionaries. Bilingual lexicography must also use this approach, as well as componential analysis, but the most important feature is contrastive analysis of the elements of the equivalence relation which ensures the accuracy and aptness in various contexts of the translation equivalents provided in the dictionary. Finally, as we have observed in this first part, there is much to be learned from transformational generative grammar, which sheds light on the status of the "word" and gives new insights with respect to deep structure relationships between word pairs in contrasted languages, and from psycholinguistic studies on bilingualism which is at the very heart of the translation problem.

We now pass to an examination of the practical methodology which is a necessary consequence of these theoretical considerations. Part II is an attempt to apply a rigorous set of guidelines to the compilation of a bilingual dictionary, from the collection and analysis of the raw materials to the solution of a number of differential equivalence problems and the application of data processing techniques to lexicographical materials.
BILINGUAL LEXICOGRAPHY:
THEORETICAL FOUNDATIONS AND PRACTICAL METHODOLOGY, WITH
SPECIAL REFERENCE TO CANADIAN FRENCH AND ENGLISH.

by

Murray Thomas Wilton

Volume II
PART II
LINGUISTIC METHODOLOGY

CHAPTER 10
THE CORPUS

1.0 It is an axiom of lexicography that the corpus of a dictionary is composed of "words" excerpted from various sources in order to arrive at the components of lexicography—definitions in the case of monolingual dictionaries, translations for bilingual dictionaries. The important thing is the source of the material. It is doubtful whether any twentieth-century lexicographer sets about compiling a dictionary without a ready-made base of some kind, and while it would be unfair to suggest that each succeeding generation of lexicographers plagiarizes the work of its precursors, there is no question that earlier endeavours serve to enhance the work of those who come later. Hence, the would-be dictionary-maker begins with a wordlist as a minimal item of source material.

1.1 This wordlist may be the entire lexicon of another dictionary which will then serve as the basis for deciding the statistically-determined items to be treated in the new work, since frequency must be an important parameter in the selection of material for all but the most extensive dictionaries.1 In bilingual lexicography
it will clearly be necessary to have a dual corpus, composed of two wordlists more or less evenly matched on the basis of frequency, although eventually, when the total dictionary has been compiled and carefully cross-referenced between the two lexicons, this matching should be achieved as a matter of course.

1.2 The compilation of a new monolingual dictionary ideally presupposes that the lexicographer will carry out an exhaustive distributional analysis for each word in order to determine all its current meanings and to establish the "semic equation" by means of accurate definitions composed of clinically objective statements in the metalanguage of traditional lexicography. Except where such research has been continued over a very long period of time and contributed to by many investigators and voluntary readers, the only way in which it can be accomplished is by the use of a concordance-style analysis using computational procedures. Any other method adopted can only be described as being based, at worst, largely on the work of other lexicographers or, at best, on the intuitions and personal experience and expertise of the editors.

1.3 Bilingual lexicographers are in a privileged position with regard to source materials. They may, with impunity, consult the lexicons of a pair of monolingual
dictionaries, one for each language, and eventually arrive at translational equivalents based on the definitions provided. This procedure obviously presupposes an editor, or team of editors, with a profound knowledge of both languages since the two hypothetical dictionaries just mentioned would be, of themselves, useless in the hands of a monolingual editor. Once again, it is highly improbable that the work of other lexicographers in bilingual dictionaries will escape the notice of the compiler of a new dictionary, whether as a source for semantic information or as a model for lexicographical style and structure.

1.4 But new dictionaries are not created merely to replace those which already exist and which may be perfectly satisfactory as they are. Presumably the editor of a new dictionary is impelled by motivations ranging from a belief in the innovatory features of his own conception of lexicography to a concern for the need to provide the most up-to-date material available. The former is a measure of his own linguistic training and practical experience, but the latter will already depend on an extensive analysis of a wide range of documentary source material, including oral encounters with native informants. Bilingual lexicographers would be unlikely to employ computers in order to excerpt the information
they seek, since this would imply carrying out the work of monolingual lexicography for two languages, instead of one, with the main task of finding equivalents still left to totally human intervention. Yet they must continue to do exactly this kind of work in order to expand on an existing corpus or to maintain its contents up-to-date.

2.0 Generally speaking, the most efficient means of gathering information for a bilingual dictionary treating two major national languages with extensive written literatures is to have access to current publications in a number of areas which might be considered germane to the content of a general dictionary. Obviously editors must be selective in taking information from highly technical publications; they must also exercise a certain degree of subjective judgement in choosing material from more popularized literary output where register is the salient feature.

2.1 The search for lexicographical information will thus be initiated on the basis that a general corpus already exists and that it is new words and expressions which are the object of the "dépouillement." To the lexicographers' practised eye, such linguistic novelties should stand out and demand his attention as enhancements to his base corpus. The new material will consist of "new words," the neologisms coined almost daily and most easily
observed in newspapers and popular magazines. Traditional methods, which cannot be improved upon, demand that the selected information should be carefully recorded (usually on file cards or slips or "fiches" of dimensions appropriate to the filing system being used) including sufficient context to ensure the correct interpretation of the meaning and details of the source materials, date, author of the card and so on.

2.2 If the material has been obtained from a monolingual publication, as is most frequently the case, there remains the task of finding suitable translation equivalents; this is the real task of bilingual lexicography. Information gleaned from bilingual publications (of the kind promulgated by the Canadian federal government) can usually be translated immediately, although the existence of a published translation does not necessarily place the stamp of authority or approval on the terminological equivalents provided. Only when the word pairs have been thoroughly checked in documentation and tested on native speakers can an editor take the ultimate step of committing them to the permanence of the printed page of a dictionary.

3.0 Before any excerpting of new material for the dictionary can be undertaken, however, the lexicographer must make important decisions on the kind of corpus on
which to base his work. As Zgusta (1971: 226) says:

The decision as to what texts should be excerpted is a consequence of the lexicographer's two basic decisions concerning the scope and the type of the dictionary: if it is to be an overall-descriptive dictionary, the repertory of the excerpted texts must be broader, in the different dimensions of the variation of language, than if we decide to prepare a standard-descriptive dictionary.

Ideally, it will be necessary to establish a balance between purely literary materials and more colloquial publications typified by weekly journals and magazines aimed at a less sophisticated reading public. There is always the danger that journalistic language will provide what appear to be examples of neology but which turn out to be hapax legomena; only the careful inspection already referred to will eliminate this possibility. On the other hand, it is scarcely credible that anything new will be acquired by restricting the processes of "dépouillement" to established literature which has by now been scrutinized closely by many generations of lexicographers. It may also be argued that popular publications are often written in a "conversational" style which captures the spirit of the spoken language rarely encountered in more formal literature.

3.1 These considerations lead naturally to the question of statistical frequency, since habitual usage is
obviously an important selectional parameter. Existing
word-frequency lists are generally inadequate for lexico-
graphy because they rarely exceed a vocabulary of some
5,000 items, a figure well below the average 60,000
entries of a general language dictionary. Moreover, even
the existence of a more extensive list is no guarantee
that the equivalent forms of the other language will
conform to the same distributional phenomena. The
quantitative aspect of selection from the corpus there-
fore reverts almost entirely to the intuitions and
experience of the lexicographer, reinforced by cautious
verification of a variety of sources.

3.2 It is usually considered a self-evident prerequisite
that the compiler of a monolingual dictionary should be a
native speaker of the language it describes. Since the
ideal of perfect bilingualism is exceptionally rare, it
is highly improbable that the editor of a bilingual
dictionary will exercise equal control over both of the
languages described and it will therefore be essential
that the collaboration of native speakers of the other
language be obtained. These may be engaged as native
informants or as full editorial collaborators with
responsibility for those parts of the second language
which are beyond the competence of the editor because, as
we have observed in Part I, the complex problems of
meaning and equivalence pervade the whole activity of bilingual lexicography and demand the constant attention of speakers of both languages.

4.0. The source materials which make up the corpus of a dictionary depend, as we have noted, upon the aims and scope of the work. In order to illustrate how documentation is selected and obtained we shall use the example of a real lexicographical venture, the *Canadian Bilingual Dictionary Project (CBDP)*\(^8\) and examine, not only the corpus itself, but the manner in which the corpus is manipulated towards the final goal of a published dictionary. The CBDP was established in 1971 in order to carry out a total revision and enlargement of the *Canadian Dictionary/Dictionnaire canadien* (CD1);\(^9\) the revised version is abbreviated here as CD2. There was thus an already existing corpus (CD1) upon which to base the new dictionary (CD2). This corpus was in the form of file cards (fiches) for the entire contents of the published version of CD1, together with a large number of fiches which contained material gathered for the first edition but never used.\(^{10}\) The unused materials constitute, in effect, a second corpus (CD0) and, almost without exception, the information has found its way into CD2. The two corpora, CD0 and CD1, were organized in alphabetical order, the latter being in the form of a
completed manuscript in the style required by the publishers, and were eventually merged to form a single body of lexicographical material.

4.1 The next decision to be made was the manner in which CD2 would be created. In a sense, the combined materials of CD0 and CD1 constituted an enlarged dictionary, but not necessarily an updated one because the data for the first edition was gathered during the mid-fifties up to about 1961. At the time of starting the revision, therefore, this material was already close to twenty years old and could not include the enormous quantity of linguistic information which has been unleashed by a technologically sophisticated and linguistically liberated society.

It had been determined that the new dictionary would have the same general aims as its predecessor; that is, it would concentrate on Canadian and North American reality and constitute a bidirectional translation tool for both of the main language groups of Canada. In scope, however, it would be different. CD1 was a concise edition, of suitable dimensions and content for young language-learners and sufficient for the lexical items most commonly encountered in everyday life in Canada and North America. CD2, on the other hand, was to be a much enlarged dictionary, covering every aspect of Canadian existence and incorporating a fairly high proportion of
more technical and specialized language in order to make it a truly efficient translation tool.

4.2 In order to ensure adequate coverage of the entire linguistic gamut, an intensive study was made of the kinds of major headings under which it would be appropriate to gather information. By establishing a kind of ideological structure it was predicted that, not only would it be possible to cover the field efficiently, but also these subdivisions would facilitate the organization of documentation and the search for translation equivalents.

Probably the best source for the elements of such a structure is the abbreviation list of general dictionaries and encyclopedias which usually contain a broad specification of human knowledge and endeavour. Other fruitful sources include the classification systems of libraries and business directories, commercial and technical guides, The Key to English Vocabulary, Roget's Thesaurus, commercial shopping catalogues, the "table des mots" of Duden Français, the "liste des tableaux analogiques" of the Dictionnaire Quillet de la Langue Française, Fernand Criqui's Mois et Concepts (lexique permanent) and the Library of Congress index. When these documents have been thoroughly investigated, a list of main headings will eventuate, corresponding to
such broad spheres as sociology, sciences, arts, administration, sports and so on. Each of these categories should, in turn, be subjected to further subdivision in order to arrive at a series of more specific components, some of which, like the individual sciences, will constitute independent fields while others, such as particular sports, will remain as subheadings under a main category. The main divisions are the "fields," corresponding to the final abbreviation list of the dictionary (Anát., Agric., Math., Soc., Pol., etc.), while the subdivisions correspond to the explanatory rubrics (Sport--boxing, equitation, golf, etc.).

There is thus a finite set of major headings and an open series of subheadings which are applied as the situation demands. Where possible, a field label is selected which is a generic term and which does duty for both languages, thereby avoiding the necessity to have two completely separate lists.

4.3 There are a number of important reasons for proceeding in this manner. To begin with, most source materials will be in the form of specialist dictionaries or other publications and this kind of organization lends itself to direct consultation of such material. Secondly, a compact file of fiches dealing with, say, "medicine" can be readily consulted and checked, whereas if they
were immediately dispatched to their alphabetical location after production they would be much more difficult to extract. Researchers with expertise in a given field can devote their energies to the perfection of this section of the dictionary; it will thus be possible to avoid unnecessary overlapping of tasks and duplication of output by ensuring that each investigator is assigned to specific fields. Finally, it is possible, even by visual inspection, to determine the dimensions of the component parts of the dictionary. If, for example there were several thousand fiches accumulated for the field "mathematics" and only a hundred or so for "administration," the imbalance would be physically visible and appropriate corrective measures could be taken.

4.4 In the case of CD2 there was a further motivation for the ideological arrangement of the material. It was predicted that the conversion of the lexicographical data to machine-readable form would require a format which would clearly identify semantic elements of the information so that sorting and retrieval programmes could be accomplished on the basis of a finite set of key labels. More remote prognostications also called for the possibility of the conversion of the entire corpus of CD2 to a form compatible with existing word banks with
which it would be in liaison for direct consultation by subscribers. Rapid access would require that the materials be organized according to the semantic categories likely to be the object of interrogation by the subscribers.

4.5 The ideological organization of lexicographical data has an important side-effect. We have observed, in Part I, that various theoretical and pedagogical considerations all point to the need for clearly-defined equivalents, so that the divergences inherent in the equivalence relation are reduced to a minimum. A crucial aspect of practical lexicography, therefore, is the identification of the semantic features which play a part in the equivalence relation and which must be explained in order to avoid ambiguity and confusion. If the lexicographical material has been collected on the basis of notional categories and subcategories, a major part of this task will have been already accomplished, with further specification coming from the metalinguistic definitions of monolingual dictionaries.

5.0 Having established a definitive list of categories, the task of acquiring appropriate documentation can be broached. For CD2, a good deal of reference material was already available as a legacy from CD1, but there were nevertheless many gaps, especially with respect to up-to-
date dictionaries and specialist publications.14

5.1 A dictionary treating French and English must obviously make occasional reference to other bilingual dictionaries, not to plagiarize them, but in order to compare methodology and verify contentious translations. The obvious major contenders in this category (for French and English) are the Harraps and Cassells productions, although much can also be learned by studying the lexicographical procedure employed in the Larousse dictionaries as well as older publications like Bellows, Nugent and Gasc.

5.2 Reference materials pertinent to Canadian and North American reality will clearly be important to a dictionary of Canadian English and French, and CD2 found itself in an enviable position with respect to the readily available bilingual publications of a number of government agencies. The most remarkable of these sources is the Bulletin de Terminologie, published by the federal governments' Bureau des Traductions, which contains an abundant supply of new terminology and technical translation. Another is the publication Actualité Terminologique produced by the Bureau des Traductions and containing useful articles on a variety of translation problems. Several other crown corporations and business enterprises possess their own linguistic service which is
prepared to give information to appropriately deserving projects (Teleglobé Canada, Canadian National Railways, Radio Canada/Canadian Broadcasting Corporation, Simpsons-Sears, and many others). The Québec provincial government operates the Régie de la Langue Française (formerly the Office de la Langue Française) which, through the Editeur Officiel du Québec, publishes various terminological glossaries aimed at regularizing the use of French equivalents in the work world. Many business firms publish bulletins describing their operations in both languages (Royal Bank of Canada Newsletter, Imperial Oil Review, Shell News, etc.), and every federal government ministry promulgates news and information in bilingual publications. This mine of information can be supplemented and verified by means of a number of Canadian published monolingual dictionaries. For English there are the Gage series Dictionary of Canadian English, the important Dictionary of Canadianisms on Historical Principles, and Winston’s Dictionary of Canadian English. For French one may consult the Glossaire du Parler Français au Canada and Louis Belisle’s Dictionnaire général de la langue française au Canada.

5.3 Terminological sources outside of Canada include the Conseil International de la Langue Française (France) which publishes La Banque des Mots, and the Bureau de Terminologie (Luxembourg) which produces Terminologie and provides
subscribers with fiches if required, *International Civil Aviation*, and many others. For neology, apart from regularly published bulletins, there is the *Barnhardt Dictionary of New English Since 1963* and the *Larousse* publication *Les Mots dans le vent*. Among many linguistic publications and translators' journals of use to lexicography are *Babel, American Translator's Association, Méta* (Montréal, Canadian Translator's Association), *Journal of the Canadian Linguistic Association, Traduire, Défense de la langue française*, etc. Other specific publications germane to French-English lexicography are: *Stylistique comparée du français et de l'anglais* (J-P. Vinay et J. Darbelnet), *Les Faux Amis* (M. Koessler), *Mots amis et faux amis* (J. Kirk-Greene), *Communication et traduction* (J. Darbelnet), and numerous others. Finally, there are hosts of specialist dictionaries which may be consulted as well as a variety of publications, many of which are listed in Appendix No. IV which describes much of the documentation used in CD2.
Notes to Chapter 10

The Trésor de la Langue Française is an important exception to this statement, since the huge dictionary is based on statistical samples from a vast concordance covering two centuries of written French texts.

Cf. various publications on monolingual lexicography listed in Bibliography 2, notably J. Rey-Debove (1971).

The Trésor de la Langue Française is the model for this kind of approach.

Bilingual dictionaries involving unwritten languages will clearly use quite different methods based on native informants, concentrated field work, questionnaires and so on.


"The exception, though it is usually the main source of the fundamental material for the dictionary, never suffices in itself, irrespective of the number of the collected slips. The factor of the lexicographer's knowledge of the language is of vital importance"; Zgusta (1971), p. 233.

Cf. the full discussion of this question in Zgusta (1971), p. 235-239.

The Canadian Bilingual Dictionary Project is a publication of The University of Victoria, B.C., and is directed by Professor Jean-Paul Vinay.


Notably for reasons of economy.


See Appendix III for a list of the field labels and their description as employed in CD2.

See the full discussion on this subject in Part II, Chapter 16.
A sample list of reference materials used in the production of CD2 appears in Appendix IV.
CHAPTER 11

STRUCTURAL METHODOLOGY

In chapter 3 of Part I we analysed the descriptive features of a number of bilingual dictionaries in order to arrive at a typological classification of bilingual lexicography. Among the structural properties examined were word order, directionality, range, scope and internal article structure. We shall now turn our attention to the methodology which underlies these features, providing linguistically-based arguments for various procedures which may be adopted in an attempt to postulate an ideal structure for bilingual lexicography.

1.0 Preliminaries - Organization for Particular Users

A lexicographical enterprise cannot begin until the editors have very clear notions of the kind of dictionary they wish to publish. Apart from obvious considerations, such as the languages to be treated, they will be obliged to make decisions on the type of user for whom the work is intended: whether the dictionary is to be designed for native speakers of one of the languages only or of both; whether it is a tool for comprehension (translation
from the non-native language), or for composition
(translation into the non-native language), or for both;
the amount of information to be given; the number of
entries and so on.

The determination of the type of user is thus a
crucial decision, since practically every lexicographical
step taken depends upon this feature. Dictionary users
can be categorized in various ways, according to age,
native language, motivation for using the dictionary and
so on. These considerations are summarized in Table VI
in which we observe that young native speakers of the
source language (SL) are clearly distinguished from young
learners of the target language (TL), that translators
who work from SL, that is, who are native speakers of SL
are different from translators who work from TL, and so
on.

1.1 The question arises as to whether all of these user
requirements can possibly be built into a single
dictionary. In his typological classification, A. M. Al-
Kasimi (1977) rejects this possibility and bases his
claim on the fact that, in spite of the statements made
by various lexicographers, most linguists declare that it
is impracticable to attempt to serve the needs of
speakers of both languages within a single volume. The
linguists concerned expressed their views at the
TABLE VI
Categorization of Bilingual Dictionary Users

orientation

source language

young
native non-native
learner
oral

adult
native non-native
translator
written

target language

young
native non-native

adult
native non-native
conference on Problems in Lexicography at Indiana State University in 1960. R. S. Harrell (1967: 51) maintained that "it is clearly impossible to pay equal attention to both X-speakers and Y-speakers in one and the same work," while S. Martin (1967: 154) concluded that "it is too much to expect such a dictionary to work both ways." The same opinion is implied in the contention of J. Rey-Debove that a dictionary for users of both languages must contain definitions of SL words as well as of TL words, a view that would create an impossibly cumbersome book.

1.1.1 Al-Kasimi considers that the limitations imposed on dictionaries by size and cost prevent the adequate explanatory information required if users of both SL and TL are to be satisfied. He uses as an example the fact that information which may be relevant in one direction (American football terminology for Arab users of an English-Arabic dictionary) will probably be irrelevant in the other, and quotes Harrell's example of the futility of having an entry for circumnavigate in addition to sail around or go around in a dictionary designed for English speakers. While it may be defensible to omit words like circumnavigate in a dictionary of limited proportions, there is another very good reason for including them: lexical items of this kind belong to a different register from those like sail around, the latter being more
informal. As we shall see in the discussion on register (below, Chapter 13), it is a matter of some importance that distinctions of this kind are made and the equivalent forms rendered at the appropriate level of formality. This is valid no matter which language user is involved.

1.1.2 The view that a dictionary cannot cater to two linguistic groups at the same time because of anisomorphic tendencies, such as Al-Kasimi's "football" example, is likewise incorrect because it is an established fact that there will be gaps in the conceptual domains of any given pair of languages and that such lacunae are filled by various expedients (definition, borrowing, calque and so on). Since the cultural reality belongs to one language group only, it is axiomatic that "translation" will occur in one direction only.

1.1.3 An important fact which is apparently overlooked by the disparagers of bidirectional lexicography is that the vast majority of those who consult bilingual dictionaries have at least some acquaintance with the second language. To compile dictionaries for the speakers of one of the languages only is to deny the existence of bilingualism and to assume that only those with a profound knowledge of both languages could possibly cope with the complexities of a bipartite dictionary. But what, precisely, are
these complexities? In order to answer this question let us now consider some actual dictionary articles:

Although [əl*-əl]-əl] conj bien que, quoique/ [+ subjunc]; même si [+ indic]: Although (I am) young, I have been around the world. Quoique (je suis) encore jeune, Tout jeune que je suis/. j'ai fait le tour du monde; Although he respected me, he . . . ; Bien qu'il me respectât, il . . .; [or avoiding imperf. subjunc.] En dépit de son respect pour moi, il . . .

Taking the information in order, the pronunciation is manifestly for the French speaker (although non-Canadian English speakers could also learn from it) and is transcribed in the *International Phonetic Alphabet* which is interlingual and adequately explained in the preface for those who are unfamiliar with it. The grammar category is represented by an abbreviation which can be checked in the abbreviation list if the user cannot immediately decipher it; almost all of the coded parts of speech are identical for both languages. The two equivalents which require a subjunctive construction and the one which takes the indicative are clearly so marked for the English user (the French user has no need for this information). The examples given not only illustrate subjunctive usage but also suggest methods for avoiding the construction. There is therefore nothing in the entry which precludes its satisfactory consultation by
speakers of either language, even though, in this particular case, the special explanations are designed to assist the English learner of French.

1.1.4 In the next entry we have an example of deceptive cognates, indicated for users of both languages by the neutral symbol \( [X] \):

\[ \text{apology} \ [æ'pələdiː] \ n \ 1. [X] \text{excuse(s)} \ f(pl) \ (de, for); \ [\gg \text{us. pl in Fr as in}] : I \text{demand an ~, J'exige des excuses.} \]
\[ 2. \ [\text{Philos}] \ [\text{defense}] \text{apologie f.} \]

For the English speaker a usage note comments on the tendency of French to use the plural and this is reinforced by an example. The translation in the second semantic division, which is a true cognate, is preceded by a label which amounts to a definition or further specification; although given in English, it is improbable that a French user would not comprehend its intention.

1.1.5 Finally, let us consider an example of differential semantic overlap:

\[ \text{coude} \ [kud] \ nm \ 1. \ [\text{Anat}] \text{elbow: un coup de ~, a dig with the elbow, a nudge; donner un coup de ~ à qqn, to nudge s.o.; ~ a ~, side by side.} \]
\[ 2. \ [\text{fig}] \ [\text{route}] (sharp) turn; [rivière] bend. \]
\[ 3. \ [\text{Techn}] \text{knee.} \]

English offers several different equivalents for a single
French lexical unit and these are specified by labels and examples which are fully comprehensible in either language.

The point to be stressed here is that the metalinguistic captions of a bilingual dictionary, like the metalanguage definitions in monolingual lexicography, are neutral and practically interlingual. At worst, the user has the option of verifying their meaning within the pages of the same dictionary.

1.2 Al-Kasimi (1977) also introduces the curious notion that, on the grounds that dictionaries should cater to one language group only, it will be necessary to provide phonetic information in the appropriate language. This means that a dictionary A + B for speakers of A should have pronunciations provided for the target words B—a practice which, to say the least, would be extraordinarily cumbersome since there might be as many as twenty different TL translations for a single SL word. Al-Kasimi counters this objection by an even more extraordinary proposal, that "it is advisable to cite one translational equivalent only, and so the objection is met" (p. 45). His thesis is based on a statement by S. Martin (1967) to the effect that single equivalents should be given whenever possible and alternatives should only be provided if accompanied by contexts explaining
why this particular form would be selected in preference to some other form. \[11\] It is unlikely, however, that a linguist of Martin's calibre intended this to be construed as meaning "one translation for one entry except in exceptional circumstances." What Martin objected to in his article was the listing of synonymous equivalents without due consideration for their need—a practice frequently found in concise or pocket editions of bilingual dictionaries. He wanted to make it clear that the user should be presented with the most frequently-encountered alternative first and less usual ones later.

1.2.1 The various analyses presented in Part I make it abundantly clear that one-to-one correspondence between the lexical items of a pair of languages, even genetically closely-related pairs, is almost non-existent and that suitably explained multiple equivalents must be provided in order to ensure the accuracy of the equivalence relation. Al-Kasimi's insistence on giving pronunciations in the target language is also based on the erroneous view that "very few dictionaries are bidirectional," (by which he means composed of two parts) \[12\] whereas the reverse is more often true. \[13\] Moreover, this practice would mean repeating the pronunciation many times over for those numerous lexical items which occur
as the equivalent of literally scores of SL entry words.

1.2.2 Overriding all these considerations, however, is the fact that pronunciations in a bilingual dictionary are of marginal value, rarely consulted by the majority of users and probably almost never in the case of the A → B example quoted above. In a two part dictionary there is no problem because phonetic transcriptions can be given in both parts; in a monodirectional dictionary designed for speakers of SL only, the absence of pronunciations for the TL equivalents is not a serious disadvantage because these may be found in the monolingual dictionary of that language which, if the enquirer is serious enough about learning the language, he will surely possess or at least have access to.

1.2.3 As a final word on the subject, it should be noted that language teachers have always stressed (and still do) the desirability of restricting the use of dictionaries translating towards the non-native language since this is the area of greatest difficulty and confusion. If such a dictionary must be consulted, it should be used in association with the opposite language version or with a monolingual dictionary of the target language so that equivalents can be carefully checked. This, alone, suggests that dictionaries should operate in both directions.
1.3 With regard to the actual selection of entries for a bilingual dictionary, the decision is necessarily based on the foregoing discussion as well as on questions of frequency and the scope of the dictionary itself. Although frequency is one of the most important parameters, it must generally be left to intuitive judgements since existing frequency lists are generally restricted to a few thousand of the commonest lexical items which would naturally find their way into all but the most diminutive of dictionaries.15

1.3.1 In making such judgements, lexicographers are undoubtedly influenced by the kinds of lexical items which find their way into monolingual dictionaries of similar proportions. They may also exercise control over certain repetitive elements, such as prefixes like anti-, sub-, auto-, pseudo- and so on which are either identical in both languages or else invariably rendered in the same way. Thus, for example, the extremely productive English prefix self- might be treated as follows:

- self- pref auto-; + de soi, + de soi-même; automatique: self-control, maîtrise f de soi; [>> usually English words in -self may be rendered by the formula N + de soi (--même); cf. also following entries].

Succeeding entries may then be restricted to those which deviate from the pattern supplied:
self-satisfied adj suffisant, content de soi.

self-supporting adj [venture] + qui couvre ses frais.

self-winding adj à remontage automatique.

1.3.2 Important assessments must also be made with respect to the quantity of technical language which can be included—a difficult decision since there is no clearly-defined demarcation between everyday language and more specialized terminology. It is a relative matter. A term like epenthesis, for example, which belongs to the daily vocabulary of all linguists, will be outside the acquaintance of the majority of non-linguists. The lexicographer's judgement must be based, in the final analysis, on whether the average user is likely to need the translation of such items or whether they should be left to the domain of specialist dictionaries.

1.3.3 Lexical items belonging to the various levels of the register scale, ranging from literary and archaic words to vulgarisms, are easier to identify but may engender other difficulties if they are not selected with discrimination. Some dictionary-makers have even faced litigation for propagating linguistic xenophobia in the form of words which are abusive to various ethnic groups, although it has emerged from such experiences that a suitable label, such as [Pejorative], is sufficient to
placate all but the most militantly anti-linguistic opinions. Similar considerations, therefore, also govern the inclusion or otherwise of terms at the extremes of the register scale. 17

1.3.4 With regard to proper names, unless there is a difference between the forms there is little point in filling the dictionary with identical pairs which provide no new information to the user. Thus, for French and English, it would be appropriate to eliminate Paris, New York, Pasteur and Churchill while retaining California/Californie, London/Londres, William/Guillaume and Mercury/Mercure. This cannot be taken as an invariable rule, however. In some circumstances identical pairs are included because of their use in syntagms which are differently rendered: the Stalin régime/le régime stalinien; or because they are the base element of derivative and fully lexicalized forms: Parkinson + parkinsonien; Trudeau + Trudeauamanie. A lexicographer may also opt to include identical pairs which relate to some aspect of reality which he is trying to highlight. For example, in CD2 it would be appropriate to include Canadian proper names such as Toronto, London (Ontario—which is not translated Londres) and Québec. Moreover, such lexical items may occur in contexts which require different equivalents: the Québec
government/le gouvernement québécois; or give rise to unusual syntactic features: in or to Ontario/à l'Ontario, en Ontario.

2.0 Word Order

Various methods of ordering dictionary entries provide lexicographers with a number of choices, each of which is advantageous in a particular situation. From a purely physical point of view, the sequence may be based on morphological elements (monemes), phonetic elements (phonemes) or graphic elements (graphemes), and each of these may be subjected to a number of further deformations such as the highlighting of certain vowel or consonant combinations and so on. Furthermore, it is possible to read the graphic image either left to right, as in conventional dictionaries, or right to left, as in inverse dictionaries. From a logical point of view, entries may be arranged according to semantic or etymological relationships.

2.1 For general dictionaries, where the disposition of the entries is largely a matter of user convenience and speed of consultation, the choice has traditionally been narrowed down to the arbitrary alphabetical arrangement on the basis of graphic structure or a quasi-morphemic organization, where words which are morphologically related are grouped together under a selected main head-
word; the latter is really just an extension of totally alphabetical order and is the system favoured by the majority of lexicographers until comparatively recently. The reason why many editors reject absolute alphabetical order is that entries which logically belong together on the basis of morphological structure and semantic contiguity are frequently separated by unrelated lexical items. For example, the English word ear occurs in a large number of compounds which logically should follow each other, but in absolute order the intervening entries dissociate the natural grouping, as Table VII illustrates.

2.2 A quasi-morphemic ordering, as in Table VIII, groups the words under a selected main entry which is the key for locating related lexical units. But this organization, which might well be described as being neither one thing nor the other, creates as many problems as it solves and often engenders a good deal of confusion. A. Rey (1965) gives the example of dictionaries using "topical" grouping within alphabetical order which create artificial links on purely etymological grounds; thus mollet (calf of leg) has nothing to do with mou (soft) in modern French, and yet it would be located under the latter in such dictionaries. Similarly, the Dictionnaire Bordas, which takes the "centre of interest" notion to extraordinary lengths, places s'attieidir under tiède and
<table>
<thead>
<tr>
<th>Word</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>ear 1</td>
<td>n [iːə]</td>
</tr>
<tr>
<td>ear 2</td>
<td>n [iːə]</td>
</tr>
<tr>
<td>ear 3</td>
<td>vi</td>
</tr>
<tr>
<td>earache</td>
<td>n [ˈɜːr.k]</td>
</tr>
<tr>
<td>*ear-catcher</td>
<td>n [ˈɜːr.ˈkætʃər]</td>
</tr>
<tr>
<td>ear-corn</td>
<td>n [ˈɜːr.ˈkɔrn]</td>
</tr>
<tr>
<td>ear-defender</td>
<td>n [ˈɜːr.ˈdefendər]</td>
</tr>
<tr>
<td>eardrop</td>
<td>n [ˈɜːr.ˈdɔp]</td>
</tr>
<tr>
<td>eardrum</td>
<td>n [ˈɜːr.ˈdɹəm]</td>
</tr>
<tr>
<td>eared 1</td>
<td>adj [ˈɜːrd]</td>
</tr>
<tr>
<td>eared 2</td>
<td>suff</td>
</tr>
<tr>
<td>eared grebe</td>
<td>n</td>
</tr>
<tr>
<td>eared seal</td>
<td>n</td>
</tr>
<tr>
<td>ear finger</td>
<td>n</td>
</tr>
<tr>
<td>ear flap</td>
<td>n [ˈɜːr.ˈflæp]</td>
</tr>
<tr>
<td>earful</td>
<td>n [ˈɜːrfəl]</td>
</tr>
<tr>
<td>earlap</td>
<td>n [ˈiːə.læp]</td>
</tr>
<tr>
<td>earldom</td>
<td>n [ˈɜːr.ˈlɒm]</td>
</tr>
<tr>
<td>earliness</td>
<td>n [ˈɜːr.ˈlɪns]</td>
</tr>
<tr>
<td>earlier</td>
<td>adj [ˈɜːr.ˈliər]</td>
</tr>
<tr>
<td>ear lobe</td>
<td>n</td>
</tr>
<tr>
<td>early</td>
<td>adj [ˈɜːri]</td>
</tr>
<tr>
<td>early bird</td>
<td>n</td>
</tr>
<tr>
<td>early-warning system</td>
<td>n [ˈɜːr.ˈwɔrənˌsɪstəm]</td>
</tr>
<tr>
<td>earmark 1</td>
<td>n [ˈiːr.mɑːrk]</td>
</tr>
<tr>
<td>earmark 2</td>
<td>vtr</td>
</tr>
<tr>
<td>earmarked</td>
<td>adj [ˈɜːrmɑːkt]</td>
</tr>
<tr>
<td>earmarking</td>
<td>n [ˈɜːrmɑːkɪŋ]</td>
</tr>
<tr>
<td>earmuffs</td>
<td>npl [ˈɜːrmɔfs]</td>
</tr>
<tr>
<td>earn</td>
<td>vtr [ərn]</td>
</tr>
<tr>
<td>earner</td>
<td>n [ˈɜːrnər]</td>
</tr>
<tr>
<td>earnest 1</td>
<td>adj [ˈɜːrnəst]</td>
</tr>
<tr>
<td>earnest 2</td>
<td>n</td>
</tr>
<tr>
<td>earnestly</td>
<td>adv [ˈɜːrnəli]</td>
</tr>
<tr>
<td>earnestness</td>
<td>n [ˈɜːrnəsn]</td>
</tr>
<tr>
<td>earnings</td>
<td>npl [ˈɜːrnənɪz]</td>
</tr>
<tr>
<td>earphone</td>
<td>n [ˈiːə.ˈfon]</td>
</tr>
<tr>
<td>earpiece</td>
<td>n [ˈɜːr.ˈpiːs]</td>
</tr>
<tr>
<td>earplug</td>
<td>n [ˈɜːr.ˈplɔɡ]</td>
</tr>
<tr>
<td>earring</td>
<td>n [ˈɜːr.ˈɪŋ]</td>
</tr>
<tr>
<td>earshot</td>
<td>n [ˈɜːr.ˈʃɔt]</td>
</tr>
<tr>
<td>earth 1</td>
<td>n [ɜːrθ]</td>
</tr>
<tr>
<td>earth 2</td>
<td>vtr</td>
</tr>
<tr>
<td>(followed by some 40 compounds with &quot;earth&quot;)</td>
<td></td>
</tr>
<tr>
<td>earthy</td>
<td>adj [ˈɜːrθi]</td>
</tr>
<tr>
<td>ear trumpet</td>
<td>n</td>
</tr>
<tr>
<td>earwax</td>
<td>n [ˈiː.ˈwæks]</td>
</tr>
<tr>
<td>earwig</td>
<td>n [ˈɜːr.ˈwɪɡ]</td>
</tr>
</tbody>
</table>
### TABLE VIII
AN EXAMPLE OF TOPICAL ORDERING

<table>
<thead>
<tr>
<th>Word</th>
<th>Pronunciation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>ear¹ n</td>
<td>[i:ə]</td>
<td>earl n [ˈeəl]</td>
</tr>
<tr>
<td>earache n</td>
<td>[ˈeək]</td>
<td>earldom n [ˈ-dem]</td>
</tr>
<tr>
<td>*ear-catcher n</td>
<td>[ˈ,kætʃə]</td>
<td>early adj [ˈeəli]</td>
</tr>
<tr>
<td>ear-defender n</td>
<td>[ˈ-də, fəndər]</td>
<td>earliness n [ˈ-nəs]</td>
</tr>
<tr>
<td>eardrop n</td>
<td>[-dɔp]</td>
<td>earlier adj [ˈ-ər]</td>
</tr>
<tr>
<td>eardrum n</td>
<td>[-dram]</td>
<td>early bird n</td>
</tr>
<tr>
<td>eared¹ adj</td>
<td>[-d]</td>
<td>early-warning system n [ˈ-wɔːnɪŋ]</td>
</tr>
<tr>
<td>-eared² suff</td>
<td></td>
<td></td>
</tr>
<tr>
<td>eared grebe n</td>
<td></td>
<td>earn vtr [ərn]</td>
</tr>
<tr>
<td>eared-seal n</td>
<td></td>
<td>earner n [ˈ-ər]</td>
</tr>
<tr>
<td>earfinger n</td>
<td>[-fɪŋə]</td>
<td>earnings npl [ˈ-ɪŋz]</td>
</tr>
<tr>
<td>earlap n</td>
<td>[-læp]</td>
<td>earnest¹ adj [ˈeənst]</td>
</tr>
<tr>
<td>earlobe n</td>
<td></td>
<td>earnest² n</td>
</tr>
<tr>
<td>earmark¹ n</td>
<td>[-mɑ:k]</td>
<td>earnestly adv [ˈ-lɪ]</td>
</tr>
<tr>
<td>earmark² vtr</td>
<td></td>
<td>earnestness n [ˈ-nəs]</td>
</tr>
<tr>
<td>earmarked adj</td>
<td>[-mɑ:kt]</td>
<td>earth¹ n [ˈeəθ]</td>
</tr>
<tr>
<td>earmarking n</td>
<td></td>
<td>earthy adj [ˈ-θi]</td>
</tr>
<tr>
<td>earmuffs npl</td>
<td>[-mʌfs]</td>
<td></td>
</tr>
</tbody>
</table>
théorème under théorie, arrangements which presuppose a reasonably mature native speaker. 21

2.3 Apart from the goal of a linguistically-oriented order instead of an arbitrary organization, the quasi-morphological grouping has the advantage that some space may be conserved by using shorthand devices to avoid duplicating repetitious information, such as pronunciation. This can be observed in Table VIII above, where the phonetic transcription of "ear" is given only once and replaced by the appropriate stress mark in all the compounds, whereas in Table VII it must be repeated after every interruption. It might also be possible to group all noun compounds together, and so on, in order to conserve space and eliminate recurring elements; but the loss in user convenience would probably outweigh the gains thus made.

2.4 It is difficult to comprehend why absolute alphabetical order should be considered inappropriate for the arrangement of dictionary entries simply because it is an arbitrary extralinguistic organization. The words must be located in such a way that they are readily accessible--this is surely the prime consideration. If it is necessary to stress linguistic "awareness" even in this feature of lexicography, then we should dispense with alphabetical order altogether and arrange the
entries according to principles based on semantics and morphology—in other words, all dictionaries should be structured in thesaurus fashion. Yet alphabetical order is a convention learned by everyone from a very early age. It is a relatively simple means of artificially organizing lexical information in such a way that it can be rapidly retrieved. It is consistent and objective and its serial order is easily comprehended by the speakers of all languages employing the Roman alphabet. Finally, it lends itself to machine-processing which, as we shall see in Chapter 16, is rapidly becoming an integral part of practical lexicography.

3.0 The Lexicographical "Word"

One of the most difficult issues in lexicography, as perhaps also in general linguistics, is the definition of "lexical unit," the element loosely referred to as "lexeme," "lexical unit," "word," "headword" or "entry word." It is beyond the scope of this chapter to attempt a complete analysis of the concept of "word" in linguistics. We shall confine our discussion, therefore, to those features which are of concern to the lexicographical conception of the "word."

3.1 For speakers of any Indo-European language, the concept of "word" is a general one which most people would define as a graphic or phonetic representation,
separated by blanks from others of the same kind. A few with linguistic training might describe it as the smallest meaningful unit which can stand alone. Still others may arrive at a pragmatic description which suits their particular purpose; for example, the delimitation of the graphic image for the purposes of automatic excerception in concordances. In a word, there is little agreement on this crucial element of the very substance of linguistic research.  

For lexicography, the problem is one of fixing the upper bound—that is, where lexical units cease to function as "words" and begin to constitute syntagms or locutions. The lower bound causes little difficulty because dictionaries have always treated the minimal graphic elements as entry words; even the selection of the form of the word, the representative of an entire paradigm as the lemma of an entry, is not a controversial subject because, for languages like French and English, a long lexicographical tradition has laid down acceptable conventions in this regard. In fact, it is probably the conventional isolation of words in dictionaries which has left us with preconceived notions about the nature of the term "word" which may, after all, be simply a convenient label which happens to suit lexicography. But, according to our earlier analysis, the "words" of a dictionary are abstract entities and some other frame of reference
may be required for "words" in context.

3.2 Until comparatively recently most lexicographers avoided conferring the status of "entry word" on anything which constituted more than a single lexical element; many went to extraordinary lengths in order to assign multiple elements to less conspicuous locations within the entry. As early as 1690, Furetière used a system of sub-entries in small capitals to account for multiple-word lexical units which he felt belonged to a single-word entry. For example, BOIS DE CHARRONNAGE appears under charroppnage, GARDE-CHASSE under chasse, and so on. Most bilingual dictionaries adopt some variation of this procedure, the most common approach being to place such lexical units as italicized examples under a main headword. Those which are considered to have become fully lexicalized, usually because of spelling as a hyphenated compound, may be treated as a single unit, and assigned to individual entry status.

3.3 However, there is now a very clearly discernible trend, at least in North American English lexicography, towards dealing with multiple-word elements on the basis of their functional properties instead of invariably relegating them to example positions simply on the basis of their formal characteristics. Various investigators have discussed these elements and described them in
different terms. For Zgusta (1971) they are "multi-word lexical units" which "carry lexical meaning as wholes (units); in this respect they function in the sentence in the same way as those lexical units which consist of one word only" (p. 144). A-J. Greimas refers to the "groupe de mots" to account for elements like pomme de terre, which semantically are single words but dimensionally are different from single graphic words and different again from compound verb forms, conjunctive locutions, and so on.25 G. Gorcy et al. describe the same elements under the heading "mot composé," which includes hyphenated words: chou-fleur, boute-en-train, franc-maçon, as well as those which do not have a hyphen even though they are perceived as compounds: pomme de terre, chemin de fer, eau de vie.26 J. Gates discusses "lexemes larger than the word," by which he means Zgusta's multi-word lexical units.27 A. Rey (1970) retains the expressions "unité de traitement" and "mot lexicographique" to depict the special form of the entries of a dictionary which are unlike the free "words" of contextual situations.28

3.4 J. Dubois and C. Dubois (1971) use the term "entrée" to denote all the lexical units of a dictionary;29 however, when they proceed to discuss their conception of lexicographical "words" they appear to confuse langue and discours:
According to the analysis in Part I, the dictionary appears to be a model of "lexical compétence," reflecting _langue_ and not _discours_; therefore the notion that lexicographical units are simply the "words" of the language is unacceptable. The authors themselves proceed to analyse the properties of dictionary entries which distinguish them from "words" to conclude that the former are "theoretical units," which is in direct conflict with their position regarding the practical, performance nature of lexicographical entries. 30

As far as compounds are concerned, J. and C. Dubois see only one solution—the traditional one, which dictates that a lexical unit is a continuous string of letters bounded by blanks at either end. For this reason they accept _coffre-fort_ and _gentilhomme_ as entries, but assign _pomme de terre_ to _pomme_ as an example, _chemin de fer_ to _chemin_, _compte courant_ to _compte_, and so on, even though these compounds do not differ conceptually from those like _timbre-poste_ where the hyphen replaces the redundant "de." 31 In both French
and English, the employment of hyphens is far from consistent. A word like *compte-rendu*, which is perceived as a single unit by all speakers, is still orthographically separated into its component parts by a redundant hyphen, and yet words like *neurochirurgie*, whose components are conceptually tangible in the meaning of the term, are totally compounded. English *dining room* and *motor vehicle* have entirely unified meaning and yet remain as separated lexical units.

3.5 The identification of lexicographical units on the basis of typographical conventions is obviously totally unsatisfactory and, in view of the chaotic state of the orthography of compound words, lexicographers must use a totally different set of criteria in order to identify the units of their discipline. At the turn of the century, A. Meillet's definition of the "word" reflected the general conception of the term: "Un mot est défini par l'association d'un sens donné à un ensemble donné de sons, susceptible d'un emploi grammatical donné."32 This interpretation of the word as a linguistic sign places no restrictions on its size: it may be a morpheme or an entire phrase and still fulfill the requirements of the frame of reference; the definition has thus fallen into disrepute as a linguistic description.

However, from a lexicographical point of view this
kind of definition (with a few modifications) may well be the solution to the identification of lexical units for a dictionary. In lexicography we are certainly concerned with sound groups which are related to a particular meaning and which operate according to certain grammatical rules related to part of speech category. We must therefore set the "lower bound" of the definition which, as suggested above, is the smallest typographical unit separated by blanks from other units. This effectively eliminates bound morphemes and affixes, although the latter may be treated on pedagogical grounds. However, it still leaves the question of the "upper bound" which must be the means of separating normal, sequential strings in context from set phrases and compounds.

3.6 Zgusta (1971: 144-152) proposes a set of guidelines for selecting the "multiword lexical units" for a dictionary. The most important procedure is the "commutation test" which enables a distinction to be drawn between freely commutable phrases: fat pig, dirty pig, and those like guinea pig which require both elements to retain full meaning. Analysis may also reveal that "semantic phenomena" indicate the difference between free combination and fixed compound; for example, the meaning of pomme de terre is not discernible from the constituent elements "apple" and "earth." Although the
same criterion may be applied to expressions like *drop a brick*, these are considered to be examples of idiomatic constructions of otherwise normal structures, usually in a figurative sense. If the proposed multiword lexical unit has single-word synonyms, this may be used as another selectional parameter: *motor vehicle/automobile*. Other methods of identification suggested by Zgusta include the existence of a single-element word in another language and syntactic operation, and he concludes his analysis by rejecting clichés of the type "divine beauty" as well as all proverbs, quotations and "similar fossilized petrified expressions." 

J. Gates (1977) emphasizes the importance of multiword lexical units in modern lexicography by pointing out that large proportions of the entries in a variety of dictionaries consist of such elements. He identifies different kinds of lexical units which will require different treatment—as full entries in the case of compounds (hyphenated or otherwise); as a subentry of a single-word entry in the case of "locutions figées" of the type *run to earth, run a temperature* (which appear under *run*); and as examples under a main headword if neither of the foregoing criteria apply.

3.7 Bilingual lexicography has generally followed the pattern of monolingual dictionaries with respect to the
entry status of lexical units; for this reason multiword entries are a rarity in bilingual dictionaries. In contrast to the figures obtained by Gates (cf. footnote 37), a sample using the same area of the dictionary (first 500 entries of R) revealed no multiword lexical units at all in CD1, CASS2 and QXF, and about 6% in HAP3 (almost exclusively phrasal verbs). Among the examples appearing under single-word main entries were numerous compounds which would be considered entries by many monolingual lexicographers: radiant heat, radio set, raise up, taxi rank, rate of exchange, rate-payer, reading-room, ready-made, rear, admiral, etc. The French section of these same dictionaries is even less productive with respect to multiple-word entries. Word-list sources for bilingual dictionaries can be assumed to be the monolingual version of each language and, since French dictionaries are very cautious about bestowing entry status on anything but single graphic units, there has been little observable change in this regard. But the influence of English lexicography must surely be brought to bear, eventually, through bilingual dictionaries wishing to keep pace with the comparatively new developments with respect to multiple-word units.

It is the physical location of these elements which is crucial in bilingual lexicography. The words quoted above were all found in most of the dictionaries sampled,
but their position as italicized examples makes them difficult to locate, especially when the internal structure of the article is not logically organized and is long and rambling. Relocated as boldface main entries in alphabetical order, or even as subentries in a quasi-alphabetical ordering, they stand out far more clearly and accordingly add to user convenience.

3.8 Zgusta (1971: 141) made a tentative proposal that the existence of a single-word unit as a translation equivalent for a multiword unit in another language might constitute a basis for determining the status of the latter. He seemed perturbed, however, that the same single-word element might have a multi-word synonym and that this would detract from the comparison—for example, English guinea pig = French cobaye = cochon d'Inde. Yet the existence of the latter should constitute further proof of the need to consider all such elements as identical in terms of lexicographical procedure. In fact, in bilingual lexicography it is just this kind of comparative analysis which should encourage editors to be more liberal in their determination of entry status.

Such units function in every way as though they were "simple words" or compounds. For example, morphological inflexions for plural follow regular rules: fat pigs, guinea pigs; les cobayes les cochons d'Inde; adjectives
occur in their normal position: les énormes cochons d'Inde, les cochons d'Inde intelligents; gender inflexions are governed by the nominal element: la bombe à neutrons, française. Certain discrepancies in morpho-logical variation are usually associated with historical factors: governors general (in addition to governor generals), knights bachelor, aides-de-camp; or because logic dictates the form taken, as in the invariable French compounds coupe-foin, porte-parole, porte-monnaie, serre-tête, compared with oiseau-mouche/oiseaux-mouches, basse-cour/basses-cours.

3.9 Another kind of multiword lexical unit which requires discussion is the so-called "phrasal verb" in English. These compounds consist of a true verbal form associated in various ways with other lexical elements: adverbs, as in put out, shove off, look over; prepositions as in wait on, care for, break off; and various other elements: fall in love, put up with, fly the coop. Lexicography has traditionally treated these verbal elements in much the same way as nominal and adjectival compounds, although with a more marked tendency to set them apart from simple examples under a single-word entry. Thus CD1 has bring about, bring around, bring back, bring down, bring forth, bring forward, bring in, bring off, bring on, bring out, bring to, bring together, bring up, as subentries under
bring (vtr). The existence of single-word synonyms and French equivalents seems to justify this procedure.

3.10 A problem with multi-word lexical units which is not discussed in any of the literature is the question of assigning grammatical categories. Most monolingual dictionaries pass over the question in complete silence, presumably leaving it to native intuition and the proximity of related single-word forms to resolve the unstated problem. One dictionary, Webster's New Collegiate Dictionary, 8th edition (WND), provides part-of-speech labels for non-hyphenated compounds on the same basis as single-word entries: rack up (vt), rabbet joint (n), inasmuch as (conj). This was not done in the earlier editions.

3.10.1 In bilingual dictionaries grammatical information of this sort is crucial for the non-native speaker because the equivalent in his own language may belong to a different category: let up (vi)/s'arrêter (vpron); date à laquelle/when, second-hand/d'occasion, sans importance/unimportant, the unmentionable/dont on ne parle pas. It is difficult to comprehend this difference on the part of lexicographers. On the one hand, they are happy to label fully lexicalized compounds like hummingbird, rainwater, man-of-war and shamefaced as noun or adjective; on the other hand, they leave parochial school, live wire, flock together, control post without commentary. The only
visible difference between these lexical units is the appearance of blank spaces between the elements of the latter; which seems to designate them as lower-level compounds by comparison with those which have earned either a hyphen or, the ultimate distinction for continuity, no division between the constituents.

3.10.2 The WND solution seems to be the only logical one, because there is no justification for considering the two groups as basically different merely on the basis of orthographical convention which, at best, is inconsistent with respect to hyphenation and compounding, and at worst is totally chaotic. It is often possible to use stress patterns in English compounds as the deciding factor. As R. Lees (1960) has demonstrated, nominal compounds have the same surface structure, even though their transformational derivation differs according to the constituents of the compound, and this surface structure almost invariably involves primary stress on the first element and secondary stress, if present, on the following element(s): bláckboàrd (as opposed to a very bláck boàrd), rainwater, éating apple, mánpower, cánnon fòdder, jáck-in-the-bóx. There is no good reason why only the non-separated lexical units of this group should be labelled as nouns.

3.10.3 Similarly, phrasal verbs must be clearly marked:
set forth is vtr in "to set forth a proposal" and vintr in "to set forth on a journey"; the French translation of the former is thus énoncer, proposer, avancer (une théorie) and of the latter se mettre en route, partir. In the case of phrasal verbs, stress patterns are different, usually with primary stress on each element: set forth, break down, and this needs to be distinguished from the stress pattern of nominal compounds: break-down, cast-off. In other words, such lexical units have their own functional as well as semantic features and the only way in which they can be adequately characterized in the dictionary is by making them separate entries where these features can be described in the usual way.

3.11 There remains the question of the multiword lexical units which do not satisfy the criteria set down for the simpler dual-element compounds; rack-and-pinion railway, bats in the belfry, day after day, bald-headed eagle, clin d'œil, voiture d'occasion, ceinture de sécurité, machine à laver la vaisselle. The establishment of the "upper-bound" for the multiword entries is not altogether clear and lexicographers may be obliged to exercise subjective judgements in order to draw the line which separates true lexical units from free combinations. In doing so, they may be influenced, to some extent, by the degree to which entry status for some elements will
add to user convenience. But the choice of grammar labels for these entries may no longer be clear-cut. It is unlikely, for example, that anyone will accept the description "adjective" for *bats in the belfry* or "noun" for *dispositif de défense au sud de la DEW-line* even though it is equivalent to *Mid-Canada line*. The CD2 project uses the label, *loc n* (= locution nominale/nominal locution), *loc adj*, *loc v*, *loc adv* for such elements. They have the merit of providing a succinct description of the probable syntactic function of the element while at the same time avoiding a totally prescriptive designation of a definite grammar category. The non-native user is thus provided with a guide to the way in which the entry is likely to operate in contexts.

4.0 Internal Article Structure

The internal organization of lexicographical articles is probably the most important feature of structural methodology. In essence, the arrangement of all the collected information about a lexical unit constitutes a kind of linguistic topography, which describes all of its potential manifestations in surface structure. It is, moreover, an independent "file" of information organized in such a way that consultation is rapid and successful—or rather this is what should be the case, but as we shall see, the internal structure
often impedes easy access to the material.

4.1 The first information encountered in any "file" is the head-word or lemma. As we saw in the preceding section, this is an abstracted representation for the entire paradigm of the word and the user is assumed to be able to generate the regularly-occurring paradigmatic realizations. Those which are irregular are usually indicated in the form of run-on information. However, for non-native speakers, in the context of a bilingual dictionary, it is essential to provide information on less common variations:

\[
\begin{align*}
\text{schema} & \quad \text{pl schemata} \quad \text{n} \\
\text{blanc, blanche} & \quad \text{adj} \\
\text{traducteur, -rice} & \quad \text{nm/f} \\
\text{bite, bit, bitten} & \quad \text{vtr}
\end{align*}
\]

The paradigmatic possibilities for English and French are quite limited and therefore the representational form used in the entry, together with run-on information, tends to be fairly inclusive. The most morphologically productive words are the verbs, although English is still quite restricted here and the principal parts supplied are normally sufficient information to generate the entire verb paradigm. The rich verbal system of French, on the other hand, means that some other expedient must
be called upon in order to provide users with all the
tense and mood realizations. This can be done in a
number of ways: most dictionaries establish a table of
the principal parts of the verbs in a separate section of
the dictionary where they can be verified; CD1 and CD2
use such a table but also include, with the verb entry, a
code number which indicates the pattern to which the
particular verb belongs, thus permitting the consultation
of a complete model for the verb paradigm; HNS provides
most of the principal parts of French verbs as run-on
material in the body of the work.

4.1.1 The canonical form selected for the lemma represents
the whole paradigm and is selected on the basis of
conventional usage—infinitive for verbs, nominative
singular for nouns, and so on. There seems to be no good
reason for altering this approach; the real difficulty,
as we observed in the previous chapter, is determining
what constitutes a lexical unit and not how it should be
represented in the dictionary. The well-established
lexicographical traditions of French and English have
eliminated the latter as a cause for debate.

4.1.2 The orthography employed in a dictionary, on the
other hand, will certainly give rise to debate. In
languages with official supervision of linguistic
development the problem is usually limited to discussions
on whether spelling reform aimed at simplification and uniformity should be introduced. French is such a language, although orthographical problems tend to be minimal. However, in a dictionary including Canadian French there will be occasion to make decisions on a number of variant spellings, many of which amount to phonetic variations of the type cataire/chataire (catnip), while others may be classified as phonetic deformations chèche (= sec, sèche), chauvérée (= chaudière), queuriosité (= curiosité). Judgements have to be made on which of the alternative forms is to be selected for the canonical version of the entry and on whether the remaining forms should be listed as alternatives. Both decisions, in the case of French, may have important repercussions with respect to the attitude of the various regulatory bodies, the Académie Française, the Régie de la Langue Française au Canada, and so on.

4.1.3 For English, orthography is a thorny problem, and perhaps nowhere more so than in Canada where usage fluctuates between British and American. For example, pairs like colour/color, centre/center are used indiscriminately, whereas programme prevails over program and tire over tyre. Regional variations are also discernible, so that even these examples cannot be taken
as generally applicable. The problem is that there is no attempt at central supervision of English comparable to the watchdog institutions for French. We are entirely at the mercy of journalists, publishers, bureaucrats and signwriters, with the result that most of us are incapable of stating, with any degree of certainty, what particular spellings should be employed in given circumstances. This is especially true of English compounds which are subject to three different orthographies—as a single word, as two words and hyphenated. If it could be established once and for all that all compounds conforming to the configuration described in the preceding section (that is, in which primary stress falls on the first syllable only and there is no occurrence of juncture between the elements of the compound) should be written as one word, we would be a long way along the road to simple consistency. The hyphen could then be retained for its real purpose—to mark the division between the elements of binuclear forms (spin-off, round-shouldered) and to separate vowels which might otherwise cause phonetic difficulties (anti-isolationism, co-opt, de-emphasis).

4.1.4 The general policy adopted by CD2 was to follow majority usage in Canada, as reflected in the nation-wide survey conducted by M. H. Scargill and H. J. Warkentyne.
But the problem of fluctuation remains as various media promulgate their own particular brand of orthography. Ironically, it may well be the news media which impose the stamp of finality on English orthography all over the world; for, with the proliferation in newspaper offices of computer terminals fed by the international wire service which is predominantly American, the chance for individual journalists to adjust spellings is rapidly disappearing. It is simpler and quicker merely to transfer the information directly to the composition output—indeed this is the motive behind this particular kind of automation which could have the unforeseen effect of standardizing the spelling of international English.

4.2 The second element of information in the internal structure of dictionary entries is pronunciation. Arguments can be produced suggesting that this material is not essential to bilingual lexicography and that it should be dispensed with. Certainly the inclusion of phonetic transcriptions in a dictionary A + B designed for speakers of A only would be a redundancy since native speakers are assumed to have sufficient control over this aspect of their language that they need only consult a monolingual dictionary of their own language in rare cases of uncertainty. The idea proposed by Al-Kasimi (1977) that the same dictionary should provide
pronunciations of the target language equivalents was discussed above (1.3) and rejected as impracticable. Even in a two-part bidirectional dictionary the phonetic information may be superfluous, either because few users of bilingual dictionaries ever consult them for pronunciation, or because well-motivated second-language learners are likely to possess the monolingual version in which pronunciations are given.

4.2.1 The opposing view, subscribed to by the majority of bilingual lexicographers, is that bilingual dictionaries are autonomous books which should contain everything likely to be of service to the second-language learner. Since most dictionary users appear to look upon pronunciation as one of the important pieces of information likely to be consulted, it follows that this should also be placed in a bilingual dictionary. Another reason for supplying phonetic spelling is that one of the languages may not as yet possess a monolingual dictionary (as would be the case for Amerindian languages, for example, although orthographic form usually does duty for pronunciation as well), or that the existing dictionary does not cover phonology to the satisfaction of the editor. It may, for example, neglect regional variations considered important in the bilingual dictionary. This is precisely the case with CD2 which
gives a succinct phonological description of Canadian English and Canadian French in the preface and then attempts to provide a phonemically-based transcription of Canadian English and International French with variants for Canadian French lexical items.

4.2.2 Before undertaking the phonetic description lexicographers, like all linguists concerned with phonology, must decide upon the system of notation to be employed. It is an unfortunate fact of lexicographical practice that there are almost as many forms of transcription as there are dictionaries, as the examples in Table IX illustrate for selected bilingual dictionaries of French and English. Although the International Phonetic Alphabet (I.P.A.) has existed for several decades and was designed to foster uniformity, it has not achieved its goal of universality. Also, dictionary-makers are not, after all, always linguists and their selection of other notations than the I.P.A. or its competitors is usually motivated by pedagogical considerations which dictate that phonetic spelling should be linked to known pronunciation patterns associated with certain spelling patterns—in other words, respelling in terms of orthography. But respelling techniques tend to favour the native speaker for the obvious reason that he is dealing with a system with which he is thoroughly
### TABLE IX
Phonetic Transcription in Selected Dictionaries

<table>
<thead>
<tr>
<th>ENTRY</th>
<th>CD1</th>
<th>CD2</th>
<th>HAP1</th>
<th>CASS</th>
<th>LAR1</th>
<th>NUG</th>
<th>BELL</th>
<th>GPFC</th>
<th>DCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>goal</td>
<td>gowl</td>
<td>goul</td>
<td>goul</td>
<td>goul</td>
<td>goul</td>
<td>gôl</td>
<td>gôl</td>
<td>gôl</td>
<td></td>
</tr>
<tr>
<td>frighten</td>
<td>'fræt'n</td>
<td>'fræt'n</td>
<td>'fræt'n</td>
<td>'fræt'en</td>
<td>'frij't'n</td>
<td>frighten</td>
<td></td>
<td>fright'en</td>
<td></td>
</tr>
<tr>
<td>measure</td>
<td>'mæsər</td>
<td>'mæsər</td>
<td>'mæsər</td>
<td>'mæsər</td>
<td>'mæsər</td>
<td>mezh'ør</td>
<td>mèjer</td>
<td>mezh'ør</td>
<td></td>
</tr>
<tr>
<td>propeller</td>
<td>pro'pelər</td>
<td>pro'pelər</td>
<td>pro'pelə</td>
<td>pro'pelə</td>
<td>propel'ér</td>
<td>propeler</td>
<td>propeler</td>
<td>µəpel'əw</td>
<td></td>
</tr>
<tr>
<td>shameful</td>
<td>'ʃeɪmfl</td>
<td>'ʃeɪmfl</td>
<td>'ʃeɪmfl</td>
<td>'ʃeɪmfl</td>
<td>'ʃeɪmfl</td>
<td>'ʃeɪmfl</td>
<td>chemef'1</td>
<td>sham'fəl</td>
<td></td>
</tr>
<tr>
<td>patinoire</td>
<td>patin'wa:r</td>
<td>patin'wa:r</td>
<td>patin'wa:r</td>
<td>patin'wa:r</td>
<td>patin'wa:r</td>
<td>patin'wa:r</td>
<td>patin'wa:r</td>
<td>patin'wa:r</td>
<td>patinwe:r</td>
</tr>
<tr>
<td>bien</td>
<td>bjê</td>
<td>bjê</td>
<td>bjê</td>
<td>bjê</td>
<td>bjê</td>
<td>bjê</td>
<td>bjê</td>
<td>byê</td>
<td></td>
</tr>
<tr>
<td>fievre</td>
<td>fje:vr</td>
<td>fje:vr</td>
<td>fje:vr</td>
<td>fje:vr</td>
<td>fje:vr</td>
<td>fje:vr</td>
<td>fje:vr</td>
<td>feeaivre</td>
<td>fyeV, fyëv</td>
</tr>
<tr>
<td>pourvoir</td>
<td>pur'wa:r</td>
<td>pur'wa:r</td>
<td>pur'wa:r</td>
<td>pur'wa:r</td>
<td>pur'wa:r</td>
<td>pur'wa:r</td>
<td>pur'wa:r</td>
<td>purvwè:r</td>
<td></td>
</tr>
<tr>
<td>serviable</td>
<td>serv'jabl</td>
<td>serv'jabl</td>
<td>serv'jabl</td>
<td>serv'jabl</td>
<td>serv'jabl</td>
<td>serv'jabl</td>
<td>serv'jabl</td>
<td>sərvyåb</td>
<td></td>
</tr>
</tbody>
</table>

Note: The transcriptions for GPFC and DCE are included to illustrate respectively Canadian French and Canadian English.
familiar; it has even been claimed in generative phonology that conventional orthography (of English) is a near perfect system for the lexical representation of words in the sense that it is often subject to consistent interpretation by rules and may operate as the underlying notation for a large number of variant pronunciations. But for non-native speakers, respelling systems tend to add to the complexities of learning a new orthographical system and offer no special help to them in mastering the orthoepy. In fact, it is usually considered that foreign users generally find a phonemic system based on IPA much easier to handle, if only because it lacks the complex array of diacritics often associated with respelling systems.

4.2.3 CD1 and CD2 adopt the I.P.A. and provide extensive information in the preface on the manner in which the system is employed for each language. (For obvious reasons, the explanation of English phonology is in French and of French phonology in English). In addition, a complete table of phonetic realizations for each language is placed in the inside cover to permit rapid consultation (See samples in Appendix V).

The selection of a notational system like the I.P.A. permits the pronunciation to be represented phonemically (broad transcription) rather than literally (narrow
transcription). Pronunciation, like the other features of a bilingual dictionary, is an abstraction which is probably never heard since the word in context undergoes changes in stress patterns which may affect individual sounds ([r] → [ə], [o] → [ə], and so on). Since it is just as impossible to provide all the distributional phonetic realizations of the words of a dictionary as it is to give all their semantic realizations, it is preferable to select a system which represents the general form. Al-Kasimi (1977) proposes that, in addition to the headword the illustrative examples should be transcribed.\(^48\) Not only would this approach create a confused and complex article, as we have seen, but also it is not in accord with his own previously stated claim that the entries should be transcribed phonemically with appropriate explanation in the preface of "the chief allophones and their complementary distribution."\(^49\)

4.2.4 Dictionary systems of pronunciation are never, however, purely phonemic since it is usually considered pedagogically essential to include features which may be automatic (such as vowel length in French or the diphthong alternation [ʌv]/[au] in Canadian English—the former occurring before voiceless consonants) or which are otherwise subphonemic (stress patterns in English which are non-distinctive). In CD1 this approach is explained thus:
We have accordingly chosen here to show a narrower type of transcription, including some subphonemic details which students are apt to overlook. Thus vowel length, which in French is automatic except for a few dozen words in [e], has been consistently shown by the phonetic sign (:). Students will thus be reminded of the sharp contrast between short syllables, e.g., type [tip], and long ones, e.g., tire [tiːr]. (p. xxiii).

In other words, the transcription is phonemic in orientation but contains elements of phonetic features which are deemed to serve the interests of the non-native language learner.

4.2.5 Another advantage of a phonemic approach to dictionary pronunciation is that the problem of social and stylistic levels can be treated more or less objectively. The preface information can provide a reasonably detailed account of the most frequently occurring realizations and leave the reader to judge which form is appropriate to his needs. In CD2 the preface describes the common variations between Canadian, British and American English and then gives the transcription for Canadian English, influenced largely by the work of Scargill and Warkentyne (1973). For French, the same dictionary provides pronunciation based on "International French" and gives a detailed description in the preface of the phonetic realizations of Canadian French. It is thus possible to employ a general pronunciation
guide which does not place the language learner in the invidious position of adopting either an excessively elevated and formal style of address or one which is too informal for normal communication.

However, in spite of the lengths to which lexicographers go in order to be accurate and impartial in their phonological treatment, it is largely a matter of satisfying their own views and at the same time avoiding the potential censure of other linguists since most second-language learners, after all, use living models or electronic aids in order to master pronunciation.

4.3 The information which follows the orthographical and phonological description of the entry word is the semantic presentation which may be associated with related features such as regional variation and register as well as commentaries on syntactic usage. From a purely structural point of view, it is important that these elements of information be located in logical positions within the framework of the dictionary article. The sequence in which the various features are encountered may be important to the correct interpretation of the material and it is certain that consistency in its organization is essential to ease of consultation as well as to comprehension. We will demonstrate in Chapter 16 that a carefully structured entry is also a necessary pre-
requisite to the successful processing of lexicographical material by computers which have a tendency to accentuate any disordered or inconsistently organized information.

4.3.1 The coherency of bilingual entries can be enhanced by following a few simple rules with regard to the location and physical presentation of the numerous labels which are employed. There are three main categories of label which apply directly to the headword: (a) regionalisms; (b) subject field; and (c) register. These labels should logically appear before the translation equivalents are encountered since, if they were placed after them, it would imply that the equivalents only were to be considered affected by the label in question. The order in which these respective labels should be encountered when more than one type occurs is basically arbitrary, although a logical progression seems to suggest that the dialect to which the lexical item belongs should first be identified, then the subject-matter (Math, Ling, Pol, etc.), followed by what amounts to a restriction of its application in terms of register (Fig, Fam, Poet, etc.) as in the following hypothetical example:


4.3.2 Since TL equivalents should conform in every respect to the SL entry, there should be no reason for using
these labels for the translations. However, the equivalents are often regional forms themselves and they should be appropriately designated, the label now following the translation instead of preceding it, as in the following example:

*chewing gum* n, *chewing m, chewing-gum m; gomme f à mâcher C.*

The inclusion of a regional form does not upset the equivalence relation but rather adds another dimension to it. Lexicographers also have a penchant for extending the equivalence relation by suggesting words which belong to a different register from the original. The label for this form, which cannot legitimately be omitted without leaving the user in ignorance about its application, may precede the translation in question as long as it follows other synonymous translations:

*remonter vi . . . . . : Cela remonte,* 
It revives you; [Fam] It's a pick-me-up.

These indicators, which we shall refer to as "higher order labels," are abbreviated forms and it is therefore possible to maintain them as closely as possible neutral with respect to French and English, thus removing, to some extent, the controversial question of which language they should be in. For languages which do not coincide as closely as English and French, it would seem
appropriate, in bidirectional dictionaries, to label in
the source language since native speakers of SL have the
greater need for support in translating into the second
language.

4.3.3 The second set of labels used in dictionaries are
those designed for meaning discrimination—"lower order
labels." Unlike higher order labels, these indicators
may apply to TL or SL independently, or to both at the
same time, and they may be subdivided into those which
provide information about the semantic variations between
a number of equivalents and those which suggest
alternative possibilities in the use of the translation
or the headword. The former should logically precede the
translation to which they apply and the latter should follow:

\[ \text{gush vi [blood, water] jaillir (à flots,}
\text{à gros bouillons).} \]

\[ \text{crop n 1. [Agric, wheat, etc.] récolte f,}
\text{moisson f. 2. [Anat] jabot m (d'un oiseau.}
\text{3. [Tech, whip] manche f (d'un fouet);}
\text{[cutting] coupe f, tonte f (de cheveux).} \]

The use of square brackets for meaning discriminators
associates them with the higher level labels, of which
they are really a sub-category, and distinguishes them
from the alternations or explanations which appear in
parenthesis after the translation.
4.3.4 A number of investigators have examined the question of meaning discrimination and the means of indicating it in bilingual dictionaries, but the universal solution has yet to be found and dictionaries continue to be somewhat inconsistent in their application of a general policy.

One of the main problems is the language in which meaning discrimination should appear. If dictionaries are prepared for one set of users only, this should present no problem—the labels should be in SL in a dictionary for speakers of SL and in TL for speakers of TL. This is the position taken by Al-Kasimi (1977) who believes that a bidirectional dictionary is a physical and practical impossibility. But most lexicographers, and a variety of linguists, do not share this view and, since bidirectional dictionaries do exist, we must address ourselves to the problem of the language of labelling.

4.3.5 J. Iannucci (1967) maintains that meaning discrimination is relevant only for translation from the source language and he therefore proposes that labels should appear in SL only. This is the procedure adopted by a number of major bilingual dictionaries and seems to be the only convenient way of attacking the problem in a bidirectional dictionary. E. Williams (1960) proposed that, if a dictionary is to serve the needs of both language users, it should address them in
each language and he established a formula which accounts for all possible cases of discrimination, including those where labels appear before and after the translation, one in each language. The latter situation, however, is probably a redundancy because explanatory material, being written in the neutral metalanguage of definition, ought to be comprehensible to non-native speakers who, as a final resort, can always find the meaning of a label in the same dictionary.

The simplest solution, therefore, and probably the most efficient, is to label meaning discriminators in the language of the entry word, except in those instances where an explanation is intended to elucidate a problem for the non-native user:

coup nm [⇒ Appears in many contexts where English has separate word, e.g., blow, kick, slap, etc. . . . . ] 1. [avec un instrument] blow: ~ de bâton, blow, whack; ~ de hache, chop; . . . . 2. [parties du corps]: ~ de pied, kick. ~ de poing, punch, cuff; . . . . . . . . . . 54

The guiding principle in this aspect of structural methodology is, as always, user convenience, and lexicographers should direct their attention to those whose need is greater in given circumstances.

4.4 A third and final order of labels consists of those which are in effect a commentary on usage designed to
further specify the entry word or one of its examples and those which are cross references. Usage labels may appear in various parts of the entry structure. As the example (in 4.3.5) illustrates for the entry "coup," the commentary may be placed at the very beginning of the entry, following the description of the head-word, where it is construed as being a commentary on the entire contents of the file. Alternatively it may appear before a semantic subdivision of the entry and apply to that portion only. Finally, it may be placed after a translation (thus constituting a remark referring to that particular element), or at the end of the total file (where it may function as a conclusion or add further elements of information):

n'est-ce pas loc. verb 1. [>> Fin de phrase qui se rend sur le modèle de is\textsuperscript{it} not? (si la phrase est affirmative)\textsuperscript{ou is it?} (si la phrase est négative), en reprenant le verbe auxiliaire de la phrase ou en employant le verbe do si le verbe principal n'a pas d'auxiliaire ou n'est pas lui-même un auxiliaire]: Il est là, n'est-ce pas?, He is in, isn't he?; ...

floor n 1. [room] plancher m, ... 2. [storey] étage m: to live on the 4th. habiter demeurer/au quatrième étage; [>> in Europe this would correspond to le troisième étage].

stony-broke adj [Fam] raide, à sec, fauché; décafé; [>> also STONE-BROKE].

It will be noticed that the commentaries are in the
language of production which, more often than not, represents the users most likely to need the information.

4.5 Cross-references are an important part of the lexicographer's opportunity to exploit the semantic relationships existing between words. As an element of internal structure, they are usually located immediately following a translation in order to direct the user to a word related to the entry:

\[ \text{janitor} \quad \text{nm} \quad \text{C} \quad \text{[Occup]} \quad \text{concierge m/f}; \]
\[ \text{[CD1]} \quad \text{gardien, -ne m/f, cf. CARETAKER.} \]

Sometimes their function is merely to direct the user to the word from which a given form is derived--especially principal parts of verbs:

\[ \text{mit} \quad \text{v} \quad \text{cf. METTRE}. \]
\[ \text{[CD1]} \quad \text{resous} \quad \text{v} \quad \text{cf. RESOUDRE}. \]

\[ \text{got, gotten} \quad \text{v} \quad \text{& pp} \quad \text{cf. GET; } \gg \text{gotten est moins fréquent que got}. \]

In other cases they are employed to direct the user to another word which is more fully treated or because the word in question is a regional form and is considered less acceptable than the normal form:

\[ \text{postillon} \quad \text{nm} \quad \text{C} \quad \text{[Poste] postman,} \]
\[ \text{[CD2]} \quad \text{letter-carrier; cf. FACTEUR}. \]
5.0 **Sense Order**

We have examined the visible features of bilingual dictionary entries and have concluded that certain aspects of the material dictate that the structural methodology should follow certain clearly-defined procedures aimed at implementing descriptive linguistic techniques and at the same time improving user convenience. The one remaining important part of internal structure is the ordering of senses which we shall now study in some detail.

5.1 The semantic description in monolingual dictionaries, which are the most usual source for the sense headings of bilingual dictionaries, organizes polysemous words in a number of ways. Historical dictionaries, for obvious reasons, classify meanings in diachronic order; technical dictionaries emphasize technical senses; general dictionaries may imitate historical order, or use frequency as the main criterion, or else combine historical as well as statistical parameters; other dictionaries may attempt a purely arbitrary arrangement based on, say, the alphabetical order of the subject headings for each sense, or a logical ordering which develops senses from the general to the particular or from the concrete to the abstract and figurative. According to J. and C. Dubois (1971), there is also a
syntactic factor to be considered, as, for example, when a verb enters into different constructions which engender different meanings. Distributional analysis, maintain Dubois and Dubois, teaches us that there are syntagmatic as well as paradigmatic relationships to be considered when seeking the sense patterns of lexical units.

5.2 One of the problems encountered by early lexicographers attempting to present a complete description of lexicographical phenomena was the successful organization of polysemous words. Emile Littré summed up the difficulties by stating in his preface: "Le classement des sens, quand ils sont nombreux et divers, est un travail épineux." Eventually he opted for a historical arrangement because he considered that the statistical parameter selected by other dictionaries (notably the earlier Dictionnaire de l'Académie), without being totally arbitrary, lacked the elements required of a systematic and national organization. Among other objections to the empirical approach, Littré pointed to the fact that the quantitative characteristics of the independent senses of a word are "transitory" and might be invalid even by the time the dictionary is published.

For L. Zgusta (1971), the arrangement of the senses depends largely on the purposes of the dictionary and on the users for whom it is destined. There is no
In reality there is no single principle by the exclusive application of which all polysemous entries of a dictionary could be arranged as to the sequence of the single senses... There is no single logical or epistemological system which would be both powerful and detailed enough to be used unequivocally and alone as the basis for the ordering of the senses and which would command a general authority and recognition (pp. 275-6).58

In describing monolingual lexicography, Zgusta presents the model used by most dictionary-makers which is based on the selection of the "dominant sense" of the polysemous word as the first one to be defined. If there is no dominant sense, the one which appears to have the broadest application and the least marked connotation is chosen, with other senses following in an order which seems best in the eyes of the lexicographer.59 Within the ordering, according to Zgusta's model, there may be sub-dominant senses which should not all be grouped at the beginning of the article but placed at strategic points throughout and followed by the more specialized meanings which are related to it.60

5.3 More problematical, perhaps, than the actual ordering of the senses is the means whereby the lexicographer arrives at the divisions. Zgusta warns of the dangers of over-categorization,61 a feature which dominates the sense partitions of Littré's work and which
modern lexicography endeavours to avoid. For example, Littré's entry lumière comprises 18 subdivisions, but these do not correspond in any way whatever to 18 separate senses. A. Rey (1965) studied Littré's entry homme, whose 26 sense categories illustrate the chaos which can result when a logical system is rejected in favour of subjective and mentalistic methods. Among the 10 first senses, meanings (1) and (10) relate to the general species ("espèce . . . individu de la race humaine"); (2) is individual man; (3) and (6) amount to contextual specializations of (1) and (2) ("homme . . . se dit de Jésus Christ; les hommes, la société"); (4) introduces a semantic enrichment of sense (3) ("homme par rapport à la bête"); (5) is a simple phrase belonging to (2) ("les enfants des hommes"); (7) illustrates examples in which "homme" loses its autonomy ("dépouiller le vieil homme"); (8) introduces a new sense ("celui qui possède les caractères moraux attribués à l'espèce humaine"); and (9) restates a distinction occurring only in various set phrases ("fonds humain"). Rey deplores the fact that the vital distinction for this entry, être humain/mâle/adulte is buried in a mass of secondary distinctions and does not appear until senses (11) and (12). 62

P. Guiraud (1967) demonstrated that the 64 senses given by Littré for the word tirer may be reduced to four basic categories: "traction," "parcours,"
"projection" and "reproduction (en série)." But because these divisions correspond to synchronic semantic criteria, they were not considered by Littré whose main guideline was etymology.

5.4 In modern lexicography, etymological relationships are considered as only marginally concerned in the organization of sense orderings, which tend to be based on features (referred to earlier) which are linguistically defensible and at the same time facilitate consultation. Probably the most advanced lexicographical venture in the world is the *Trésor de la Langue Française* (TLF), not merely because it has excerpted millions of words from a variety of contexts spanning two centuries of literary output in the French language, but because its editors are professional linguists who have applied the findings of modern linguistics to their practical endeavour. Thus, with regard to the organization of polysemous entries, the ordering is based on an empirical examination of semic structure supported by distributional, syntactic and derivational analysis—a procedure which eliminates the intuitive analysis employed by earlier generations of lexicographers and still widely practised today. Although frequency is a guiding factor in the arrangement of senses, there is an overall set of rules which applies a taxonomic logic to the final
layout. Thus, for example, literal meanings always precede figurative or extended senses; concrete precedes abstract; the general and statistically frequent precedes the particular, technical and rare. In cases where a number of special uses appear under the same sense subdivision, the order is alphabetical and not historical or etymological.

Rejecting mentalistic definitions entirely, the TLF establishes the structure of each entry on the basis of a careful distributional analysis, so that each definition is related to actual use. No definition is even attempted until the precise framework has been established to ensure that the meaning in question will be clear. Thus, for example, the introduction to the entry abandonner begins with the following description:

I. Emploi transitif
   A. [L'agent est une personne]
      1. [l'objet est une chose]
         (a) [Le lien antérieur avec l'objet était un lien de possession réelle].
         (Def.): --Renoncer à un pouvoir, . . .

5.5 Although bilingual dictionaries should follow similar motivations in the establishment of the internal ordering of senses, this is usually carried out in a manner which is far from systematic. As we observed in Part I, Chapter 3, there is almost no correspondence between various bilingual dictionaries on this question,
either because each has used a different monolingual dictionary as a model, or, and unfortunately this seems to be the more likely alternative, because they have no system whatever.

The dominant factor in the organization of polysemous entries should always be user convenience. Bilingual dictionaries are consulted, not for information about the diachronic status of the meanings of a source language entry or the factors which influenced their synchronic development, but in order to find appropriate equivalents. The easier these equivalents are to find, the more efficient the dictionary will be. Accordingly, no matter how strongly a lexicographer feels about the desirability of demonstrating etymological or diachronic relationships, in the final analysis he must bow to the needs of his readers. If the model selected is a monolingual dictionary of recent compilation, with a clearly articulated policy with regard to the treatment of polysemy, the result should be satisfactory.

5.6 On the other hand, there is no reason for rejecting logical sense ordering on the basis of some arbitrary convention simply because it is extra-linguistic. In fact, a good case can be made for purely arbitrary organization on the grounds that it is consistent and clearly comprehended by all users. Crowell's Spanish-English
Dictionary, for example, orders polysemous senses according to the alphabetical order of the subject labels, when present, and by alphabetical order of the first key word of examples when no labels are required, as the following example illustrates:

```
hacienda f large farm, country property, landed estate; property; treasury; Am ranch; Arg cattle, livestock;
hacienda de azucar, PR sugar mill;
hacienda publica, public treasury; public finances; public funds;
el Ministerio de Hacienda, the Treasury, the Exchequer.
```

Alphabetical progression is the most common convention adopted in order to locate the entries themselves; it has proved to be the simplest and most efficient disposition of the entries and the one which is universally recognized for languages using a Romanized alphabet. Therefore, its use for the internal sequence of senses seems to be a logical application.

As we shall see in Chapter 16, internal alphabetical ordering of polysemous entries lends itself to automatic processing of the material and permits a number of tasks to be manipulated by the computer—notably the merging and ordering of inverted entries.

5.7 We have already noted the inconsistency of bilingual dictionaries with respect to the adoption of an internal
organization for polysemous entries (Part I, Chapter 3) and we shall therefore turn our attention now to other aspects of entry arrangement. The actual ordering of senses and the philosophy which dictates their identification are only a part of the total methodology, because entries in a bilingual dictionary seldom consist of only a given set of appropriately segregated meanings. More often than not, the sense divisions are accompanied by examples which either reinforce the equivalence relation depicted in each division or else constitute derivational meanings. The most logical location for these examples is following the senses which they depend upon, and yet many dictionaries choose to describe the polysemous senses independently and then attach a series of locutions or sentences for which the relationship is lost, as we observe in the following example from OXF:

\textit{temperament} sm. 1. Constitution, temperament, temper, character, disposition, humour; 2. medium, middle course, compromise; (comm.) tally, tally system; \textit{avoir du}, to be ardent; to have a strong, amorous disposition; \textit{- bilieux}, choleric or bilious temperament; \textit{- lymphatique}, phlegmatic or lymphatic temperament; \textit{- sanguin}, sanguine temperament; \textit{par \textsuperscript{a}}, constitutionally, temperamentally; (comm.) \textit{\textsuperscript{a}}, by instalments.

The numerical division of the entry into two categories is the first feature to draw attention, but the commercial sense (comm.) is subsumed under 2, even though the
meanings are not even remotely similar. In other words, there should have been at least three sense divisions. The examples which follow relate to meaning (1) and to the commercial sense, although nothing is provided to assist the reader in seeing this relationship. Nothing is given, moreover, to suggest a contextual application of meaning (2). There is nothing intrinsically wrong with the order of the meanings; it is simply the vagueness of the divisions and the lack of continuity which may be criticized as contributing to confusion in the mind of a young learner of the second language. A more effective organization would be the following:

\textit{tempérament} s.m. 1. [Méd, Psych] constitution; temperament, temper; character; disposition, humour: \textit{avoir du} - , to be ardent; to have a strong amorous disposition; - bilieux, choleric, bilious/temperament; - lymphatique; phlegmatic, lymphatic/temperament; - sanguin, sanguine temperament; [Loc] par - , constitutionally, temperamentally. 2. [compromis] medium, compromise middle course. 3. [Com] tally, tally system: à - , by instalments.

The reorganized structure maintains the overall order (this could have been altered to a consistent system) but arranges the examples so that they may be quickly found in their logical situation—within the sense division to which they belong. Extra labels have been included in order to make the equivalents more comprehensible to a
French-speaking user, who has the greatest need to unravel the complexities of his polysemous word in translating towards English. The "either/or" segments have been replaced by a more suitable convention using oblique slashes to identify variant renderings. A more consistent use of punctuation enables semantic distinctions to be implied within sense levels; for example, the medical use of constitution as opposed to the psychological sense of temper, character and disposition.65

Not only do these changes result in a more readable and logically ordered entry which will enhance consultation but, as we shall see in Chapter 16, careful structural methodology is rewarded in another way if computational technology is employed—the efficient processing of the material in retrieval and sorting programmes and in the important inversion procedure. In the next three chapters we will examine the application of the formal aspects of methodology to specific problems in syntax and grammar, semantics and comparative stylistics.
Notes to Chapter 11

1. The illustrations will in general be based upon research conducted for CD2.


5. "But this procedure [the use of illustrative examples in both languages] cannot be followed by dictionaries aiming at a broader scope, otherwise the dictionary would become too cumbersome, bulky and expensive"; p. 23.

6. P. 22.


8. Al-Kasimi does not explain how the "football" problem is to be resolved for Arabic speakers.

9. Adapted from CD2 entries.

10. P. 45.


12. P. 45.

13. Cf. Zgusta (1971: 303): "An 'intentionally pure' dictionary like, for example, the English-Czech one quoted above (destined only for the comprehension of the English texts), is a rather rare phenomenon."

14. See the discussion below, 4.2.

15. For detailed discussion on the nature of the lexical frequency, see Guiraud (1954), Muller (1964 and 1968), Zipf (1935 and 1949), etc.

16. Cf. Zgusta (1971: 309): "(......) such a dictionary, especially if it is a smaller one, should be rather reserved in its inclusion of colloquialisms, slang expressions or even vulgarities and similar levels of language; and even bigger dictionaries of this type should be extremely cautious in this respect, lest the user be put into a ridiculous or painful position."
See the detailed discussion below, Chapter 11.

A. Rey (1965), p. 81.

Ibid., p. 81.

Ibid., p. 82.

Dictionnaire du Français Vivant, eds. M. Cohen et al., Paris: Bordas, 1972. For further discussion on the ordering problem, see Part I, Chapter 1 (3.6) and Part II, Chapter 16.

K. Togeby, "Qu'est-ce qu'un mot?", Travaux du Cercle Linguistique de Copenhague, Vol. 5 (1949), pp. 97-111: "... les linguistes ne sont même pas d'accord de ce que c'est qu'un mot, élément fondamental de l'objet de leur étude."

Cf. Part I, Chapters 4, 5, 6 and 7.


Pp. 61-2.

P. 62.

André Meillet, Revue de métaphysique et de morale, 1913, p. 11.

34 Ibid., p. 146.
37 Pp. 37-8; the statistics for a sample of the letter R revealed proportions ranging from 35% in Webster's Third to 13% in American Heritage Dictionary.
38 Ibid., pp. 47-49.
40 See the discussion in Part I, Chapter 7.
44 According to a survey conducted by C. L. Barnhart (1967: 161) pronunciation is considered third in importance after meaning and spelling and before synonymy, usage and etymology. This, of course, applied to monolingual dictionaries.
45 Al-Kasimi (1977) goes into this question in some detail, Chapter 3.1, pages 34-48.
46 A. C. Gimson (1973), p. 120.
47 "Experience shows that the foreign user (who has a different set of prejudices) finds such a notation infinitely simpler to grasp"; ibid, p. 120.
48 p. 44.
49 p. 38.
50 This is the position taken by J. Sledd in "Breaking, Umlaut and the Southern Drawl," Language, 42 (1966), pp. 18-41.
51 p. 73.
52 James E. Iannucci (1967), p. 204: "Since discriminations are required by the native-to-foreign use, they should be in the native language of the user, that is, in the source language."


54 Excerpted from CDI.


56 Pp. 94-98.

57 Preface to Dictionnaire de la langue française, p. 130.

58 This view seems to be shared by most lexicographers—of Householder (1967), p. 281.


60 P. 277.

61 P. 273: "Another important observation is that the lexicographer should not see polysemy where there is none. A not infrequent error is that differences caused by different applications of a lexical unit are taken for its different senses."

62 P. 68.

63 P. 171.

64 This description of the TLF organization of senses is based on P. Imbs, "Le Trésor de la Langue Française, Bulletin de la Société de Linguistique de Paris, 61 (1972), pp. 85-106.

65 Cf. Chapter 16, where the question of punctuation is related to machine processing.
References for Chapter 11


1.0 A common misconception among linguists and laymen alike is the notion that dictionaries are concerned with discrete lexical units exclusively in their function as the meaning-bearing elements of language and that syntax and morphology belong in a grammar. Some linguists have gone so far as to suggest that lexicography's concern with lexical information is such that the only words which should be found in a dictionary are those which have non-grammatical functions. M. Mathiot (1973) provides evidence which she claims supports the theory of the independence of lexicon and grammar and proposes that "words" which do not have lexical status, that is, grammar words (or function words) as opposed to content words, do not belong in a dictionary:

The second manifestation [the existence of grammar words] has been observed long ago: it is at the bottom of the distinction between function words and content words. The tendency in this case among both lexicologists and lexicographers is to disregard this distinction completely because it goes against the age-old conception of the lexicon as the inventory of all the words of the language.
In the approach that I follow, words with only grammatical status are not considered to be part of the lexicon. Consequently the main problem presented by the grammar-for lexicology is to exclude from the lexicon the elements that do not belong to it (p. 40).

Although the identification of "words with only grammatical status" constitutes a challenging problem for theoretical lexicology, the practical treatment of these items, according to Mathiot, is simply a matter of whether or not they should be included in a dictionary and if so, in what form.  

1.1 However, it is unlikely that a clear distinction can be made between grammar words and true lexical items in Mathiot's sense. She proposes to identify the two kinds of "words" representing "two separate dimensions of language" and then arranges these items into the appropriate categories. But her determination of various "non-content" words as function words only is unacceptable since it presupposes that the sole basis for according lexical status is that there should be designation of what Mathiot refers to as "entities and actions." Entries for prepositions like à in French and to in English consume enormous quantities of dictionary space, not simply because it is traditional procedure to deal fully with the grammatical operators, but also, and
significantly, because these kinds of item are such an important part of general language that they enter into countless constructions expressing temporal and spatial relationships. They are therefore just as much meaningful elements, in this sense, as are full content words. All discrete lexical units, considered in isolation, tend to suffer a reduction in semantic density which is recovered in context, as we have observed in Part I. The function words suffer most because their semantic content, in terms of tangible or perceptible identity, is lower and they do not reach their full potential until their combinatorial qualities are realized in context.

1.2 Conversely, content words may operate in a grammatical capacity, in the sense that what one language expresses by means of a function word may be expressed by another language employing a content word. This phenomenon is clearly seen in "chassé-croisé," the stylistic device identified by J-P. Vinay and J. Darbelnet (1958): blown away/emporté par le vent, the notion of "distance away from" expressly stated in the English adverb away is implied in the prefix of the French participle emporté. Conversely, the English participle blown, denoting the "means" of the action, is paralleled by an adverbial construction in French--par le vent.
H. A. Gleason (1973) summarizes the view proposed in this chapter, that grammar is an indispensible part of lexicography, in the following way:

There was a period in recent linguistic history when much was made of a distinction between "function words" and "content words." That was a useful step in establishing that language is a great deal more than vocabulary. But, as we came to know, it is an oversimplification. There are few items that fall cleanly into one category or the other. Most combine, in various proportions, the two manners of operation (p. 31).

The consequence of this state of affairs is that dictionaries must account for a good deal of what some linguists would claim to be the sole preserve of grammar, and that conversely, grammar must include much of what they would label purely lexical material. M. H. Scargill (1973) notes that Otto Jespersen recognized this overlapping of lexicon and grammar:

With regard to prepositions, it is quite right that dictionaries should account for the various uses of at, for, in, etc., just as they deal fully with the various meanings of the verbs put and set.

Grammar, according to Jespersen, must deal with the "general facts" concerning the manner in which function words operate and express various syntactic relations. M. A. K. Halliday et al. (1964) also emphasized the distinction which must be drawn between grammar and
lexicon, but recognized that this distinction is not always easily found:

If we can account for the prepositions of English by grammatical statements, we shall have said a great deal more about them, and about the way they work, than we do by entering them in a dictionary. But it is of course precisely these middle range choices that are the hardest to account for in a grammar.6

1.4 The theoretical position thus quite clearly stresses the independence of grammar and lexicon, but at the same time recognizes the grey area between the two which prevents their total separation. In practical terms, lexicography is concerned with words as discrete meaningful elements, but there is just as much concern for meaningful combinations of lexical units, as opposed to sentence structure per se which belongs to the study of syntax.7 The discussion which follows will thus centre on various grammatical problems which arise in bilingual lexicography (whether related to morphological or syntactic features), and which we subsume under the heading "differential grammar."

2.0 Grammatical Description

It is almost a truism that every part of a bilingual dictionary entry contains something of grammatical description. The selection of the canonical form of the headword itself involves decisions directly related to
grammar because, as we noted in Chapter 11 (2.3), the lemma is a representation of an entire paradigm whose elements are either unstated because they are regularly derived, or else indicated according to some convention—usually in an abbreviated form which requires native intuitions to generate the full paradigm. Phonology is treated in various ways according to different philosophies about the part it should play in lexicography. Part-of-speech labels are provided as an essential feature of total description and, in bilingual lexicography, as an indispensable adjunct of the equivalence relation. The translation equivalents themselves are accompanied by morphological and syntactic information (such as gender, irregular inflexions, government of verbs, and so on), and by anecdotal illustrations of the word in context which frequently serve to emphasize some feature of syntax which a non-native user cannot otherwise interpolate from the details provided about the headword. Sometimes lexicographers even introduce explicit usage notes relating to syntactic or morphological peculiarities of one of a pair of equivalents.

2.1 In Part I a good deal of attention was devoted to the question of accurate grammar category labels for bilingual dictionary entries, and evidence was provided from grammatical theory to support the position that
these labels, as well as being an important part of total lexical description, are also vital to the effectiveness of the equivalence relation. 9 Although native intuition can usually identify the subcategories of a class such as verbs (transitive, intransitive, auxiliary, pronominal), it is not sufficient in a bilingual dictionary to use the simple label verb used by many monolingual lexicographers. 10 What is required is a succinct description of these "higher-order categories" which is at the same time adequate to ensure that non-native users are informed of differential features. In other words, the grammatical description must indicate not only the main category to which a word belongs, but also the subclass.

2.1.1 Some statistical research has been done on the relative frequencies of the various part-of-speech categories in French and English and the results are of particular interest to bilingual lexicography. Gilbert Barth (1961) examined categories in French, English and Spanish and concluded that there is a kind of comparative stylistics for parts of speech, such that the predilection of one language for a turn of expression involving a given category necessitates various transformations in translation in order to maintain the statistical balance. 11 Barth's materials revealed the preference for nouns and verbs in French compared with adjectives,
adverbs and prepositions in English. For example, an English translation of a French text contains more adverbs than the original, because English has a far greater number of adverbs and uses them far more frequently. Where English will explicitly refer to spatial and temporal dimensions ("that book up there"), French leaves them unstated and implied by context ("ce livre-là").

The overall conclusion of Barth's investigation was that French translation carries out part-of-speech transpositions on about 7% of the words of an English original and that English transposes about 5% of a French text in the course of translation. These figures are significant enough to warrant the attention of bilingual lexicographers in naming part-of-speech categories as well as in verifying the translations.

2.1.2 While it would be simple and flattering to contemporary lexicography to suppose that the specification of grammar categories is a modern development, such is not the case. In 1747 Samuel Johnson stressed the importance of designating the subcategories of verbs:

The verbs are likewise to be distinguished according to their qualities, as actives from neuters; the neglect of which has already introduced some barbarities in our conversation, which, if not obviated by just animadversions, may in time creep into our writings.
Monolingual dictionaries frequently omit these, crucial subclasses, which are summarized in Table X, and assume the ability of native intuition to fill in the gaps. Bilingual lexicographers can make no such assumptions and should, in fact, be explicit to the point of redundancy in order to ensure non-native comprehension.  

2.2 D. Bolinger has proposed that the notion of transitivity is not of itself sufficient description for the operation of verbs and that a further specification—"saturated/unsaturated"—is required to indicate those verbs which require an extra complement for completion.

For example, sentence (1) below is complete, but sentence (2) is incomplete, even though all the elements of a sentence are present—subject, verb, complement—and in spite of the fact that both verbs are labelled as "transitive":

(1) I saddled the horse
(2) *I saddled the man.

Bolinger would describe saddled in (1) as "saturated" and saddled in (2) as unsaturated and requiring a phrase like with responsibilities. Similar conditions apply to verbs like set in, set one's watch, set the table, but set the book requires on the table, etc.; and French commander in commander un livre, commander le bataillon, but not
### TABLE X

**Subcategorization of Word Classes**

<table>
<thead>
<tr>
<th>Main Class</th>
<th>Subclasses</th>
</tr>
</thead>
<tbody>
<tr>
<td>verb</td>
<td>transitive/intransitive/pronominal; phrasal; saturated/unsaturated; tense; voice; mood</td>
</tr>
<tr>
<td>noun</td>
<td>singular/plural; collective/mass/count; proper; variable/invariable; masculine/feminine</td>
</tr>
<tr>
<td>adjective</td>
<td>attributive; positive/comparative/superlative; pronominal/postnominal; variable/invariable; possessive; numeral; masculine/feminine; interrogative</td>
</tr>
<tr>
<td>adverb</td>
<td>verbal modifier; adjectival modifier; phrasal modifier; verbal postposition</td>
</tr>
<tr>
<td>pronoun</td>
<td>personal/impersonal; pronominal; disjunctive; demonstrative; definite/indefinite; possessive; reflexive; masculine/feminine; relative; interrogative</td>
</tr>
<tr>
<td>preposition</td>
<td>nominal; verbal; motion/agent/location</td>
</tr>
<tr>
<td>article</td>
<td>definite/indefinite; partitive; masculine/feminine; variable/invariable</td>
</tr>
<tr>
<td>participle</td>
<td>present/past/future; attributive/adjectival; masculine/feminine; variable/invariable</td>
</tr>
</tbody>
</table>

**commander un garçon** which must be completed by à . . . de faire quelque chose.

Although the reflexive use of verbs is restricted in English to a few set phrases like *it beoves me, to betake oneself, to give oneself over to*, as well as contrived forms like *to kill oneself, to kick oneself, to cut oneself,* and reciprocals like *to love one another,* the pronominal verbal system of French is extraordinarily
productive. Generally speaking, a French reflexive verb is equated with an English intransitive: se lever (to get up), s'embrasser (to wrap up for warmth), s'entendre (to agree). But this is not always the case: se passer de quelquechose (to do without something), se rappeler quelquechose (to remember something). The reason for the imbalance is twofold. In most cases the French reflexive pronoun functions as a direct object of the verb (in its normal preverbal position) and results in a completed (or "saturated") phrase equivalent to an English intransitive construction. But sometimes, as in se rappeler, the reflexive pronoun is intransitive and requires a complement: Je me rappelle cet incident (I recall that episode), and therefore is equivalent to an English transitive. Verbs like s'embrasser, se passer, betake oneself are "unsaturated" and require a complement, which in French happens to be introduced by the preposition de. These are special constructions which bilingual dictionaries must account for by adopting an entry structure such as the following:

s'embrasser vpron to gain possession (of, de); to pick up, grab, seize, lay hold of/(sth, de qqch).

All the relevant syntactic details are included neatly without unnecessary disjunction of normal lexicographical form. For verbs like se rappeler, it would be
appropriate to add to the label "vpron" an indication that the reflexive in this case is indirect: vpron (indir), since this information not only clarifies the transitive translation, but also dictates the absence of participial concordance: elles se sont rappelé l'événement. As a last resort, and this is the most usual expedient chosen by lexicographers, some illustrative examples may be included. Sometimes, in fact, exemplification is essential since there may be no direct translation of the verb out of context; for example, s'entendre à quelquechose (to be very good at something). We have, then, the verbal structure schematized below which must be accounted for in bilingual lexicography.

```
VERB
    transitive
    intransitive
    pronominal
        un-
        saturated
        saturated
        saturated
        direct
        indirect
```

2.3 Substantives are likewise much more complex in functional realizations than the simplistic dictionary label noun reveals. While most lexicographers attach the indication of "plurality" or "singularity," which is often self-evident from the lemmatized form, few, if any, provide other specifications such as the fact that some nouns are uncountable or are used in a collective sense
while others are invariable in the plural, and so on. The feature "countability" can usually be implied in the absence of any other indication, leaving the way clear to specify only those less frequent features which are a potential source of differential confusion.

2.3.1 Some nouns, for example, are used in the plural only and are true plural forms in terms of their syntactic function: English earnings, remains, dregs, tongs, thews, trousers, West Indies; French annales, Antilles, ténèbres, appointements, oreillons, entrailles. Yet their true sense is collective and the equivalent form in the other language may be either singular: earnings (traitement, salaire), gallows (potence), bellows (soufflet), dregs, lees (lie), thews (force, puissance), trousers (pantalon), coveralls (salopette, tablier-blouse), dishes (vaisselle); appointements (salary), meubles (furniture), ténèbres (darkness); or plural: remains (restes), thews (tendons, muscles), annales (annals), oreillons (mumps), entrailles (entrails, guts), Antilles (West Indies), arrérages (arrears). The use of the label npl in such cases is thus far from accurate and should be supplemented by the indication "collective" in order to inform the non-native speaker that, in spite of being syntactically plural, these words are generally uncountable: *five dregs, *three mumps (although trois
meubles is possible because in the singular meuble means "a piece of furniture").

2.3.2 Similarly, large numbers of nouns used in the singular only must be accounted for in the methodology of bilingual lexicography. Examples are: English stamina (vigueur), wealth (richesses, biens), smelt (éperlan), clothing (vêtements), the press (la presse), broccoli (choux brocolis), consumption (consommation), flesh (chair), asparagus (asperges); fish (poisson, poissons); French: moyen (means), mélasse (molasses), dépouille (mortal remains).

Like their plural-only counterparts these words must have their "non-count" or "mass" features specified. Words which are used exclusively in the singular can be labelled n(sing only) or n(sing); if they have a plural form but are generally used in a collective sense in the singular, they must be labelled n(coll) or n(sing coll). Those, like fish, with specializes plural forms require special treatment:

fish n [pl sens coll: fish, espèce fishes] . . .

2.3.3 Another group of substantives forms a subgroup comprising words which have one meaning in the singular and another in the plural: English term (duration, semester, linguistic form, etc.), terms (relationship,
conditions); **stair** (step), **stairs** (stairway); **glass** (material), **glasses** (spectacles); **order** (command), **orders** (ordination); **measle** (swine lava), **measles** (disease); **good** (merit, kindness), **goods** (possessions); **tripé** (meat), **tripes** (entrails); French **largesse** (generosity), **largesses** (gifts); **bien** (good, merit), **biens** (possessions, wealth); **ciseau** (chisel), **ciseaux** (scissors, shears); **lunette** (telescope), **lunettes** (spectacles); **menotte** (child's hand), **menottes** (handcuffs). These words may be entered in their singular form with the plural noted as an additional sense division, the practice followed by most dictionaries, or else two separate entries may be established. The latter approach would be especially logical with nouns like **order/orders**, **glass/glasses**, **ciseau/ciseaux**, where the semantic differences are sufficiently separate to warrant individual treatment.

2.3.4 Finally, there are large numbers of French compounds which are invariable in the plural, that is, they have no formal morphological variation even though syntactically and semantically they function in singular or plural: **pied-à-terre** (lodgings), **porte-parole** (spokesman), **cache-sexe** (G-string), **ramasse-miettes** (brush and shovel), **pousse-café** (after-coffee liqueur). These must have the indication "invariable" attached to the part of speech
label, whereas a number of others: pèse-personne (scales), contre-manifestation (counter-demonstration), couvre-livre (book jacket), timbre-poste (stamp), surprise-partie (party) must have their variable and often irregular plural forms indicated as part of the lemma thus:

pèse-personne  pl  pèse-personnes
contre-manifestation  pl  contre-manifestations
couvre-livre  pl  couvre-livres
timbre-poste  pl  timbres-poste
surprise-partie  pl  surprises-parties.

Table XI illustrates the way in which some of the nouns discussed in this section are handled in various dictionaries.

2.4 Probably no other categorial division among the parts of speech is as badly abused and as poorly understood as the adjectives, which constitute another class of words whose subcategories are critically important in bilingual lexicography. The simplest variable element which requires attention is the inflexion or morphological variation for comparative and superlative degrees. If the preface has described the configuration of regular forms it remains for irregular forms to be shown in the lemma in the same manner as irregular plural nouns:
# TABLE XI
Nominal Category Labels in Selected Dictionaries

<table>
<thead>
<tr>
<th>ENTRY</th>
<th>FULL LABEL</th>
<th>CD1</th>
<th>CASS</th>
<th>HAP3</th>
<th>HNS</th>
<th>LAR2</th>
<th>GASC</th>
<th>BELL</th>
<th>PROB</th>
<th>DCE</th>
<th>BHD</th>
</tr>
</thead>
<tbody>
<tr>
<td>dregs</td>
<td>npl (sing rare)</td>
<td>npl</td>
<td>npl</td>
<td>s (us pl)</td>
<td>npl</td>
<td>spl</td>
<td>pl</td>
<td>npl n(us pl)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>trousers</td>
<td>npl (coll)</td>
<td>npl</td>
<td>npl</td>
<td>sing</td>
<td>npl</td>
<td>spl</td>
<td>pl</td>
<td>npl n (construed as pl)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gallows</td>
<td>n (coll)</td>
<td>npl</td>
<td>n</td>
<td>s (often with constr)</td>
<td>ninv</td>
<td>s</td>
<td>Ø</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>stamina</td>
<td>n (coll sing only)</td>
<td>n</td>
<td>n</td>
<td>s</td>
<td>n</td>
<td>s</td>
<td>Ø</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>flesh</td>
<td>n (coll sing only)</td>
<td>n</td>
<td>n</td>
<td>s</td>
<td>n</td>
<td>s</td>
<td>pas de pl</td>
<td>n</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>glass</td>
<td>n</td>
<td>n</td>
<td>n</td>
<td>s</td>
<td>n</td>
<td>s</td>
<td>Ø</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>glasses</td>
<td>npl (coll)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>spl</td>
<td>pl</td>
<td>npl</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>oreillons</td>
<td>nmpl (coll)</td>
<td>nmpl</td>
<td>nmpl</td>
<td></td>
<td>smpl</td>
<td>mpl</td>
<td>sm</td>
<td>pl</td>
<td>nmpl</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ténèbres</td>
<td>nf (pl only)</td>
<td>nfpl</td>
<td>nf (us. only in sing)</td>
<td>sfpl</td>
<td>fpl</td>
<td>sfpl</td>
<td>pl</td>
<td>nfpl</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ciseau</td>
<td>nm</td>
<td>nm</td>
<td>nm</td>
<td></td>
<td>sm</td>
<td>m</td>
<td>sm</td>
<td>Ø</td>
<td>nm</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ciseaux</td>
<td>nmpl</td>
<td>nmpl</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>smpl</td>
<td>Ø</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>cache-nez</td>
<td>nm (inv)</td>
<td>nm</td>
<td>nminv</td>
<td></td>
<td>sminv</td>
<td>m</td>
<td>sm</td>
<td>Ø</td>
<td>nminv</td>
<td></td>
<td></td>
</tr>
<tr>
<td>porte-parole</td>
<td>nm (inv)</td>
<td>nminv</td>
<td>nminv</td>
<td></td>
<td>sminv</td>
<td>minv</td>
<td>-</td>
<td>-s</td>
<td>nminv</td>
<td></td>
<td></td>
</tr>
<tr>
<td>surprise-partie</td>
<td>nf (pl in full)</td>
<td>-</td>
<td></td>
<td></td>
<td>sf (pl given)</td>
<td>f (pl given)</td>
<td>-</td>
<td>Ø</td>
<td>nf (pl given)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: Dash means word not given; Ø means word given without label; BELL incorrectly gives porte-paroles.
Those like initial, dual, several, extra, which are not subject to comparison should have this feature shown in the grammar label.

2.4.1 Much more contentious, as we observed in Part I, is the use of the label "adjective" for words which do not exhibit the normal properties of adjectives—modification by intensifiers (very, much), use with verbs like "seem," inflexion for comparative and superlative, and so on. In Chapter 7 of Part I it was concluded that a distinction must be drawn in lexicography between words which function in a partially adjectival manner—going (in "a going concern"), vernacular, bank, copper, cathedral—and those which are pure adjectives. The former, which frequently correspond to a pure adjective in French (vernaculaire, bancaire, cathédral), operate in attributive positions in such a way that the resulting combination of what amounts to a noun group (bank clerk, cathedral entrance, going concern) may frequently be considered as a compound noun and entered as such in the dictionary according to the criteria outlined above (Chapter 11, 3.6). Distributional analysis should enable a division to be established between attributive nouns
susceptible to combination with an indefinite number of second elements (copper, brick, sea, earth), which therefore appear in the dictionary as independent entries labelled "attributive," "adjectival," "nominal adjective," or some other suitable caption, and those which, together with their second element, constitute compounds in a closed series (bank clerk, bank rate, bank manager, bank teller, and so on) and merit treatment as discrete lexicographical units.

2.4.2 The vacillation in lexicographical methodology with respect to those words is clearly illustrated by the Supplement to the Oxford English Dictionary which has the following entry for "batik":

```
batik n & attrib (passing into adjective)
```

Like compounds consisting of separate words, attributive nouns are generally relegated to the grey area of uncertainty between the original part of speech and the one which has been partially embraced. The OED's commentary "passing into adjective" seems doomed to perpetuity, since it is improbable that the word will ever acquire all the features of a true adjective which remain elusive even for words of venerable lineage in the language.
2.4.3 From the point of view of bilingual lexicography, the specification of restrictions on the use of such words in adjectival contexts is indispensible; the non-native user will then be prevented from believing that the existence of a true adjective in his own language justifies using a carelessly labelled equivalent in the other language as if it, too, were a true adjective. When developing an adjective entry whose equivalents are not true adjectives, suitable labels should be appended to the translations and, where appropriate, exemplifying material included:

\[
\text{bancaire adj [Fin] concerning banking; [attrib] bank, banking: services bancaires, banking services.}
\]

2.4.4 A. A. Hill has suggested further specification of adjectives on the basis of whether they occur before the noun only (prenominal), following the noun only (postnominal) or in both positions. Examples of prenominal adjectives are: English late (= deceased), former, latter, every; French \textit{feu} (= deceased), \textit{ancien} (former), \textit{chaque} (every). Postnominals, by which Hill means following an existential verb, include: English content, averse, close (= mean). French adjectives, apart from words like \textit{feu}, are positioned according to stylistic criteria which cannot be globally accounted for in a dictionary. Those which regularly occur before the noun
(grand, petit, gros, vieux, jeune) may also be used postnominally (in Hill's sense): il est grand, la maison est vieille, and so on. The much larger group of adjectives which regularly follow the noun (une maison gigantesque) and which may always be used postnominally (la maison est gigantesque), may often occur prenominally for stylistic effect (une gigantesque maison).

2.4.5 A case can therefore be made for the specification "prenominal" or "postnominal," with the absence of this indication implying that the adjective operates in both positions. The use of a label of this kind would permit a neat distinction to be made between those French adjectives which change their meaning according to their position: prenominal--ancien (former), simple (mere, simple), pauvre (wretched), propre (own), même (same), brave (worthy), sale (rotten, mean), grand (great); postnominal--ancien (ancient), simple (simple, single); pauvre (poor, unproductive), propre (clean), même (very), brave (courageous), sale (dirty), grand (tall).

2.4.6 Adjectives may also be particularized according to the features "numeral" (cardinal or ordinal), "demonstrative" and "possessive," which appear in some dictionaries but are not consistently treated. In differential grammar the last two are particularly important because of the disparity between the systems of
French and English for concord (French possessive adjectives agree with the thing possessed rather than the possessor, as in English); the specification should therefore serve as a reminder of the special use of this subcategory of adjectives.

2.5 Adverbs, Bolinger informs us, also form subcategories made up of those which are verbal modifiers, those which are adjectival or adverbial modifiers and those which may modify an entire phrase or sentence. The first two are significant in lexicography since it is of some concern to a dictionary user whether a TL equivalent displays the same selectional restrictions or functional latitude as the SL word. Between English and French, adverbial equivalence with respect to these two subcategories is generally very close. Thus, English hardly and its French equivalents à peine, ne ... guère, presque operate with verbs and adverbs as in sentences (3) and (4), but not adjectives (sentence (5)):

(3) He is hardly brilliant/Il n'est guère brillant
(4) He hardly ever comes/Il ne vient presque jamais
(5) *A hardly affair/*Une presque affaire.

2.5.1 However, this functional equivalence in terms of modifying features does not extend to every aspect of their use. French à peine, for example is not restricted
to its position with relation to the verb:

(6) À peine suis-je dans la rue que . . .
(7) J'ai à peine commencé
(8) Je commence à peine
(9) C'est à peine si . . .

But English hardly always precedes the verb except in sentences like (3), using the verb "to be." It would therefore be appropriate to attach a specification to these adverbs indicating whether they are "preverbal" or "postverbal."

2.5.2 Those which modify other adverbs always either intensify or diminish the force of the other word: very fast, extremely quickly, scarcely better, trop tard, bien à propos; they therefore constitute a special class of adverbs for which the name "intensifier" is proposed to specify those which augment the impact of an original adverb, "diminisher" for those which reduce it.

2.5.3 In view of the disparity between languages with respect to the proportionate distribution of various parts of speech, we may expect some differential problems to arise in translation. English, as we observed above, possesses a significantly larger number of adverbs than French and the lacunae must be filled by means of adverbial phrases: jauntily/d'une manière enjouée,
contentedly/d'un air content, excitedly/d'un ton excité, satisfactorily/de façon à satisfaire, single-handed/sans aide. These phrasal equivalents will be more restricted than their single-word counterparts in terms of sentence position and this can only be explained in the preface where the general rules for adverb position can be outlined. The preface may also be used to indicate the fact that French tends to avoid using long and unwieldy adverbs by circumlocutions conforming to the configurations [d'une manière + ADJ], [de façon à + VERB], [d'un air + ADJ], [avec + NOUN], [sans + noun] and so on.

2.5.4 English adverbs which combine with verbs as a single sense unit and constitute "phrasal verbs" should be treated as lexicographical units and given entry status in the dictionary. From a distributional point of view, the adverbal component of these verbal compounds functions quite differently from the single adverbs found in context with verbs: come around (= regain consciousness) is not syntactically equivalent to come quickly or come soon. The latter examples contain an independent adverb modifying the verb in spatial or temporal terms; the former is a semantically indivisible unit in which the total sense is dependent on this combination and no other. Al-Kasimi's suggestion that expressions like leave out are simply examples of adverbs in idiomatic
combination with verbs is thus incorrect. He bases this notion on Bolinger's description of various restrictions in adverb use: leave out the details but not leave quickly the city, and comes to the erroneous conclusion that dictionaries should mark these adverbs in some way as being capable, or otherwise, of forming "two-word verbs." It would clearly be far simpler, to say nothing of being linguistically more sound, to recognize instead the existence of verbal units like take off, fall behind, gather in, fly away and so on, especially when it can be shown that they are often equivalent to single-word translations: décoller, traîner, recueillir, s'envoler. Even more significantly, confusion is often engendered by dictionaries which refuse to recognize these forms because it is not clear whether they are transitive or intransitive. Thus, take off may appear as an example under take (vtr), whereas quite obviously it is intransitive in the meaning "depart," "fly off" and transitive in the sense "remove."

2.6 Pronouns, as they are loosely designated in most dictionaries, actually form the important subclasses shown in Table XII. The extraordinary variety of pronominal forms is no doubt related to the statistical frequency of these words and their important place in overall syntactic structure. But, just as the most
TABLE XII

Pronoun Subclasses in Lexicography
frequently-occurring "content" words are the most difficult to describe adequately in a bilingual dictionary, so the pronoun subgroups constitute elements of differential grammar which must be clearly explained and appropriately labelled.

2.6.1 The most important subcategory comprises the personal pronouns which are further subdivided into conjunctive (weak, or unstressed) and disjunctive (strong or stressed). Conjunctive pronouns may function as the subject of a sentence or as the direct or indirect object. If the French personal pronoun *me* is to appear as a dictionary entry, which surely it must because, unlike a verb paradigm, the elements of pronoun structures have independent status and require special illustration, then it will be necessary to distinguish clearly between the four subclasses to which it belongs: conjunctive direct object (= me), conjunctive indirect object (= to me), reflexive direct object (= myself), reflexive indirect object (= to myself). Similar remarks would apply to the treatment of English *who*, *whom* (relative, subject, object, interrogative), French *lui* (personal indirect, disjunctive subject and object) and so on. Lexicographical treatment of all these words involves accurate description (grammar category) supported by adequate exemplification.
2.7 The article in French and in English, usually subdivided into definite, indefinite and partitive, creates little difficulty with respect to labelling conventions because of their near equivalence. In general, Figure 4 represents that equivalence at the formal level.

<table>
<thead>
<tr>
<th>ARTICLE</th>
<th>DEFINITE</th>
<th>INDEFINITE</th>
<th>PARTITIVE</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENGLISH</td>
<td>FRENCH</td>
<td>ENGLISH</td>
<td>FRENCH</td>
</tr>
<tr>
<td>sing</td>
<td>the</td>
<td>a, an</td>
<td>some, Ø</td>
</tr>
<tr>
<td>pl</td>
<td>the</td>
<td>some, Ø</td>
<td>des</td>
</tr>
</tbody>
</table>

Figure 4. Formal equivalence of French and English Articles (Ø = zero).

Functionally the equivalence is also close. English the may specify: the man who is coming, or generalize: the camel is a strange animal, in the same way as French: l'homme qui vient, le chameau est un animal bizarre. Yet these are important syntactic variations which may not be captured by simple lexicographical entry structure and which must be described in full. Each article for words of this kind actually constitutes a miniature grammatical description of the dominant contextual realizations. The Petit Robert (PROB), for example, has the following main subdivisions for the entry le, la, les (art. déf.):
I. Devant un nom:

1. (devant un nom générique) le chien est un mammifère carnivore, the dog is a carnivorous mammal; . . .

2. ("article de notoriété," devant un nom désignant un objet unique très connu, . . .) le soleil, the sun, avoir la fièvre, to have a fever; . . .

3. (devant les noms déterminés par un complément, ou une proposition) la lutte pour la vie, the struggle for life; . . .

4. (valeur démonstrative ou exclamative) Oh! le beau chien, Oh! what a nice dog!; . . .

5. (valeur distributive: devant un nom désignant une unité) cent francs la pièce, 100 francs each; (devant un nom de division du temps) le lundi, on Mondays; . . .

6. (après certaines prépositions et devant un nom de nombre) vers les huit heures, around 8 o'clock; . . .

7. (devant les noms propres de personne) la Callas, (Maria) Callas; . . .

8. (devant un nom propre pour en faire un nom commun) la Citroën de mon père, my father's Citroen; . . .

9. (devant un mot substantif) le manger et le boire, eating and drinking.

II. Devant un qualificatif:

1. (devant un adjectif qualificatif se rapportant à un nom déjà exprimé) les affaires politiques et les militaires, political and military matters; . . .
2. (répété devant des adjectifs) la grande et la petite industrie, major and minor industry; . . .

3. (à la + adjectif féminin, formant une loc. adv.) à la française, French-style.

III. Avec le superlatif: (V. plus, moins; mieux, pire, pis) la plus belle fille du monde . . ., the most beautiful girl in the world; . . .

IV. l'un . . . l'autre (V. autre); le même (V. même); l'on (V. on); tout le (V. tout); le mien (V. mien); la plupart (V. plupart).

2.7.1 A bilingual entry with French as the source language could use this general structure to establish a very full treatment of the definite article in French, including the most common contextual manifestations. We observe that, in spite of the formal equivalence just noted, there is a good deal of disparity, with English often using an indefinite article or a "zero" article in place of the French definite article. These are realizations which can only be described by means of illustrative material. The PROB article gives little attention to the important differential feature in which French uses the definite article for parts of the body: fermez les yeux, while English employs possessive adjectives: shut your eyes. This characteristic would have to be high-lighted with a separate subdivision in the bilingual entry.
2.7.2 Similar considerations apply to the other article forms in both languages where the same kinds of differential problems arise. It will be important to note under the English indefinite article, for example, that French drops the article with an unqualified noun indicating profession, nationality, religion, and so on: *il est médecin* (he is a doctor), *elle est Anglaise* (she is an Englishwoman); and also omits it before nouns in explanatory apposition: *il habite aux Trois-Rivières, ville que je connais bien* (he lives in Trois-Rivières, a town I know well).

3.0 Gender and Number Problems

Closely related to the question of the identification of grammatical categories and their treatment in bilingual lexicography is the important feature of grammatical gender and morphological concord. It is not simply a matter of their formal properties and the manner in which dictionary entries describe them; of equal concern is the broader spectrum of pedagogical motivations and psycholinguistic features. For example, the existence in one language of a total gender system which the other language does not possess constitutes a problem for lexicographical methodology which parallels the difficulties experienced by speakers of either language in attempting to adapt to the system of the second
language. French and English illustrate precisely this situation. The dual gender structure of French is grammatical, almost totally arbitrary and, except in a few cases, unrelated to sex; in contrast, English has a defective triple gender system which is almost exclusively natural; French gender structure extends over nouns, verbs, adjectives, pronouns, and articles and forms part of a complex system of sentence-wide concord; the English system manifests itself only in a small number of nouns and in the pronouns.

3.1 Similar remarks apply to number agreement which, although more important than gender in English, is much less extensive than in French. Number in English is more explicitly marked in oral language than in French where final -s, for example, is unpronounced except in cases where liaison comes into play; on the other hand, French plurals in -aux [o] constitute a severe difficulty for English speakers whose natural tendency is to develop -al into -als. In written language French number and gender extend over the entire phrase or sentence with a very high degree of redundancy. For example, in the sentence: les petites poulîches élégantes se sont sauvées the "femininity" of the basic nominal element is reduplicated in three other elements and its "plurality" occurs in five others. In the English equivalent: the dainty
little foals ran off gender is totally absent and number is marked only in the noun foals. Spoken language, of course, has yet another set of marks for gender and number and these, also, constitute an area of syntactic differentiation.

3.2 Before discussing differential gender and number problems and the methodology which can be employed to reduce interference in the lexicographical context, we will turn our attention briefly to some psycholinguistic considerations which have a bearing on the subject. Norbert Wiener's research in the 1940's, which eventually spawned the new science of cybernetics, applied the concepts of information theory to the problems of communication between living organisms. Information to be communicated, according to this theory, is viewed as a structure which may be transmitted in a number of ways—ink marks on paper, sound waves, electrical impulses, and so on; the success of the transmission of a message is measured in terms of the degree to which it reduces uncertainty (with regard to the content of the information) in the receiver of the message. Questions relating to the probabilistic nature of language (the statistical possibilities of various linguistic choices) and the stochastic process by which a message is decoded by its receiver all come into play—during the reception of the communication the reader or listener applies his know-
ledge of various features of language structure in order to reduce the number of choices left to be made for the full comprehension of the message.

3.2.1 The predictive characteristics of language decoding are significant to second language learning (and by analogy to lexicography) because they indicate the manner in which speakers and hearers organize the purely formal features of linguistic data. As we have seen, a high degree of redundancy, especially in written language, tends to reinforce the speaker's ability to retain certain linguistic features which are important elements in the transmission of information; they also enable the reader/listener to make accurate predictions about elements of the message which are yet to come and, freeing him from total subserviency to formal structural detail, allow him to give more attention to the content of the message.

3.2.2 Obviously the choices which must be made with respect to gender and number agreements in the production or reception of information will be part of the linguistic competence of the native speaker. To the extent that differential factors make these choices divergent in any given pair of languages, there will be varying degrees of interference and difficulty for the non-native speaker. For example, French adjectives with masculine inflexions
in **if** must be changed to **ive** with a feminine noun; this is a binary choice which is instinctive for a French speaker and which he will make long before reaching the element which is subject to the concord transformation. An English speaker, on the other hand, is dealing with an artificially imposed structure which does not come naturally to him and hence the process is laborious and the choices real. His language has no such system of adjective agreement because English adjectives are totally invariable. Faced with the choices which must obligatorily be made in sentences (10) to (12), the English speaker reacts in various ways related to the structure of his own language:

(10) J'ai remarqué un type<sup>offensif</sup> offensive
(11) J'ai rencontré plusieurs personnes<sup>offensifs</sup> offensifs
(12) Je déteste son affectation<sup>offensif</sup> offensive

In (10) he is working virtually in his own system: "type" has a gender which is motivated by sex and to which he can therefore relate his own equivalent word; the masculine agreement which is required tends to come quite easily because the adjective is usually learned in this form. Sentence (11), on the other hand, contains a noun which is introduced by a determiner without gender; personne has an unmotivated gender which is always
feminine even when applied to males. If the sense denotes a male, the English speaker's instinctive tendency will be to equate personne with the natural gender of English person and the choice of the correct inflexion is therefore more strained. Example (12) is even more complex because there are two conflicting influences: the possessive adjective appearing in the masculine form before a following vowel (irrespective of the word's gender) and the -tion desinence which invariably indicates feminine gender.

3.2.3 Psycholinguistic studies bearing on the problems of grammatical gender have concluded, with good reason, that native French speakers are very sensitive to word endings as a means to gaining control over gender. It is even conceivable that speakers of inflexional languages like French actually store lexical items in reverse order, thus placing the desinence in the prominent position. The enormous capacity of the neural areas of the human brain is such that it is easy to believe that the mechanisms for storage and retrieval of information could follow a number of such patterns. The fact that even a young child soon masters the number and gender systems of his language is sufficient evidence of the potential of our mental faculties in this regard.

3.3 Grammatical gender in French is activated by the
nominal component of vocabulary and eventually influences all word categories subject to gender inflexion (adjectives, pronouns, articles, participles). Although gender may be related to sex among animate beings (un homme, une femme, une jument, un lion), the vast majority of nouns have a totally arbitrary gender (le soleil, la lune, une maison, une victime, une souris, un rat). Synchronically, gender in French nouns can be considered a binary and arbitrarily organized structure which creates great difficulties for speakers of languages with a differently organized structure. The tripartite gender system of English, for example, is motivated by the oppositions male/female and animate/inanimate. All inanimates, apart from a few personified words (ship, Mother Earth, fatherland, mother country), are neuter and all animates are either masculine or feminine according to the sex of the designatum. A few words have a grammatical feminine desinence which is either a vestige of an earlier inflexional stage in the language or else the result of borrowing (goddess, heroine, sultana, artiste, danseuse, fiancée), and some use a prefix to specify sex (she-goat, man-child, bull-moose). But the majority are not discriminated by formal structure (queen, king, hen, rooster).

3.3.1 From a lexicographical point of view, English gender
can be dispensed with through a succinct description in the preface. The main area of difficulty concerns the possessive adjectives and pronouns, although it is simply a matter of convincing a French reader that, unlike their French equivalents, English possessives agree (by sex) with the possessor, not the thing possessed.

3.3.2 French gender must be very fully treated, both in the interests of accurate linguistic description and as an aid to non-French users. Bilingual French and English dictionaries are generally fairly explicit in their use of gender labels; they always appear after French substantive entries as an essential element of the part-of-speech label and frequently appear as well in following examples; they usually appear, also, as an element attached to the translations of English substantive entries and in succeeding examples. Pedagogical motives suggest continuous reinforcement of gender indications by whatever means are at the disposal of the lexicographer. For example, the translation of dotted line, appearing as an example under "line" or "dotted," should appear as ligne f pointillée which provides a dual reinforcement of the gender of ligne. When the adjective does not provide inflexional information the gender label is even more crucial: croix f double, rayons mpl cosmiques. Sometimes, specific gender problems can be
discussed in a usage note attached to an entry:

\[\text{ gens nm/fpl } [> \text{ gens is always masculine except when immediately preceded by an adjective with a distinct feminine form; thus: de vieilles gens, toutes ces bonnes gens; but: les braves gens sont hospitaliers, ces vieilles gens sont ennuyeux, des gens polis, tous les gens] people, ...}\]

If this kind of information is considered appropriate in a monolingual French dictionary,\(^{27}\) then it is even more crucial for it to be included in a bilingual entry.

3.3.3 French adjectives and determiners with distinct masculine and feminine forms should always be shown clearly in the lemma, even when regularly formed. Most dictionaries which follow this practice adopt a consistent system for indicating the morphological change which occurs. Unfortunately, the convention used in some dictionaries is simply to add the feminine form of the desinence after the canonical form is given as masculine, with no indication of the morpheme boundary where the desinence must be added: craché, ée; cracheur, euse; fanfaron, onne; jumeau, elle; actif, ive. The use of a dividing character of some kind, where needed, makes the position clear: crach\(|\text{eur, -euse}; fanfaron, -ne; actif, -ive; jum\(|\text{eau, -elle. It also enables a distinction to be made in certain cases, such as the ending -eur which may become -euse or -rice; the rule is that words which...}]}
end in -teur or -deur (acteur, spectateur, ambassadeur, donateur) become -trice, -drice, while those which end in -eur and in which a preceding -t- belongs to the base morpheme and not to the desinence (menteur, chanteur), change to -euse. Those which end in -eur, but are derived from Latin comparatives in -ior (extérieur < exterior; meilleur < mellior), change to -eure. This difference should be shown thus:

menteur, -euse nm/f
acteur, -trice nm/f
ambassadeur, -drice nm/f
meilleur, -e adj

3.3.4 From a pedagogical point of view the reinforcement which this methodology supplies makes it a sound practice. The question remains whether the same should be done for the French translations of English entries; in view of the space occupied by such material, it can probably only be answered by publishers and those who decide how much room can be sacrificed to motivations of this sort. Furthermore, it will obviously be a highly repetitive feature and may add a measure of extra complexity to material which is already very intricate and difficult to read. A compromise solution is to leave all translations in the masculine form if the feminine is regularly derived and to furnish only irregular feminines:
false adj faux (m), fausse (f).
dry adj sec (m), sèche (f).
public adj public (m), publique (f).
announcer n speaker (m), speakerine (f).

3.3.5 The front matter of a bilingual dictionary in which gender is perceived to be a differential grammatical problem would benefit by the addition of a statement indicating the manner in which gender can be rationalized for the non-native speaker. Since it is clear that French gender is to a large extent related to various affix patterns, these could be organized in the form of a chart (as in Table XIII) which would act as a guide to probable gender/affix relations; common exceptions should be mentioned in the preface and also labelled in the text. In the same preface it would be appropriate to discuss other gender-related problems, such as the homograph pairs whose meaning varies with the gender: le tour (trick, turn, tour), la tour (tower); le vase (vase), la vase (muck); le moule (mould, model), la moule (mussel). In the body of the work such pairs are exponentially separated with commentaries where necessary:

vase¹ nm container, vase.
vase² nf [X] slime, muck; [>> voir anglais VASE].

3.3.6 Further reinforcement of the need for gender
### TABLE XIII

**French Gender Desinences**

<table>
<thead>
<tr>
<th>DESINENCE</th>
<th>GENDER</th>
<th>EXAMPLES</th>
<th>EXCEPTIONS</th>
<th>TOTAL</th>
<th>%m</th>
</tr>
</thead>
<tbody>
<tr>
<td>-acle</td>
<td>m</td>
<td>miracle, spectacle</td>
<td>bâcle, bernacle</td>
<td>17</td>
<td>65%</td>
</tr>
<tr>
<td>-age</td>
<td>m</td>
<td>potage, sondage</td>
<td>rage, page, cage</td>
<td>933</td>
<td>99%</td>
</tr>
<tr>
<td>-ail</td>
<td>m</td>
<td>chandail, soupirail</td>
<td></td>
<td>20</td>
<td>100%</td>
</tr>
<tr>
<td>-eau</td>
<td>m</td>
<td>beau, chapeau</td>
<td>eau, peau</td>
<td>220</td>
<td>99%</td>
</tr>
<tr>
<td>-ège</td>
<td>m</td>
<td>privilège, siège</td>
<td></td>
<td>19</td>
<td>100%</td>
</tr>
<tr>
<td>-eil</td>
<td>m</td>
<td>pareil, sommeil</td>
<td></td>
<td>7</td>
<td>100%</td>
</tr>
<tr>
<td>-eur</td>
<td>m</td>
<td>jongleur, parleur</td>
<td>douleur, frayeur</td>
<td>1300</td>
<td>95%</td>
</tr>
<tr>
<td>-ier</td>
<td>m</td>
<td>pommier, belier</td>
<td></td>
<td>668</td>
<td>100%</td>
</tr>
<tr>
<td>-ien</td>
<td>m</td>
<td>canadien, macédonien</td>
<td></td>
<td>212</td>
<td>100%</td>
</tr>
<tr>
<td>-ment</td>
<td>m</td>
<td>froment, testament</td>
<td>jument (= 1281 advs)</td>
<td>1027</td>
<td>100%</td>
</tr>
<tr>
<td>-oir</td>
<td>m</td>
<td>miroir, savoir</td>
<td>cour, tour,</td>
<td>91</td>
<td>100%</td>
</tr>
<tr>
<td>-our</td>
<td>m</td>
<td>retour, four</td>
<td></td>
<td>37</td>
<td>95%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>DESINENCE</th>
<th>GENDER</th>
<th>EXAMPLES</th>
<th>EXCEPTIONS</th>
<th>%f</th>
</tr>
</thead>
<tbody>
<tr>
<td>-ade</td>
<td>f</td>
<td>promenade, salade</td>
<td>stade, grade</td>
<td>183</td>
</tr>
<tr>
<td>-aille</td>
<td>f</td>
<td>volaille, paille</td>
<td>coup-paille,</td>
<td>69</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>perce-muraille</td>
<td></td>
</tr>
<tr>
<td>-aison</td>
<td>f</td>
<td>saison, floraison</td>
<td></td>
<td>52</td>
</tr>
<tr>
<td>-ance</td>
<td>f</td>
<td>nonchalance, nuance</td>
<td></td>
<td>194</td>
</tr>
<tr>
<td>-ée</td>
<td>f</td>
<td>fée, corvée</td>
<td>musée, scarabée</td>
<td>210</td>
</tr>
<tr>
<td>-elle</td>
<td>f</td>
<td>tutelle, cervelle</td>
<td>labelle,</td>
<td>182</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>vermicelle</td>
<td></td>
</tr>
<tr>
<td>-ence</td>
<td>f</td>
<td>carence, immanence</td>
<td>silence</td>
<td>190</td>
</tr>
<tr>
<td>-esse</td>
<td>f</td>
<td>déesse, justesse</td>
<td></td>
<td>114</td>
</tr>
<tr>
<td>-ette</td>
<td>f</td>
<td>fillette, assiette</td>
<td>squelette, lettre</td>
<td>397</td>
</tr>
<tr>
<td>-ice</td>
<td>f</td>
<td>matrice, cicatrice</td>
<td>service, supplice</td>
<td>235</td>
</tr>
<tr>
<td>-ie</td>
<td>f</td>
<td>patrie, chacuterie</td>
<td>paraplui,</td>
<td>694</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>incendie</td>
<td></td>
</tr>
<tr>
<td>-ière</td>
<td>f</td>
<td>filière, théière</td>
<td>cimetière,</td>
<td>400</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>arrière</td>
<td></td>
</tr>
</tbody>
</table>

**Note:** The statistics are based on A. Juillard, *Dictionnaire inverse de la langue française*, The Hague: Mouton, 1970.
awareness can be achieved by attaching usage notes to the entries of words which exhibit some kind of anomaly, such as the desinences usually indicative of a particular gender but which are irregular, or the invariability of some nouns and adjectives:

brouhaha nm racket, hubub; [>> note gender].

squelette nm skeleton; [>> note gender --words in -ette are normally fem.].

thick-set adj breviligne (inv.).

connaissance nf [person] acquaintance; [>> always fem. even referring to males].

preux adj m [Hist] gallant; [>> this adj is always m].

3.3.7 French adjectives with senses which vary according to whether they are prenominal or postnominal also need to be considered under the general heading of gender and concord problems. Their lexicographical treatment should be as follows:

cher, chère adj [before noun] dear, beloved; [following noun] dear, expensive.

vrai, -e adj [before noun] untrue; [following noun] true: un vrai conte, a tall story; une histoire vraie, a true story.

3.3.8 Those which are invariable in certain combinations require appropriate commentary or exemplification:
bleu, -e pl -s adj blue; [⇒ invar. in compds]: des yeux bleu acier, steel-blue eyes.

demi, - e adj half; [⇒ invar. in compds]: une demi-heure, half an hour; une heure et demie, an hour and a half.

noir, - e adj black: une veste et un pantalon noirs, a black coat and trousers

French adjectives with a special masculine form before a following vowel must have this form indicated in the lemma or in a usage note:

vieux, vieille, pl vieux, vieilles adj [⇒ masc sing before singular noun beginning with vowel of h mute: vieill] old . . .

3.3.9 The irregular forms of nouns and adjectives should also appear as separate entries with a cross-reference to the appropriate lemma. This procedure both serves as a reinforcement and also ensures that the user is directed to the information he needs:

pire adj worse; cf. MAUVAIS.

messieurs nmpl gentlemen; cf. MONSIEUR.

The same methodology may be employed for the irregular principal parts of verbs:

mis, -e v & pp cf. METTRE.

résous, résout v cf. RÉSoudRE.

broke v cf. BREAK.
3.4 Number problems receive similar lexicographical treatment to those of gender. Irregular forms are indicated in the lemma (cf. 2.3 above):

nez (pl. inv.) nm nose.
trou pl -s nm hole.
beau-frère pl beaux-frères nm brother-in-law.
arc-en-ciel pl arcs-en-ciel nm rainbow.
marsupial, -e pl -aux, -ales adj marsupial.
goose pl geese n oie f; cf. GANDER.
mouse pl mice n souris f.

For irregularities in the target language the same conditions apply with respect to space considerations as those observed above for gender. Sometimes particular difficulties should take the form of a usage note followed by example:

most adj la plupart (de); [» plupart is always followed by a plural verb: la plupart des gens sont partis] . . .

nu, -e adj bare, nude; [» nu is invariable in compound use before noun, variable after it]: nu-tête, bare-headed; les pieds nus, bare-footed.

maint, -e adj many, manifold; [» note that maint may be used collectively]: maint autre métaphore, many other metaphors.
cent adj (one) hundred, a hundred: [>>
takes an -s in plural if not followed by
another numeral]: trois cents, three
hundred; deux cent trente, two hundred
(and) thirty.

4.0 Syntactic Problems

At the beginning of this chapter a discussion on the
dichotomy grammar/lexicon revealed that most linguists
recognize a clear theoretical division between the two
but admit the existence of an area of no-man's land in
which both operate. The conclusion was that lexicog-
graphy is basically concerned with the lexicon but cannot
ignore grammar entirely. Indeed, some linguists have
proposed that bilingual dictionaries should include a
description of the two languages for the benefit of non-
native users. But a complete contrastive grammar is
hardly the task of lexicography; if it were to be
exhaustive its size would rival the dimensions of the
body of the dictionary itself; if it is brief and
inadequate its purpose is questionable. Bilingual
dictionaries should probably compromise by summarizing in
the preface important syntactic features which can be
said to belong to the realm of differential lexical
description. The total grammar of each language is the
preserve of grammar texts or contrastive studies and is
the responsibility of language teachers.

4.1 A number of linguists have expressed doubt about the
possibility of drawing two different grammatical structures together. André Meillet, for example, suggested that there is no common ground between the grammatical systems of various languages. His formula is repeatedly illustrated in contemporary linguistics and is an accepted phenomenon of anthropological linguistics. The differential characteristics of comparative syntax vary in degree with the proximity or remoteness of genetic relationship. B. L. Whorf, for example, describes the Hopi verbal system with its nine voices (intransitive, transitive, reflexive, passive, semi-passive, resultative, extended passive, possessive and cessative), its nine aspects (punctual, durative, progressive, spatial, segmentative, inceptive, punctual-segmentative, projective and continuative), and emphasizes how the Hopi view of the real world is necessarily conditioned by the complexity of its verbal system which bears little resemblance to that of European languages.

4.1.1 But in spite of the disjunction which exists in grammatical structure, even between closely-related languages, translation is not totally inhibited. As Mounin (1963: 264) explains, an identical concept which may be treated syntactically in very different ways in different languages nevertheless remains the same concept;
the structure used to depict it is similar to the arbitrary nature of linguistic signs which still designate the same conceptual reality. He quotes as an example three different structures to depict the notion of "thunder peal"; Hopi uses the "word" rehpi with no reference to an agent of any kind; French has the expressions il tonne, ça tonne, which indicate an impersonal agent (it's thundering); and Italian uses a single "word" tuona in which the third person singular affix -a refers to an unspecified, but not necessarily impersonal, agent. 32

4.1.2 French and English are probably nowhere more divergent than in syntactic structure. Even the grammatical operators do not stand in anything like a one-to-one correspondence and the consequence of this is that bilingual dictionaries must give large amounts of descriptive material in the entries of function words. Although French de is generally equivalent to English of, from, there are many contextual and syntactic variations which must be accounted for. Each article for a grammar word in a bilingual dictionary constitutes, in effect, a succinct contrastive study, as the following sample (adapted from CD2 materials) illustrates:

\[
\text{de prép A. [rapports sémantiques]}
\]

\[
\text{I. [contenu] of: une tasse - thé, a cup of tea.}
\]
4.2 In section 3 of this chapter we examined the formal grammatical properties which must be accounted for in bilingual lexicography. We now turn our attention to purely syntactic structures which require special consideration.
4.2.1 Spatial concepts are analysed in different ways in French and English. The numbering system, for example, corresponds roughly at the level of discrete quantities, but is subject to considerable syntactic divergence. French cardinal numbers always précède ordinals:

(13) Les deux derniers siècles.
(14) Les trois premières voitures.

English word order preference, is the reverse:

(15) The last two centuries.
(16) The first three cars.

It may sometimes use the same order as French:

(17) The three first cars.

although this is really analogous to other sentences in which the second element is superlative:

(18) The three best cars.
(19) The best three cars.

This kind of divergence can be treated as exemplifying material under the appropriate headword.

4.2.2 Expressions of quantity requiring a partitive article in French can be explained by means of a following parenthetical statement when occurring in the
TL position:

enough adv assez (de).

and by examples when in the SL position:

tant adv so much, so many: de bruit, so much noise.

4.2.3 Physical measurements are expressed by avoir and être in French, by to be only in English. The pattern is indicated by examples under the headword:

avoir vtr [Mes] . . . de + adj, to be . . . + adj; la tour a 50 mètres de haut, the tower is 50 metres tall.

être vi [Mes] . . . adj + de . . . , to be . . . + adj; l'eau est profonde de 10 mètres, the water is 10 metres deep.

Age, time, dates, weather and dimensional measurements can all be treated in this manner.

4.2.4 One of the most troublesome differential syntactic problems is the use of prepositions, in particular their government by verbs in infinitive and predicate constructions. Although a rough guide can be proposed for those French verbs requiring à br de before a following infinitive or predicate, the construction is one which must be learned and which dictionaries should illustrate by parenthetical information or example:
renoncer vi to give up (sth, à qqch):
~ à une entreprise, to abandon a project.

refuser vi to refuse; ~ de faire qqch,
to refuse to do sth.

spend vtr [time] passer (le temps):
~ time doing sth, passer le temps à faire qqch.

French transitive verbs translated by English verbs which
require a preposition before a predicate, or which, more
accurately, constitute compound verbs (See discussion in
Chapter 11 above, as well as Part I Chapter 7) are
organized as follows when French is SL:

écouter vtr & intr to listen (to):
~ la radio, to listen to the radio.

This format captures the intransitive forms of both verbs
(when English to is not required) as well as the
transitive construction; the requirement of the English
preposition in this case is shown parenthetically in the
translation and reinforced by the following example.

English compound verbs as SL entries are dealt with in
accordance with earlier discussions on the status of
phrasal verbs:

pay for ['peɪ,faɪə] vtr payer.

fall over ['fɔːlˈouvər] vi [person]
tomber (à la renverse); [thing] se renverser.

Double preposition constructions may be indicated
parenthetically, by examples or by both:

defendre vtr to forbid (s.o. to do sth, à qqn de faire qqch): Je vous le défends, I forbid you to do it.

Variant preposition use is shown parenthetically:

defend vi défendre (de, on).

s'interesser vpron to be interested (in, à)

s'occuper vpron to take care of (sth, de qqch).

think vi [person] songer, penser/(à, of, about); [opinion] penser (à, of).

Specific idiomatic difficulties must be illustrated in examples:

jouer vi 1. [Mus] to play: - de [+ instrument], to play [+ instrument]; - du violon, to play the violin. 2. [Sport] to play: - à [+ sport], to play [+ sport]; - au golf, to play golf.

4.2.5 Most grammatical anomalies, as we have already observed, can be fairly neatly handled within the entry for the word concerned. Each function word will thus be an abbreviated contrastive grammar of that lexical unit, varying in complexity from those like French de (illustrated above, 4.1.2) to those like contre/against which are almost perfectly equivalent. Among content words there will be various constructional differences
which need to be described. It is insufficient, for example, to inform dictionary users that French manquer = English to miss because there is an important syntactic difference between them which is related to divergent conceptual structure. They must be handled as follows:

\[
\text{manquer vi to miss; to be/missing, missed/; [>> le verbe anglais est senti comme transitif]: Vous nous manquez beaucoup. We miss you greatly.}
\]

\[
\text{miss vtr [person, place, etc] manquer; [>> in this sense the French verb is intransitive and used passively]: I've missed you, Vous m'avez manqué.}
\]

4.3 The important differential syntactic features in French and English have been very thoroughly investigated by J-P. Vinay and J. Darbelnet (1958) under the heading of "contrastive stylistics." Although conceived as a translator's manual, their text is a veritable mine of information for bilingual lexicography; its material has been catalogued in such a way that the various kinds of equivalence categories may be applicable to many languages other than English and French.

4.3.1 Under the heading of "transposition" (transforms), for example, they characterize the very productive process whereby translation units are shifted from one morpholexemic class to another; from noun in English, say, to verb in French. Transforms of this kind are very
common and may be optional, where the choice is stylistically motivated:

As soon as he got up — Dès son lever (NP) — Dès qu'il fut levé (VP)

or else obligatory, where a particular constraint determines that one morpholexemic category be used to the exclusion of all others:

(20) No parking + Défense de stationner. 33

4.3.2 In a bilingual dictionary, compulsory transforms should invariably be given as examples; optional transforms also constitute very useful illustrative and reinforcement material and should be included if possible. There is, in effect, very little difference between the transformation which converts a lexical unit (or "word") like English offhand into French au pied levé and that which requires that a verb phrase be shifted to a noun phrase:

(21) Before he returns + avant son retour.

Yet, although all dictionaries describe transformations (or rather "modulations") of the type: gooseneck = col de cygne, lifebelt = ceinture de sauvetage, few go very far into the syntagmatic realm of optional transformations.
4.3.3 J. Casares (1950) gives explicit details on various kinds of syntagmatic "locuciones" which he would include as dictionary material, although he comes to no satisfactory conclusion on whether or not to incorporate occasional combinations such as buque de vapor, buque de vela, buque de hélice, and so on. However, he was discussing a monolingual dictionary for use in Spanish-speaking countries and these two strictures explain why he is inclined to omit such syntagms:

Se trata, evidentemente, de combinaciones occasionales, puesto que de vapor pued juntarse con el mismo sentido a "buque," "barco," "máquina," etc. En cambio la combinación caballo de vapor entra en el grupo de las de carácter estable, porque la acepción de unidad de medida equivalente a 75 kilómetros no reside en "caballo" mi en "vapor," sino en la combinación de ambas palabras (p. 101).

4.3.4 A. Rey (1965) also emphasizes that there is no clear boundary between syntagmatic and lexical units. Sometimes a syntagm may appear to be a lexicalized entity, such as "feu vert" in the expressions donner le feu vert, au feu vert, and so on, and yet be commutable between its elements feu vert, feu rouge, feu jaune; in such instances, argues Rey, the expression should appear as an example under "feu." The result of this uncertainty among lexicographers is that traditional methodology has caused the number of lexical units in a given language to
be significantly underestimated; for example, the *Petit Larousse* gives entry status to *arrière-grands-parents* but not to *pomme de terre*. These kinds of item constitute a problem for the nomenclature of the dictionary, whereas other more properly syntagmatic structures, if they are to be included, form part of the material of the entry itself.

4.3.5 Among syntagmatic units, those which are totally free and whose elements are fully commutable belong to the area of optional combinations and may not require treatment, even in a bilingual dictionary, as source language material. They are indefinitely large in the number of combinations possible and in the potential for expansion, as well as flexible in terms of their internal structure; they form, in effect an "open" class of syntagms. However, since we have postulated lexical competence as the theoretical model for bilingual lexicography, it may be considered appropriate for an example of syntagmatic structure to be provided whenever a number of combinations are possible. Although this takes the dictionary into the realm of performance, the concrete examples provided really do no more than emphasize the abstractive power of the material. For example, in the expression *il est friand de bonbons* (he loves candy) the element "bonbons" is fully commutable and serves only as
a factual example to exemplify the use of a syntactic structure which is not transparent for the non-native user. The fact that the example would appear under "friand" but not under "bonbon" confirms the status of the two main components; the use of parentheses and et cetera, in such cases, would help to isolate the commutable elements and emphasize the abstractive quality of the example: il est friand de (bonbons, etc.).

4.3.6 Syntagmatic units belonging to the closed class of expressions whose elements are in a unique and clearly determined relationship must always be included in a bilingual dictionary, at least as far as statistical frequency allows. There is no way for the non-native user to predict, for example, that libre à vous, means "you may," "it's your choice" (although the context may make the meaning clear); this information must therefore be provided. The kinds of expression susceptible to this treatment include "locutions figées" (set phrases), idioms, proverbs and syntagms with unusual structure. The choice of whether or not to include them will be based on frequency and probability of occurrence, on the equivalent behaviour of the corresponding language, the type of dictionary, the public for whom it is destined, and so on. Examples are: French fruits de mer (seafood), tant bien que mal (not too well), il en est ainsi (that's
the way it is); English I couldn't care less (m'en fiche), to have a good mind to . . . (avoir bien envie de . . . ); Canadian French passer un Québec à qqn (to deceive s.o.), se mettre au blanc (to progress), ne pas être vite sur ses patins (to be dull-witted).

4.4 A final class of syntactic features which may be of concern to bilingual lexicography involves the selectional and subcategorization rules which block certain combinations of lexical units. The selectional rules as characterized in a transformational generative grammar describe the selectional restrictions (based on feature categories such as animate/inanimate, concrete/abstract attached to each lexical unit) which determine what combinations are acceptable. Selectional restrictions thus belong to semantic structure and will be discussed in the following chapter. Subcategorization rules describe the limitations on certain syntactic combinations and as such concern us here.

4.4.1 Chomsky (1965: 106-7) illustrates strict subcategorization rules by schemata which, for example, classify verbs into subcategories of transitive, intransitive, pre-adjectival, pre-sentence, and so on. Violation of these strict subcategorization rules leads to anomalous sentences such as:
(22) *John compelled.
(23) *John became Bill to leave.

4.4.2 Bilingual dictionaries already indicate some of these rules by means of explicit labels: vtr, vintr, n(coll), adj(inv), and so on (as described in section 2 above). Others are expressed implicitly; in the absence of the indication "coll" (collective) or "sing" (singular), nouns are assumed to be countable; parenthetical information may suggest whether an animate or inanimate object is required:

\[
\begin{align*}
\text{lighten vtr} & \text{ éclairer, illuminer/(qqch).} \\
\text{promote vtr} & \text{ donner de l'avancement à (qqn).} \\
\text{obey vtr} & \text{ obéir à (qqn, un ordre).} \\
\text{déranger vtr} & \text{ [person] to trouble, disturb, bother/(s.o.); [device] to put out of order, cause to break down.}
\end{align*}
\]

In general, where the same restrictions apply to both languages in a bilingual dictionary, it is possible to leave them unstated and assumed as predictable. Thus, for example, transitive verbs, like eat/manger which may be passivized, and intransitive verbs like come/venir which may not, require no statement to this effect. Conversely, if one language allows such a rule while the other does not this phenomenon should find expression in the dictionary entry for the lexical item in question.
For example, French intransitive verbs which govern the dative may not be used personally in the passive, whereas English equivalent verbs may be passivized; the French construction in such cases is active and impersonal:

(24) We were offered a drink/On nous a offert une boisson

(25) He will not be asked any questions/On ne lui posera pas de questions

(26) They are forbidden to . . ./Il leur est défendu de . . .

Sometimes a reflexive verb does duty for an English passive:

(27) The buildings were mirrored in the water/Les bâtiments se reflétaient dans l'eau.

Alternatively, it may be possible to make the agent of a passive clause the subject of an active one:

(28) A motorcycle pursued by a police car . . ./Une motocyclette, que poursuivait une voiture de police . . .

4.4.3 Similarly, there are examples of transitive verbs which may not be used passively. For example, English to have (= be in possession of) is used only actively except in the case of a few idiomatic expressions: to be had up for (= accused), to be had (= fooled), a good time was had by all (= enjoyed). Certain set phrases are active or passive only: to keep tabs on (active), to be
4.4.4 Lexicographical treatment of such syntactic anomalies is generally inadequate; at best, non-native speakers are provided with an example of a verb like défendre used in a sentence which illustrates that an active form is the equivalent of an English passive construction. But this information may not of itself be sufficiently unambiguous because, although défendre may never be used passively, certain other French intransitive verbs which normally operate only actively may be used in the passive by using an inanimate subject: obéir, désobéir, pardonner, as in tes péchés te seront pardonnés.

4.4.5 For this reason, the usual lexicographical structure of the entry for such words is also insufficient of itself:

obéir vINTR to obey (s.o., an order, à qqn, à un ordre).
ask vTR demander (à qqn de faire qqch, s.o. to do sth).

Although the required construction is provided, there is no information which suggests whether or not a passive is possible. Even when an example illustrates the active/passive relation, this may not be sufficient to give a full characterization of the possibilities:
désobéir vintr to disobey (s.o., an order, à qqn, à un ordre): On a désobéi à l'officier. The officer was not obeyed.

It may be argued that knowledge of the syntactic features of certain classes of words is part of overall grammar and outside the domain of lexicography. Yet such information is as much an element of the totality of the equivalence relation as features like gender, part of speech and preposition government, all of which find their way into dictionary descriptions. The problem of how to depict such restrictions can be resolved in various ways; an abbreviated label can be attached to the grammar code: v. intr (act. only); a usage note may precede the translations or follow an example of the correct construction: [>> active only--use on, reflex. verb or transform].

4.4.6 Finally, the following are some examples of specific syntactic anomalies and the way in which they should be described in bilingual lexicography:

peut-être adv perhaps, maybe; possibly; [>>, as sentence initiator, invert following verb or use que]: est-il déjà arrivé, qu'il est déjà arrivé, perhaps he has already arrived.

jamais adv never, not . . . ever: Elle n'est jamais rentrée, She never came back; [>> when subject of sentence, no inversion as in English:] je ne l'aurai fait, never would I have done that.
can, could v. aux [› can possède un sens prés. et futur; could a une valeur de passé et de condit.] pouvoir, savoir: I can come tomorrow; Je pourrai venir demain; she could do it if she really wanted to, Elle pourrait le faire si elle le voulait vraiment; [› it is not always necessary to translate can, could]: I can't see them, Je ne les vois pas; We could hear them coming, Nous les entendions qui arrivaient.

run vintr courir; [› for run in, run out, run over, etc., use model entrer en courant]: He ran down the hill, Il descendit la colline en courant.

it pron pers (3) neut [following idea] ce; il; cela; [› in some constructions it is not translated]: I find it difficult to . . ., Je trouve difficile de . . .

ou pron rel where . . . from, where . . . to, etc; [› où often replaces sur lequel, dans lequel, etc.]: La rue où je demeure, The street on which I live.

feu, -e adj late, deceased; [› remarquer l'ordre des mots dans les deux langues]: feu la reine, the late queen; [› French word Order may imitate English; the adj becomes variable]: la feue impératrice, the late empress.

oneself pron [with verbs] se: to cut - , se couper; [with impersonal subjects, reciprocal verbs of motion] soi: to think of - , penser à soi; what one does - , ce qu'on fait soi-même.

y pron there; [› note use of y for à + infin]: Est-il décidé à le faire?—Oui il y est décidé, Has he decided to do it?—Yes he has.

faillir vintr to fail: - faire qqch, to almost do sth, J'ai failli tomber à terre, I nearly fell down.
beware vtr & intr [⇒ ce verbe n'est usité qu'à l'infinitif et à l'impératif] prendre garde (de, of): Beware of the dog!, Attention! Chien méchant; Beware the ides of March!, Méliez-vous des ides de mars!.

hope vtr espérer (faire qqch, to do sth); [⇒ in subordinaté with espérer use indicative, not subjunctive]: I hope (that) he will come, J'espère qu'il viendra.
Notes to Chapter 12

1 M. Mathiot (1973), pp. 43-44.
2 Ibid., pp. 39-44.
8 See above, Chapter 11, section 2.4.
9 Chapter 7, Part I.
10 H. A. Gleason (1967) emphasizes that "all pertinent grammatical identification" must be given for each item and that the dictionary should "index the grammatical statement" (p. 102).
12 According to W. F. Mackey (1967: 168) the proportion at the thousand-word level is 12% English to 3.6% French.
15 Speaking of Webster's Third New International Dictionary, Dwight Bolinger (1968) claims that "the dictionary fails to label the subcategories of mass noun and count noun. (This is essential information to any foreign learner of English) . . ."; p. 290.
16 Ibid., pp. 291-2.
17 Bolinger's examples, ibid., pp. 290-292.


21 Carlota S. Smith uses the term "intensifier" for a class of adjectives like mere, utter, real, which behave more like modifiers than adjectives. She does not distinguish between those which increase and those which decrease the impact of the sentence. "Determiners and Relative Clauses," in Modern Studies in English, eds. D. A. Reibel and S. A. Schane, New Jersey: Prentice-Hall, 1969, pp. 249-250.


24 It may be argued that "dainty" is marked as [+ feminine] since it generally, although not exclusively, relates to females. But this is a semantic feature related to sex and therefore it has no connection with grammatical gender.


27 The usage note and some of the examples in the preceding entry are adapted from PROB.


29 L. Urdang says that "a completely adequate grammar or dictionary of a living language may not be feasible; perhaps we are imposing on ourselves demands that by their nature never can be met," "Review of Problems in Lexicography," Language, 39 (1963), p. 592.

30 "Les systèmes grammaticaux de deux langues sont ... impénétrables l'un à l'autre," Linguistique historique et linguistique générale (tome 1), Paris: Champion, 1926, p. 82.


35 A. Rey (1965), pp. 72-3.

36 Ibid., p. 73.

37 Ibid., p. 84.

38 Ibid., p. 84.


40 Ibid., p. 148.

41 Chomsky's examples, ibid., p. 149.

42 Adapted from CD1 and CD2.
References for Chapter 3


CHAPTER 13

DIFFERENTIAL LEXICOGRAPHY: SEMANTICS

1.0 Semantic Preliminaries

The term differential lexicology was first proposed by J. Darbelnet (1970) to describe the disparate features of the lexical systems of pairs of languages; it would form a sub-component of the overall discipline of lexicology and its practical application would find expression in bilingual lexicography:

... il existe une discipline linguistique que l'on pourrait appeler la lexicologie différentielle et dont la lexicographie serait le champ d'application. Les insuffisances des dictionnaires tiennent en partie à ce que leur préparation a précédé la constitution de cette discipline. Ce n'est que lorsque nous aurons patiemment serré de plus près les écarts de sens, de valeur et d'emploi entre des mots qui en gros se correspondent que nous pourrons vraiment améliorer l'efficacité des dictionnaires bilingues (p. 102).

We have chosen, for this chapter and the one preceding it, to use the title "differential lexicography" and to subdivide it between grammar and semantics, a division which enables the term to apply to all facets of lexicographical deviation between pairs of languages.
Although the preceding chapter emphasized the importance of differential syntax, this aspect of bilingual lexicography is really only a prerequisite to the real goal which is the correct interpretation of the equivalence relation through an accurate analysis of semantic features. Whereas the goal of monolingual lexicography is the establishment of the "semic equation," bilingual lexicography seeks to solve what amounts to a pair of "simultaneous equations" whose unknown terms are the sometimes elusive common elements of meaning which enable a relationship to be established between word pairs. If the mathematical image may be pursued a little further, the difficulty of establishing the relationship is a measure of the complexity of the individual semic equations which may be linear (simple meaning) or polynomial (complex polysemous structure).

1.1 When Darbelnet suggests that the task of bilingual lexicography is not to resolve the question of definition but that of equivalence, he probably intends this to be construed as meaning that definitions belong to monolingual lexicography but that they must be clearly understood in the two languages of a bilingual dictionary before the equivalence relation can even be discussed. The degree of equivalence, as we observed in various sections of Part I, depends upon the semantic structure
of the individual words and the degree to which such structures in two different languages overlap. The problem for bilingual lexicography is to reconstitute this structure in the artificial environment of the dictionary and to take whatever steps are necessary in order to remove all sources of ambiguity and confusion. One of the methods suggested in Part I to achieve this end is to use a combination of componential analysis (to establish the semic structure of individual words) and lexical field analysis (to determine the relationships between semantically contiguous words). Although this approach may not be necessary for many words with monovalent semic structures which can be rapidly verified in monolingual dictionary definitions, there are significant numbers of lexical items which would benefit from such an analysis.

1.2 The conclusions drawn from the descriptive and theoretical introduction in Part I emphasize that the materials of a bilingual dictionary are necessarily an abstraction of reality and that the equivalents represent potential equality which may not be realized in real contexts. The nature of dictionaries is such that their information is removed from the global lexical network of associative relations (in which they naturally occur) and artificially restructured in arbitrarily organized new structures for the convenience of those who need to find
them in one place. This means that one of the essential
tasks of bilingual lexicography is not only the
constitution of the equivalence relation, which is a
theoretical entity, but also the establishment of the
practical means of its implementation. These practical
measures, which relate to the structure of dictionary
articles as discussed in Chapter 11 above, will be
analysed in further detail in the next section of this
chapter. The remainder of the chapter will be devoted to
the methods of establishing the equivalence relation, the
kinds of equivalence likely to be encountered, the
question of culture-bound words and lacunae, idioms and
figurative language, register problems and selectional
restrictions.

2.0 The Formal Properties of
the Equivalence Relation

Under the heading of "structural methodology" in
Chapter 11, we examined in some detail the numerous formal
features of dictionary entries all of which, in one way
or another, have a bearing on equivalence. The status of
the headword itself is a matter of no small importance
because what may constitute a lexicographical "word" in
one language may be a phrase of several "words" in the
other. In a dictionary we are dealing, not with "words"
as linguistically definable entities, but with lexical
units corresponding to clearly-defined meanings or groups
of meanings and therefore the same lexical divisions should be set up in both languages as far as is possible. If dishwasher and machine à laver la vaisselle both designate the same reality, then there is no logical reason for assigning entry status to the former but not to the latter. Lexicographers still cling to traditional notions with regard to multi-word lexical units, no matter what their semantic status, and the consequence is that machine à laver la vaisselle will be found in no dictionary as an entry but only as an example attached to machine or to vaisselle. The fact that one language forms compounds more readily than the other should not be an impediment to this aspect of equivalence. French actually possesses many compounds formed on the same basis as "dishwasher": essuie-glace, porte-parole, coupe-circuit, mange-tout, all of which find their way into dictionaries as independent entries, and it is not inconceivable that there will eventually be a compound "lave-vaisselle" to join lave-glace and lave-main. In the meantime, the five-word equivalent should have entry status on the same basis as pomme de terre and chemin de fer which are structurally and conceptually identical to it.

2.1 Equivalence also manifests itself, as we have seen, in the part-of-speech labels used in bilingual
dictionaries and great care must be taken to signal cases of non-equivalence such as the use of a collective in one language for a plural in the other, a transitive for an intransitive, and so on. These features form part of the total structure of each lexical unit and where exact equivalence does not occur lexicographers must be particularly vigilant about their description, always bearing in mind the special difficulties which may be encountered by non-native speakers of each language. Differential characteristics with respect to part of speech are shown in explicit labels as well as by means of succinct usage notes in the language of the user most likely to be affected:

suspend vtr surseoir (à des négociations, etc.).

bagage(s) nm (us. pl) luggage, baggage (coll. sing.).

un- pref neg [⇒] Used predicatively, adjs. in un- may often be translated by a Fr. verb in the neg. (this must be done in the case of Fr. translations formed by placing 'non' before an affirmative adj.): I am unafraid, + Je n'ai pas peur; an unfurnished room, une chambre non meublée; This room is unfurnished, + Cette chambre n'est pas meublée.

2.2 The most important formal feature of the equivalence relation is what we refer to as discriminatory labels, the information attached to bilingual dictionaries in order to specify certain characteristics or to indicate
restrictions in application of the equivalents provided. In Chapter 2 we discussed the physical location of these labels and provided arguments for their appearance in the language of the entry word as a general rule and in the target language in cases where the information is of particular interest to its speakers. The actual material contained in these labels is part of normal lexicographical methodology. They may be extensions of translations in the form of definitions, descriptors, usage notes or commentaries; they are written in the meta-language of lexicographical description which, in the case of definitions is as close to interlingual as possible. Defining labels appear in square brackets before the translation:

motte nf [de terre] clod; [butte] mound; [beurre, etc.] lump, pat.

Labels which specify or suggest contexts, as well as those which act as alternatives, appear in parentheses following the translation:

hunt vtr chasser (le canard, l'original, etc.); parcourir, battre (le terrain); poursuivre, traquer (une bête); rechercher (un indice, une indication).

2.2.1 The role of the defining labels is to disambiguate otherwise unspecified translations. In many smaller bilingual dictionaries, space considerations have been
considered more important than the need to inform the reader about the correct choice of the alternative equivalents which are often strung together as though they constituted a list of synonymous translations, any of which might be employed in any context:

motte f. mound; clod; lump; turf [gazon].

The only clue to possible differentiation in the first three translations is the use of semi-colons between them, a convention which usually signifies the existence of different semantic levels; the tagged final equivalent is a curious anomaly in an otherwise weakly described entry.

Words like French motte, as well as thousands of others in both languages, illustrate the well-known phenomenon of the non-equivalence of semantic extension in pairs of words; this divergence in the sphere of influence of pairs of lexical elements is due to their differing semic structure and may vary in scope from almost imperceptible discrepancy (the rarest kind) to virtually total disjunction. The word motte is a fairly generic term denoting "a piece of something," but no single English word, in spite of the existence of words like chunk, clump, bit and so on, covers exactly the same semantic area and this is why the alternatives must be given in the dictionary. But without some kind of
specifying label the equivalents are practically useless. The French speaker, composing in the non-native language, must be directed to the appropriate English word for each of the contexts in which his single word does service; except for the equivalent turf, the Larousse article clearly does not do this.

2.2.2 The material contained in these disambiguating labels amounts to a definition of the various senses of the headword in abbreviated form. The Petit Robert, for example, defines motte thus: "morceau de terre compacte, motte de gazon, motte de beurre," and the definition labels of the bilingual dictionary rearrange these in appropriate form, in French because they are intended to aid French users in this particular case of "bifurcation." J. Iannucci (1957) has proposed that the only way in which meaning discrimination can be dealt with effectively in bilingual dictionaries is to provide full definitions of the SL words and relate the TL equivalents to the numbered subdivisions of the monolingual description, as shown in Figure 5 below. Although this novel idea is somewhat unwieldy and consumes large quantities of valuable space, the basic premise that clearly defined equivalents are the only acceptable ones in a bilingual dictionary holds true and is advice that should be heeded by all bilingual lexicographers. The fully defined
semantic levels depicted in Iannucci's example could easily be reduced to abbreviated forms with no loss in comprehension for a native speaker who, as a last resort, can always consult his monolingual dictionary if the definition is still confusing to him.

**Source Dictionary**

**bolt n** 1. a movable bar which when slid into a socket fastens a door, gate, etc. 2. the part of a lock which is protruded from and drawn back into the case, as by the action of a key. 3. a strong metal pin, often with a head at one end and with a screw thread at the other to receive a nut. 4. a sudden swift motion or escape. 5. sudden desertion of a meeting, political party, program, etc. 6. a woven length of cloth. 7. a roll of wall paper. 8. a sudden dash, run, flight, etc. 9. a jet of water, molten glass, etc. 10. an arrow, esp. one for a crossbow. 1. a shaft of lightning; a thunderbolt.

**Target Dictionary**


Figure 5. A Method of Dealing with Meaning Discrimination.

Since this material is evidently designed for an English-German dictionary for speakers of English, it is hardly necessary for this amount of definition detail to be given. The first task of a bilingual lexicographer is to decompose the elements of meaning into major topical
subdivisions: [Mech, Techn, Pol, Weap, Meteor, Fig], and then into descriptors corresponding to refinements of these major topics: [bar, lock, pin; movement, desertion, etc.]. The results may be summarized as follows:


The essential details of the English definitions are retained, either by means of abbreviated descriptions or by being subsumed under the major subject headings; for example, the English label "Text" (textiles) leaves the English user in no doubt that *bolt* in this sense refers to a roll of material, whereas "Constr" (construction, building) on its own is insufficient to specify the particular meaning intended here and the label "wallpaper" is appended.

The formal role of the defining labels in bilingual lexicography, therefore, is to differentiate between the sense patterns which are characteristic of SL polysemous lexical units. The labels are of two kinds: higher-order labels, which particularize the major categories, and lower-order labels which specify secondary subdivisions of major categories.
2.2.3 The contextual labels, those which usually appear in parentheses following the translation, serve a quite different purpose. They may indicate an alternation between two or more elements in an equivalent lexical unit:

rebuild vtr (faire) rebâtir.
insist vintr tenir (absolument).
biffer vtr to strike (out).
neat adj (net et) ordonné.

They serve to illustrate syntactic alternations:

se mettre vpron to begin (to do sth, faire qqch).
research vtr enquêter (sur qqch, sth).
spite n rancune f (envers qqn, against s.o.).

and they suggest contexts in which the elements concerned may operate:

split n fente f (dans le bois)
spread n sandwich pâte f (d'anchois, etc.).
perdre vintr to/lose, be a loser/(at cards, etc.).
rembourrer vtr to upholster (a piece of furniture, etc.).

2.2.4 The first of these uses is simply a conventional shorthand method to conserve space by avoiding repetition.
The second relates to syntactic phenomena which were discussed fully in Chapter 12. It is the last feature which is of concern to us here because, like the defining labels, the parenthetical material operates in a specifying or disambiguating role and enables the dictionary user to discriminate between alternative translations on the basis of contextual information. This kind of information, as we have seen, emphasizes the abstractive power of translation equivalents; the items in parenthesis are not intended to limit the range of extension of the lexical unit concerned but to suggest real contexts in which it may operate.

2.2.4 Apart from definition and context indicators, bilingual dictionaries use the major labels which provide information on dialect, subject matter and register. As we shall see in Chapter 15, regional forms do not have to be matched with regional forms and therefore these labels are only marginally concerned in the equivalence relation. If the SL headword is a regionalism this must, of course, be clearly identified as such, but the translation equivalents may be in standard form or in regional form according to individual circumstances; sometimes, for example, it is appropriate to provide a dialect equivalent when the lexical item concerned designates a reality common to the geographical area or cultural
milieu shared by the two dialects:

komatik n (C) [Amerind, Transp: dog
sled] cométique m (C).

In other cases the equivalent of a regionalism may be a
standard form only or it may be represented by a standard
and a dialect form:

raquetteur nm (C) [Transp] showshoer;
raquetteur (C).

2.2.5 Subject field labels, as we saw above (2.2.2),
operate in combination with defining labels in order to
separate discrete senses of polysemous words. Whereas
the field labels provide the first subdivision, however,
the defining labels (or descriptors) act as further
specifiers of particular connotations of the main
divisions:

bone n [Anat] os m; [dead person]
ossèments mpl; [fish] arête f.

It goes without saying that translation equivalents must
conform to the SL word with respect to subject field
label. If an equivalent does not exist in TL, the usual
expedient is to furnish a succinct definition.

2.2.6 Lexicographers rarely deviate from the axiom that
translation equivalents must agree in terms of subject
matter, but the same is not true of words where a
particular register is involved; even though equivalence at this level is equally important, many dictionaries provide the information that a SL word is colloquial but then equate it with a more formal TL equivalent:

**policier** nm [Fam] detective novel.

The equivalence relation in such cases is incomplete; presumably the user is left to determine the appropriate word himself, a situation which would work, in the case of this example, for native English speakers and a few French speakers with an exceptional command of English.

As a formal feature of the equivalence relation, register can be handled as follows:

**policier** nm [Fam: detective novel] whodunnit, penny-dreadful; thriller; voir ROMAN POLICIER.

In the restructured article a defining label has been added, which users of both languages will be able to comprehend, three colloquial equivalents have been given and a cross reference to the more formal French word permits the reader to check further if necessary. This kind of arrangement of the material not only fulfills the requirements of the equivalence relation but also enables all users, even the young and linguistically naive, to reach the correct solution to an equivalence problem.
2.3 A final formal feature of equivalence is the question of the development of semantic relationships by means of cross-references. These elements of bilingual dictionary articles function in a number of different manners. In cases where absolute alphabetical order separates related lexical items, the affinity can be reconstituted by an appropriate reference. For example, expressly will be detached (in absolute alphabetical order) from express (verb, adjective and noun) by the following entries:

*express¹ vtr
express² adj
express³ n
expressage n
express boiler n
express delivery n
expressed adj
expressed almond oil n
expressible adj
expression n
expressional adj
expressionism n
expressionist n
expressionistic adj
expressionistically adv
expressionless adj
expressionlessly adv
expressive adj
expressively adv
expressiveness n
expressivity n
expressless adj
express letter n
*expressly adv

The relationship between expressly and express \(^2\) (adj) is shown by a cross-reference:

expressly adv [on purpose] exprès;
[stated] expressément; cf. EXPRESS \(^2\).

2.4.1 A cross reference may be employed to direct the user to another word which will provide further information on a lexical item which engenders a problem of a differential nature. For example, the frequently occurring phenomenon of deceptive cognates in French and English is such that, not only must the deception be signalled, but also the user must be directed to verify for himself the real meaning of the form in question:

actuality n [X] réalité f, existence \(f\) effective; \([\Rightarrow\) cf. ACTUALITÉ in Fr. section].

Sometimes the cross-reference serves to direct the user to a more appropriate term, as in the case of unaccept-
able Anglicism in Canadian French:

`plug` nf. C [Abus] cf. PRISE DE COURANT.

As the discussion on regionalisms will show, words of this kind must be recorded in the interests of the reader who encounters them and is unable to deduce their meaning from context; the cross reference enables the dictionary editor to recognize their existence while at the same time passing a veiled judgement on their status by directing the reader to the more usual term.

2.4.2 Finally, cross-references are used to establish semantic relationships considered important by the lexicographer. Usually the reference is designed to provide supplementary information:

- canoe nm [Naut] canoe; [» in Canada use CANOT, q.v.]
- as conj . . .; [» for special uses cf. FAR, FOLLOW, MUCH, SOON, WELL, YET].
- distinguish vtr distinguer (entrer, between); [sorts, etc.] déterminer; cf. DISTINGUER, DISCERNER.

It may also serve to suggest the full form of an element of a verb paradigm or the singular of an irregular plural:

- veux nmpl cf. OUI.
- threw v cf. THROW1.
3.0 **Differential Equivalence Problems**

In Chapter 5 of Part I we examined the phenomenon of semantic isomorphism which enabled us to pinpoint the theoretical basis for the divergence in meaning structure between pairs of languages. It was very clear that, even in closely related languages like French and English, lexical equivalence is unstable and difficult to establish. We will now proceed to identify different kinds of equivalence and the manner in which they are treated in bilingual lexicography.

3.1 Because of the complexity of lexical structure and the different ways in which it manifests itself in different languages, it will be no surprise that exact equivalence is a rare occurrence. One-to-one correspondence, when it does occur, is a characteristic of the words which appear at the limit of the scale of frequency distribution, that is, whose infrequent use tends to stabilize their semantic structure so that they do not develop the polysemous extension characteristic of lexical elements at the opposite end of the scale. Zgusta (1971) attributes this situation to the anisomorphism of languages which renders such "absolute equivalence" virtually impossible since it requires that "the two lexical units be absolutely identical, in all components (designation, connotation, range of applica-
tion). The kinds of words which will exhibit such absolute equivalence will be rare words and lexical items which, from the point of view of general language are of infrequent use even though they may constitute an important and frequent element of technical texts: mantissa/mantisse, oligarchy/oligarchie, aporism/aphorisme, ceint/girt with. Some pairs in this category are, in effect, the same word borrowed from the other language or coined simultaneously by the same technology: fuchsia, radon, hapax, bric-à-brac, hydrofoil. By its very nature, absolute equivalence is not a differential problem in bilingual lexicography; the only problematical issue concerns the decision whether or not to include such words.

3.2 The most common kind of equivalence is that which occurs between pairs of words which agree in some aspect of semic structure but diverge in others. This phenomenon is called "partial equivalence" by Zgusta (1971) because it is the result of the "partial identity" of the lexical items. It is determined by a distribu-
tional analysis which reveals those contexts in which pairs of lexical items function in an identical manner; the degree to which word pairs can be inserted into identical contexts is a measure of their equivalence. Polysemous words in two different languages rarely share
the same sense divisions and this is why partial
equivalence is not only a frequently-occurring phenomenon
but also one which causes the most difficulty in
establishing translation equivalents in a bilingual
dictionary.

Differences in the semantic extension of words are
illustrated by the diagrams in Figures 6 and 7. In
Figure 6 it is the French word *chou* which has a much
wider extension than its simpler English equivalent
*cabbage*, which is used only in the literal botanical
sense and as a slang metaphor for money.

**CABBAGE**

- *feuille de _, [newspaper] rag*
- *bête comme (un) _, simple, puerile*
- *être dans les choux, to be embarrassed*
- *entrer dans le _, attack, collide with*
- *faire _ blanc, to miss the mark*
- *faire ses choux gras, to profit from sth*
- *aller planter ses choux, to retire to the country*
- *ménager la chèvre et le _, to sit on the fence*
- *mon (petit) _, sweetheart, honey*
- *ce qu'elle est _, she's really nice!*
- *à la crème, cream puff*

**ROSETTE**

Figure 6. An example of differential extension: "chou."
By contrast, Figure 7 illustrates an example of a highly polysemous English word each of whose independent senses is represented by a different word in French. The word *bill*, in accordance with its position among the words of high frequency, has developed multiple meanings, but the most frequent of these (the commercial sense) has as an equivalent, French *compte*, which has a quite different range of meanings apart from the one shared with *bill*.

![Diagram](image)

Figure 7. An example of differential extension: "bill."

Thus, although the equation $\text{bill} = \text{compte}$ stands, it must be made clear to dictionary users that each word has its own semic structure which is by no means the equivalent of the other, as Figure 8 illustrates.
These structural differences are the most important factor in differential equivalence and must be constantly in the mind of the lexicographer as he develops the article for the word in question in such a way that the dictionary user is directed towards accurate discrimination between the senses of polysemous words.

Unfortunately, bilingual dictionaries rarely, if ever, take sufficient account of the phenomena of lexical structure and tend to leave the user to make many decisions which may be beyond his competence. B. Hunter Smeaton (1959) summarizes this situation as follows:

Most bilingual dictionaries, however, abrogate what might be considered their responsibility at a point far short of what might be hoped for. In the better (and costlier) dictionaries, one must grant, the uses of the entry word are parenthetically qualified, and sometimes illustrated in sentences. Even the best of them, however, do not take into account more than a small part of the structural ramifications of a language—that part which lexicographical tradition recognizes. One never completely escapes the notion that a word in language A equals a word in language B. 14.
3.3 A third kind of differential equivalence is what might be called "negative equivalence," where no lexical equivalent exists in the other language. The most usual reason for this state of affairs is cultural divergence which creates gaps in the lexical representation of the language which does not share a particular concept. These may take the form of what Zgusta refers to as "onomasiological gaps," where the concept is comprehensible in the other language but the term does not exist. Examples would be French Cegep (Collège d'enseignement général et professionnel, Québec), Cégepiste (member of the C.G.T. = Confédération générale du travail), Zupéen (relating to the Z.U.P. = Zone à urbaniser en priorité), Péquisté (member of the P.Q. = Parti québécois), germanopratin (= concerning the area of the rue St Germain-des-prés), Igame (= Inspecteur général de l'Administration en mission extraordinaire). In some cases an equivalent term exists but the derivational forms do not: O.N.U. = U.N.O.; but onusien (= concerning the United Nations Organization); drugstore = drug(-) store, but drugstorien (= concerning drugstores). Both languages use the word beatnik (originating in American English) in the same way, but French has derived in addition beatnikisé and beatnikisation. In general French produces a wealth of derivational forms for words which originally came from English and which do not exist.
in this state in English: speaker → speakerine (female announcer); footballer → footballeur → footballeuse (female football player); lift (British = elevator) → liftier (elevator operator) → liftière (female elevator operator); bikini → monokini (= topless bathing costume). Often the denoted concept is well known in both languages but only one has a lexical expression for it: chambrer (= to bring wine to room temperature); aquariophilié (= aquarium fish fancier).

3.3.1 Negative equivalence most often concerns the so-called "culture-bound" words which designate a cultural reality known in only one of the pairs of languages. Obviously this phenomenon will manifest itself most often between pairs of languages whose cultures are very divergent and be comparatively rare in languages which are closely related in socio-cultural terms. French and English are examples of the latter and it is virtually impossible to find examples of lexical items in one language which denote concepts which are completely incomprehensible to the other; although the concept itself may be lacking, its absence does not constitute a serious communication barrier. For example, the English term lunch box (or lunch pail) has an equivalent in Canadian French boîte à lunch, but in European French the expression is completely unknown on three counts: the
lexical combination does not occur, "lunch" would never be put in a "boîte" and Frenchmen always go home to eat lunch anyway. But this does not mean that European Frenchmen are unable to conceptualize a "lunch box" any more than their Quebec counterparts who have actually calqued the English word and applied it to the same reality.

3.3.2 For French and English, therefore, it is not the complete lacunae which cause the greatest difficulty since gaps in the lexical apparatus of one language can usually be satisfactorily filled by means of explanatory definitions in the target language. A more insidious problem is the fact that many translations are in fact only approximate equivalents and this is why bilingual lexicographers must determine the precise definitions of matching word pairs with scrupulous attention to detail. Our analysis of the characteristics of lexical structure revealed very clearly that even word pairs which appear to represent an identical concept more often than not differ in a number of subtle ways; the words are not simply different labels for the same thing:

Even in those areas where the two cultures overlap and where the material extra-linguistic world is identical, the lexical units of the two languages are not different labels appended to identical notions. In the overwhelming majority of cases, the designata are differently organized in the two languages.
For example, the equation *pain = bread* is approximately correct; the two words designate the same general kind of food, but on the other hand English bread has a different appearance and nutritional value from French bread. Similarly, *breakfast = petit déjeuner* and both refer to the morning meal, but the breakfast eaten in France bears little resemblance to the meal consumed in North America. 17

3.3.3 Sometimes it is not the concept that is lacking but the manner in which it is viewed by the two linguistic groups. A French butcher and an English butcher will carve up the same carcass in different ways which are related solely to the traditional methods of producing meat cuts for culinary purposes in the two linguistic areas. The resulting portions of meat not only differ in appearance but are differently named even when they appear to coincide and therefore constitute a serious translation problem. Other examples of lacunae resulting from metalinguistic factors are suggested by J-P. Vinay and J. Darbelnet (1958): *charcuterie* (pork butcher's shop) is comprehensible in Britain but not in North America where it does not exist (although the notion is clear enough); *hocher la tête* (shake the head in doubt) is difficult to render since the gesture is not used by English speakers, nor is there a single word for English
3.3.4 Vinay and Darbelnet also identify a category of differential equivalence in which one of the languages lacks a generic term for some concept: nuts (noix, noisettes, amandes, etc.); awards (prix, bourses d'études, distinctions honorifiques, etc.); utilities (l'eau, le gaz, l'électricité, le téléphone); promenade (walk, ride, row, drive, sail); ici (here, over here, down here, up here, in here, out here, back here, etc.).

The dichotomy general/particular frequently comes out in favour of English which usually specifies by means of individual lexical units whereas French must add a particularizing word to a generic term: siffler (whistle, hiss, swish); grincement (grating, screeching, squeaking, squealing, scraping, rasping, grinding); luire (glimmer, gleam, glow, glisten, glint); humide (damp, humid, dank, moist, clammy, dewy).

3.4 A final kind of differential equivalence problem concerns the deceptive cognates (or "faux amis") whose common genetic origin in the two languages leads the naive language learner to believe that he is dealing with a pair of nearly identical signs which designate the same concept or reality. Unfortunately, large numbers of such
pairs have undergone different diachronic development which has left them graphically equivalent but semantically divergent. The divergency may vary from a single sense difference: French journal = journal (periodical); but also and more frequently newspaper, whereas English journal more usually refers to the organ of a learned society (revue savante); oblitérer means to obliterate in figurative senses, but more often has the meanings to postmark, to cancel (a stamp), which are not shared by English. Sometimes derivational forms do not share the senses exhibited by base forms: loger = to lodge, but logeur = landlord, except in Canada where it may mean a lodger. Partial divergence is more often the rule, but total dislocation is not unknown: French versatile (changeable, fickle), English versatile (personne aux talents variés, souple); French fastidieux (dull, tiresome, tedious), English fastidious (difficile, exigeant; délicat).

The incidence of such spuriously associated pairs is a measure of the genetic proximity of the languages in question. A very large proportion of the lexical inventories of French and English have a common source, but the effects of centuries of isolation and antagonism have bestowed different meanings on so many of them that their synchronic status constitutes a serious translation problem. This aspect of differential equivalence is so
important to bilingual lexicography that the whole of Chapter 14 is devoted to an analysis of its causes, a categorization of its main features and the means of dealing with the problem in dictionaries.

3.5 The lexicographical treatment of the various categories of equivalence summarized in this section will clearly depend on the circumstances surrounding each particular lexical unit, but the general procedure must take into account all the features which enable meanings to be accurately discriminated and which have been discussed in earlier chapters.

3.5.1 Because absolute equivalence is not a differential problem, its lexicographical treatment will be restricted to the application of formal methodology which ensures equivalence of grammatical category, assignment of appropriate shared field and register labels, and so on. The equivalence relation in such cases is reduced to a simple equation of the type:

\[
\text{mantissa } n \quad [\text{Math}] \quad \text{mantisse } f.
\]

\[
\text{girt } \text{adj} \quad \text{ceint} \ (\text{de qqch, with sth}).
\]

3.5.2 Unhappily, the vast majority of word pairs do not belong to the absolute equivalence category and must be described with much more attention. If the language which tends to generalize is the SL the TL equivalents
must be defined carefully by means of discriminatory labels. For example, English to skin applies to a wide variety of objects but the French translations are specific:

skin vtr 1. [person, animal] écorcher, dépouiller; [banana, etc.] peler; [vegetable, chicken] éplucher. 2. [Fig] dérober, duper.

French piquant means a sting or prick of some kind which English specifies:

piquant nm 1. [plante, etc.] prickle, thorn; [insecte, etc.] sting; [porc-épic] quill, spine; [fil] spike. 2. [Fig] piquant, bite; [réponse, etc.] sharpness, pungency.

3.5.3 When the generalizing language is the TL, the entry is reduced to more simple equivalence; the pattern will usually be repeated for all the specific terms for which this language presents a general word, with the appropriate particularizing agent:

slam n [door] claquement m; the ~ of a door, le bruit m d'une porté que l'on ferme.

thud n bruit m sourd, bruit mat.

buzz n [conversation] murmure m: a ~ of voices, un bruit confus de voix.

pop n [Onomat] explosion f, détonantion f: the ~ of a cork, le bruit d'une bouteille qu'on débouche.
3.5.4 In the frequently-occurring cases where a word in one language has a broad semantic extension for which multiple equivalents are required, the article must be carefully subdivided into major and minor categories and the appropriate defining and particularizing labels must be included:


3.5.5 The various kinds of "negative equivalence" require particularly careful treatment. Lexicalized acronyms should be identified as abbreviations and the full form given together with appropriate subject labels; if no suitable translation exists, the lexicographer must resort to a translation-definition:

Cégep nm C [Educ = Collège d'enseignement général et professionnel] + community college.
Enarque nm Fr [Educ = former student of the École nationale de l'administration, E.N.A.] higher civil servant.

NDPer n C [Pol, de N.D.P. = New Democratic Party] membre m du Nouveau Parti Démocratique.

Derivative forms unknown in the other language must be carefully studied and suitably elucidated:

germanopratin, -e adj Fr [Urban] [>> adjectival derivative of St. Germain-des-prés] + of or concerning the area of the rue St. Germain-des-prés.

récréologie nf [Recré, Sci] + the science which studies leisure-time activities.


liftier nm [Occup] elevátor operator; liftman.

monokiniste nf topless bather.


deep-six vtr US [Slang = get rid of] bazarder.

high-muck-a-muck n [Slang, Chinook jargon] grosse huile f, gros légume m.

3.5.6 Words which designate cultural realia restricted to one of the language areas only must be interpreted in the dictionary by means of a translation-definition which, in effect, constitutes a miniature essay:
bowdlerization n + expurgation d'un texte (pièce, roman, etc.) pour en supprimer les passages considérés comme peu modestes (par pruderie); [>> d'après Thomas Bowdler (1754-1825), éditeur anglais d'une édition expurgée de Shakespeare].

block n [Urban: street divisions] [espace entre deux rues], pâté m de maisons: two blocks away, à deux rues de là; deux coins de rue plus loin.

hocher vtr to shake: ~ la tête, to shake one's head; [>> this gesture is one of doubt, in France, rather than negation, as is the usual case in English-speaking countries].

3.5.7 The partial equivalence which results from the fact that the two cultures have a different set of attitudes for the same reality must be explained as far as possible:

breakfast n [Culin] + petit déjeuner m; déjeuner m [C]; [>> le repas du matin en pays anglophone n'est pas du tout le même repas que celui des Français d'Europe; celui des Français canadiens peut être "petit" mais le plus souvent c'est le "breakfast" des Anglais; on peut distinguer entre petit déjeuner à l'anglaise et petit déjeuner à la française].

3.5.8 Lexical lacunae in the TL are resolved by means of a definition:

sibling n [Ethn] + le frère ou le sœur.

curb n [Road] + bordure f de trottoir, bord m du trottoir; chaîne f (du trottoir)
facilities mpl [Urban, etc.] [>> traduire selon le contexte]: to have the for, être/installé, outillé/pour: [>> ou préciser: bâtiments, pistes, locaux, etc.]: harbour-, installations fpl portuaires.

inadvertently adv + par/inadvertance, mégardé/.

cagneux, khagneux n [Educ] -> high school student preparing for examinations for entry to the École normale supérieure (lettres): [>> the word is an ironical derivation of cagne = lazy].

hypo-cagne nf [Educ] -> a high school class in the year before preparing for entry to the École normale supérieure.

dépouillement nm [texte] + careful scrutiny of documents, etc., in order to extract particular information.

3.5.9 Deceptive cognates require special treatment which will vary according to the circumstances of the divergence. The categories of these pairs and their lexicographical methodology are dealt with in detail in Chapter 14; a few examples here will suffice to demonstrate the problems and their solution:

violon. nm [Mus: instrument] violin; joueur, X] violinist, violin player.

Restauration nf [Hist, Pol, X: des Bourbons en France] the Restoration (1827); [>> ne pas confondre avec "the Restoration" des Stuarts en Angleterre (1660)].

croche nf [Mus, X] quaver, eighth note; [>> voir l'anglais crochet].

crotchet n [Mus, X] noire f; [>> cf. French croche].
4.0 **Idiomatic and Figurative Equivalence**

Among the equivalence problems which bilingual lexicographers are called upon to resolve are examples of idiomatic language, metaphor and words and expressions used in a figurative sense. Idioms are usually regarded as "lexical clusters" or "complex lexical items" whose meaning cannot be interpreted or predicted from the individual meanings of the words of which they are composed.\(^{21}\) **Metaphor** is the figure of speech in which a meaning is transferred from one item to another; there is generally a 'natural' connection between the primary referent and the secondary referent to which the lexical item in question applies.\(^{22}\) Dictionaries usually refer to metaphors as **figurative** usage.

4.1 Idioms are, by definition, idiosyncratic and this is the important justification for their inclusion in a bilingual dictionary. De Saussure referred to idioms as "locutions toutes faites" (ready-made utterances), by which he meant those expressions which are learned and used by native speakers as "unanalysable wholes" and which "resemble 'incomplete' sentences in that they do not correspond directly to sentences generated by the
grammar, but differ from them in that their description does not involve the application of the rules established to account for the vast mass of the more 'normal' utterances.”

Idioms are characterized, not only by the abnormality of their semantic import, but also in many cases by the existence of some syntactic peculiarity. For example, the expression to kick the bucket and its French equivalent casser sa pipe are employed almost invariably in the past definite tense and may never be passivized; ellipsis often reduces an otherwise recognizable expression to a state where analysis is almost impossible: pas de quoi < il n'y a pas de quoi < il n'y a pas de quoi . . . mentionner (don't mention it, you bet).

4.2 Although idiomatic expressions are classifiable as sentences in the sense that they are distributionally independent, they do not conform to the usual rules for the generation of sentences with acceptable combinations of lexical elements. This may prevent their inclusion in a formal grammar but, as Lyons points out, they must be accounted for if only "by listing them in the dictionary with an indication of the situations in which they are used and their meaning." If this advice is appropriate for monolingual dictionaries, it is even more important to bilingual lexicography where such expressions are
totally opaque to the non-native speaker and must be elucidated by a suitable equivalent idiom, if one exists, or by an explanatory definition in the target language.

4.3 In addition to idiomatic expressions, Lyons identifies types which appear to be grammatically unstructured but "which yet can be combined in sentences according to productive rules." For example, what's the use of [-ing]?, for (somebody)'s sake. The former expression represents what Lyons refers to as "sentence-schemata" because they form the base configuration for an indefinitely large number of sentences of the type: What's the use of/worrying, doing that, keeping chickens in the back yard, saving money by stuffing it in a mattress/??. He calls phrases of the second type "phrase-schemata" because they can not only be varied by filling the empty slot: for/my, your his sister's, the Prime Minister's/sake, but also may be used variously within sentences: I did it for your sake, Please do it for my sake, for goodness' sake stop talking, and so on.

4.4 There is little need to debate the question of whether such expressions should find their way into a bilingual dictionary. The problem is rather the manner in which they should be represented and the search for accurate and appropriate equivalents. The first decision to be made concerns the physical location of the phrase;
in spite of their lexical and semantic homogeneity, a case cannot be made for making them into entries and they must therefore be embedded in the article of one of the lexical units of which they are composed. The selection of this element, which applies to all syntagmatic expressions included in the dictionary, is based on parameters such as the specificity of the words and the need to include the expression under more than one heading. For example, the locution à mon corps défendant would be placed under "corps" or under "défendre" or under both; it would clearly not appear under "à" or "mon" whose generality would in theory make them eligible for the acceptance of literally thousands of expressions. Similarly, the English equivalent in self-defence would be located under "self-defence," if this were an entry, or under "defence" or under both.

Like any other information in a bilingual dictionary, idioms must be accompanied by whatever labels or explanations are necessary to avoid ambiguity and to ensure the accuracy of the equivalence relation:

bone n . . . . [Fig, Loc] to have a to pick with s.o., avoir maille à partir avec qqn.

face n . . . . [Fam, Loc] [embarras] Was my red!, Je ne savais plus où me mettre.

pipe nf . . . . [Fig, Pop] casser sa , to kick the bucket.
Sometimes the TL offers an equivalent which seems tailor-made for the SL original, as in the case of the last example where a similar kind of imagery is involved. In other cases, like the "red face" example, the lack of a similarly colourful equivalent is filled by an expression which conveys the notion of "embarrassment" implied in the idiom; the attachement of the word "embarras" as a definer makes the meaning clear to a French speaker.

4.5 Metaphor may be considered, from a lexicographical point of view, as a special case of polysemy since the extended meaning, usually from a literal to a figurative sense, applies to the same lexical unit and does not constitute a separate homonymous element. Metaphorical locutions thus differ from idioms in that they may have a literal interpretation as well as a metaphorical or figurative one:

Toute metaphore peut en effet se ramener à son sens fondamental, à ce que Bally appelle le terme d'identification.

For example, to flog a dead horse, mettre les points sur les i, to hang in the balance, and so on are readily interpretable as literal utterances even though they are more usually employed in figurative contexts.

This kind of polysemy is indicated in dictionaries as a figurative sense. In the case of lexical units, the
indication appears as a labelled, extended meaning which usually, although not invariably, follows the literal senses which are generally the more frequent or are historically antecedent:

\begin{itemize}
  \item elevation \textit{n} 1. altitude \textit{f}, cote \textit{f} d'altitude. 2. [Math] élévation \textit{f}.
  \item 3. [Mil] angle \textit{m} de hausse. 4. [Ind] montage \textit{m} (du grain, etc.). 5. [Geog] site \textit{m}, éminence \textit{f}. 6. [Fig: thought] élévation \textit{f}; [importance] grandeur \textit{f}, dignité \textit{f}.
\end{itemize}

If the figurative sense occurs as a natural extension of a particular meaning subdivision and in a syntagmatic group, it should be placed accordingly:

\begin{itemize}
  \item hot \textit{adj} 1. [temperature] chaud:
    \begin{itemize}
      \item [Fig] to be in water, être dans le pétrin,
      \item [pepper] cuisant;
    \end{itemize}
    \begin{itemize}
      \item 3. [Fig] violent;
      \item acharné: \textit{to be in pursuit}, serrer (qqn) de près.
    \end{itemize}
\end{itemize}

4.6 Figurative expressions, from the point of view of translation in a dictionary, fall into two groups; those for which there is an equivalent expression which employs virtually the same imagery:

\begin{itemize}
  \item clockwork \textit{n} [Mech] mouvement \textit{m} (d'horlogerie); rouages \textit{mpl} (d'une montre): [Fig] \textit{to run like clockwork}, marcher, se passer comme sur des roulettes.
\end{itemize}

and those for which the target language has no metaphorical equivalent, in which case a rendering of the sense is
given or else a close figurative equivalent may be suggested: \textit{la marche a suivre} (the procedure); \textit{as cool as a cucumber} (avec un sang-froid parfait); \textit{before you could say Jack Robinson} (en moins de rien; en deux temps trois mouvements; en criant ciseaux C). 29

4.7 Under the general heading of "expressions toutes faites" we will also consider the set phrases ("locutions figées") or clichés which are an important part of the lexical composition of all languages. All of these multi-word expressions constitute what Vinay and Darbelnet refer to as "unités de traduction" (translation units) which are generally non-decomposable wholes. 30 As far as possible, translation equivalents should be similar constructions: \textit{mener une vie de chien} (to lead a dog's life); \textit{you could have heard a pin drop} (on aurait pu entendre voler une mouche). 31

Since the dictionary user needs to be informed that he is dealing with a particular kind of expression, and not merely a contextual example, bilingual dictionaries should employ a distinctive label to distinguish set phrases from those which are not. The label [Loc] used in CD1 and CD2 seems to cover this requirement adequately:

\textit{partie} \textit{NF part} . . . [Loc] \textit{en grande} largely; \textit{prendre à}, to take to task.

\textit{just} \textit{adv} juste, exactement . . . [Loc]
I'd just as soon . . . , Cela m'est égal de . . . .
Locutions of this kind are frequently associated with a particular subject matter or language level and these should clearly be marked as associated labels:

mad adj fou, insensé; furieux: [Loc, Fam] to work like _, travailler comme un perdu; to be _ about golf, être un joueur de golf enragé.

chair n ... [Admin, Fig: meeting] to take the chair, présider (une réunion, etc.); to leave the chair, lever la séance.

suffoquer vtr to suffocate, to choke: [Fig, Loc]/être, rester/suffoqué, to be flabbergasted.

5.0 Language Level Equivalence

Translators are duty bound to preserve all stylistic aspects of the original, including those which relate to register, a somewhat delicate matter in lexicography as we shall see. In order to maintain an impartial and faithful rendering of an original text, a competent translator will take into consideration all the nuances and subtleties which the author injected into his writing and which endow it with the particular stamp of his handwork. These elements of the original, according to Vinay and Darbelnet,^32 make up the "tonality" of the text and, in contrast to a general range of language levels, determine the choice of translation equivalents in the same register. In lexicography it is generally assumed that the entry words and their equivalents belong to the
common core of language and that deviations from the "norm" are indicated by labels which specify the level to which they belong.

5.1 The problem with determining these levels and marking them in the dictionary is that someone has to make what amounts to a value-judgement on the status of each word, and this is usually a purely subjective assessment which is unlikely to satisfy all speakers of the language or all users of the dictionary. This is by no means a new problem. Even Samuel Johnson wrestled with the controversial issue of what to include in the dictionary and how to describe it:

With regard to questions of purity, or propriety, I was once in doubt whether I should not attribute too much to myself in attempting to decide them, and whether my province was to extend beyond the proposition of the question, and the display of the suffrages on each side; but I have since determined by your Lordship's opinion, to interpolate my own judgement, and shall therefore endeavour to support what appears to me to be the most consonant to grammar and reason; . . . And I may hope, my Lord, that since you, whose authority in our language is so generally acknowledged, have commissioned me to declare my own opinion, I shall be considered as exercising a kind of vicarious jurisdiction, and that the power which might have been denied to my own claim, will be readily allowed me as the delegate of your Lordship. 33

Johnson actually went to great lengths to give his
personal opinions on many of the words in his dictionary; sometimes a neutral word was endowed with pejorative tones, as in his definition of *excise*: "a hateful tax levied upon commodities and adjudged not by the common judges of property, but wretches hired by those to whom the excise is paid"; and *lexicographer*: "a writer of dictionaries, a harmless drudge, that busies himself in tracing the original, and detailing the signification of words." While he did not use actual labels to indicate levels of register, Johnson had a series of repetitive phrases which amount to the same thing: "a low word," "a barbarous contraction," "an unauthorized word," "a word out of use and unworthy of revival," "a word introduced unnecessarily from the French by a Scottish author."

5.1.1 The difficulty with all restrictive or prescriptive labels in a dictionary is that what may be seen to be unacceptable or limited in application by one generation, or even by one level of the same generation, may be viewed quite differently by another. Thus, for example, Chaucer employed the word *fart* with almost reckless abandon, and four centuries later Johnson defined the word without comment on its usage. Yet Mansion (HAP3) found it necessary to append the warning "not in decent use" to his entry for the word earlier in this century and RHD labels it "slang and vulgar." Presumably modern
speakers of English, and those who would learn the language, are being directed by lexicographers to employ euphemisms such as "break wind" or medical circumlocutions like "expel a flatus through the anus," expressions which are unknown to a large majority of native English speakers. Although it may be necessary for lexicographers to pronounce some kind of judgement on the entries of their dictionaries, they should be particularly wary of those words which are subject to the vagaries of social acceptability and continuous change. R. L. Wagner (1967) sums the situation up thus:

Les vocabulaires, correspondant à des situations, partagent le lexique en aires. Mais au sein de chaque vocabulaire d'autres répartitions s'opèrent entre les mots. Tel, par exemple, leur partage selon les niveaux de langue. Une société repose sur des conventions. Celles qui portent sur la tenue engagent la langue autant que les gestes ou tout autre signe qui caractérise une certaine façon de se comporter. Le contenu de ces conventions est changeant; cela tient à ce que varient d'une manière concomitante les milieux qui les édictent. On ne s'étonne donc pas que des mots, admis aujourd'hui dans le vocabulaire d'un homme distingué, aient été naguère ou jadis proscrits comme laids ou vulgaires, l'inverse étant aussi vrai.

5.2 Bilingual lexicographers are in a particularly difficult situation with regard to register. Not only do they have the responsibility for the selection or rejection of various lexical items, but they must also
ensure that the non-native speaker is given all pertinent information about the synchronic status of such words and provided with an appropriate equivalent. Once the decision has been taken to include a word, probably on the basis of statistical frequency of occurrence, it must be assigned to a level in an arbitrarily established hierarchy of sociolinguistic usage, such as that suggested by Vinay and Darbelnet (1958) and translated below in Figure 9. The schema is based on Ch. Bally's adaptation of Saussurian terminology in his *Traité de stylistique française* (1951) and includes divisions between "bon usage" (acceptable style) and "vulgar language" on the one hand and between aesthetic and functional specialization on the other.35

<table>
<thead>
<tr>
<th>ACCEPTABLE STYLE</th>
<th>VULGAR</th>
</tr>
</thead>
<tbody>
<tr>
<td>poetic language</td>
<td>popular language</td>
</tr>
<tr>
<td>literary language</td>
<td>slang</td>
</tr>
<tr>
<td>written language</td>
<td>administrative, judicial, scientific, etc.</td>
</tr>
<tr>
<td>familiar language</td>
<td></td>
</tr>
<tr>
<td></td>
<td>jargon</td>
</tr>
</tbody>
</table>

Figure 9. Language Levels
While a translator faced with a text must make many decisions himself on the intent of the original writer, dictionaries are obliged to indicate those usages which deviate in some way from the theoretical "norm." In terms of the hierarchy described in Figure 9, the labels should include "poetic," "literary," "familiar," "popular," "slang" (French argot), "vulgar," or some variation of these. Some modern monolingual dictionaries, in an attempt to inject an element of increased objectivity into the designation of language levels, have reduced the number of labels to two: "formal" and "informal," thereby "focussing upon the style or atmosphere of the language rather than the mode of communication--speech or writing." 36 Webster's Third was bold enough to eliminate altogether indications of language level, basing their practice on the belief that there is, after all, no clear line of demarcation between the levels which often overlap each other and change from one social group to another and from one generation to another. The editors considered that the presence of adequate citations accompanying each entry would be sufficient to enable users of the dictionary to make their own judgements based on the evidence presented to them. But the hostile reaction which greeted this and several other of Webster's innovations should sound a warning to other lexicographers who feel inclined to
shift the burden of responsibility for register indication onto the user.

5.3 Unlike monolingual dictionaries, bilingual dictionaries have a particular responsibility to non-native users, somewhat akin to the accountability of editors of school texts for what they suggest to young readers. As a general rule, the non-native user of a bilingual dictionary cannot be left to make his own judgements about the suitability of a given word to a given context in terms of its sociolinguistic affective value. But if a label is appended and an equivalent translation given he should be in a position to determine whether or not the word in question is appropriate to his purpose.

It is an unfortunate fact that bilingual dictionaries frequently fall short of the goal of the equivalence relation in this important feature, either because they fail to give an adequate description of the lexical item in question or, and this is a more serious defect, because they do not provide an equivalent of the same language level. The French expression en rester baba, for example, is rendered by HAP3 as to be struck all of a heap which belongs to the rather restricted parlance of elderly ladies and can hardly be described as "familiar"; a more appropriate rendering would be: to be
thunderstruck, to be flabbergasted. The same remarks apply to vache in its popular sense ("personne méchante, qui ne passe rien, se venge ou punit sans pitié") which HAP3 gives as disagreeable fellow, a neutral and very mild expression; a closer equivalent would be: cad, rotter or punk, louse, although the latter two words are probably a degree lower on the scale of language register. The word salaud, which Robert describes as "popular" and applicable to a person who is "despicable and morally repugnant," is labelled by CDI as "vulgar" and translated by stinker, bastard; the latter are slang words, certainly not vulgar and therefore appropriate equivalents for salaud according to Robert's description of it; in this case, therefore, it is the label that is inaccurate. HAP3 describes foutre as "vulgar and not very decent," declining to offer an English equivalent; the expression foutre le camp is given, however, and inadequately rendered by to decamp which is a completely neutral expression. In this particular case, however, since the four-letter word equivalents of the French phrase are seldom committed to the printed page (or rather, this was the situation at the time of publication of HAP3), the use of a neutral term is perhaps justifiable; armed with the information that the term is "vulgar" and the definition equivalent, an English speaker may be best left to decide for himself how to
translate the French expression in the context in which he has encountered it. Alternatively, less-offensive but equally slang expressions could easily be suggested: to split (the scene), to scram, to take off.

Sometimes a neutral SL word is translated by a neutral TL word and followed by a familiar or slang term. This is not to be construed as being a breach of the equivalence relation, but rather an attempt on the part of the lexicographer to develop semantic relationships considered of benefit to the user, or to remind readers of the correct level of synonyms which may suggest themselves:

```
mangle vtr lacérer, déchirer; [Fam] charcuter.
```

In other cases, a familiar lexical item is rendered first by a neutral term, by way of definition, and then by a word at an equivalent level:

```
junket n [Fam] voyage m officiel/paié par le gouvernement, [Fam] aux frais de la princesse/.
```

5.4 The general rule, therefore, with respect to register levels in a bilingual dictionary is that, the decision to include a word or expression having been made by whatever criterion is being applied, there must be a
determined effort either to provide an adequate equivalent of the same or similar register. If this is not possible, and if it is necessary to provide an unmarked equivalent, at least the marked word must be clearly labelled so that the user is in a position to make his own assessment. It is, after all, just as important that levels of usage between equivalent lexical units correspond as it is that they conform with respect to synchronic status or subject-matter. An archaism should only be equated with a modern equivalent if no parallel archaism exists in the other language; in this case, just as with register, the translation is really a definition:

\[ +\text{smite vtr frapper, cogner.} \]
\[ +\text{liesse nf joy, glee} \]

The same guidelines apply to neologisms which may have parallel development in the TL:

\[ \text{pig n } [\text{Police, Slang} \text{gorille m}.} \]
\[ \text{scene n } \ldots [\text{Loc, Slang} \text{to make the} \ldots \text{avoir le ticket.} \]
\[ \text{camé nm } [\text{Drogues, Pop} \text{junkie.} \]
\[ \text{alcotest nm } [\text{Admin, Police} \text{breathalyser.} \]
\[ \text{monokiniste nf } + \text{topless bather.} \]

Finally, although called upon to exercise judgement
on the sociolinguistic status of lexical items, lexicographers should generally refrain from making subjective comments about them. For example, HAP3, in describing the Anglo-Saxon institution known under the name of "finishing school" and for which there is no simple French equivalent, added a somewhat unexpected and uncalled-for remark to his translation-definition: "école où l'on paracheve l'éducation des jeunes filles (croit-on)." This does not preclude, however, an informational remark, even though it may be construed as whimsical or even humorous, as the following example from CD2 illustrates:

\textit{drudge n; souffre-douleur m, femme f/de peine: harmless \_\_\_, lexicographe m [Samuel Johnson dixit].}

6.0 Practical Equivalence Problems

In Chapter 10 of Part II attention was drawn to the desirability of collecting material for a bilingual dictionary according to subject-headings. This procedure enables control to be exercised over content because information is accumulated in an orderly and predetermined manner which ensures that quantities remain constant and conform to acceptable frequency distributions. Since the lexicon of a closed field, such as mathematical terminology, tends to be relational and
complementary with the same lexical items outside the field, a number of interesting and informative facts emerge from its close study. These kinds of data can be transformed by the lexicographer into commentaries and cross-references likely to be of benefit to the user. For example, it may be possible to develop semantic correlations by means of cross-references for the SL and synonyms for the TL; important syntactic data relating to the use of a particular terminology may emerge; repetitive information, such as a constantly recurring suffix, can be highlighted; differential problems can be observed in grouped combinations which make their lexicographical solution more obvious, and so on.

6.1 Musical terminology in French and English presents a number of problems contingent on the fact that some of the word pairs are deceptive cognates. The following is a selection to illustrate the difficulty:

<table>
<thead>
<tr>
<th>ENGLISH</th>
<th>FRENCH</th>
</tr>
</thead>
<tbody>
<tr>
<td>quaver</td>
<td>croche, crochet</td>
</tr>
<tr>
<td>crotchet</td>
<td>noire</td>
</tr>
<tr>
<td>minim</td>
<td>blanche, minime</td>
</tr>
<tr>
<td>breve</td>
<td>brève, carrée</td>
</tr>
<tr>
<td>semiquaver</td>
<td>double croche</td>
</tr>
<tr>
<td>semibreve</td>
<td>ronde, semi-brève</td>
</tr>
<tr>
<td>demisemiquaver</td>
<td>triple croche</td>
</tr>
</tbody>
</table>
6.1.1 The crotchet/quaver problem referred to earlier is intensified as we follow the progressive note divisions in each language: a semiquaver, instead of being a "semi-croche" is a double croche because, while English views the consecutive subdivisions as fractions of the preceding ones, French refers to the number of "crochets" carried by the note; in other words, the English note system is purely mathematical whereas the French is physical as well as mathematical. The result of this divergent attitude towards the same reality is that English speakers must contend with a double deception: the translation of quaver (or eighth note) by croche(t), which does not mean English crotchet, and the doubling instead of halving of the sixteenth, thirty-second and sixty-fourth notes. For French speakers the discrepancy is no less ambiguous, except that they must bear in mind that croche(t) does not equal English crotchet.
6.1.2 The denomination of musical instruments is also a source of confusion, particularly because French almost invariably uses the same word for the instrument and for the person who plays it: cor = horn, horn-player; trompette = trumpet, trumpeter. In English the instrument may designate the player only in the restricted sense of a conductor addressing a section of the orchestra, although this really constitutes a kind of synecdoche rather than an actual lexical meaning difference. A divergence in the semantic extension of the word musique may cause ambiguity to both language groups and the translation of French horn by cor d'harmonie seems, to English speakers at least, an unwarranted deviation from the natural equivalence displayed by the pair English horn/cor anglais. All of these various anomalies, as well as others which would be revealed by further study of the terminology, may be dealt with in a bilingual dictionary according to the guidelines set out in earlier sections of this chapter. In the case of the lexical items described here, the use of a symbol signifying "deceptive cognate" followed by a brief commentary or cross-reference should be sufficient to eradicate ambiguity; further clarity would be introduced by the addition of the illustrations of the notes themselves with a description of their respective values.
6.1.3 Another closed field demonstrating similar divergences is the terminology of chess and checkers:

<table>
<thead>
<tr>
<th>ENGLISH</th>
<th>FRENCH</th>
</tr>
</thead>
<tbody>
<tr>
<td>chess</td>
<td>échecs</td>
</tr>
<tr>
<td>chequers</td>
<td>(jeu de) dames</td>
</tr>
<tr>
<td>chequerboard</td>
<td>damier</td>
</tr>
<tr>
<td>chessboard</td>
<td>échiquier</td>
</tr>
<tr>
<td>pawn</td>
<td>pion</td>
</tr>
<tr>
<td>rook, castle</td>
<td>tour</td>
</tr>
<tr>
<td>queen</td>
<td>reine, dame</td>
</tr>
<tr>
<td>bishop</td>
<td>fou</td>
</tr>
<tr>
<td>king</td>
<td>[chess] roi; [chequers] dame</td>
</tr>
<tr>
<td>knight</td>
<td>cavalier</td>
</tr>
</tbody>
</table>

The self-evident ambiguities in this series demand the same attention as those described for musical terms.

6.2 Neology, as we observed above (5.4), is a constant cause of translation difficulty. If the concept described by the new terminology is common to both language groups there is usually no problem in establishing the equivalent: draft dodger/embusqué, energy crisis/crise de l'énergie, direct distance dialling/interurbain automatique, lunar module/train spatial, to go on a (drug) trip/aller en voyage, wildcat strike/grève sauvage, pollutant/polluant, organ
transplant/greffe, streaking/nuvitesse, mini(-skirt)/mini, midi/cache-genou, maxi/frôle-cheville. But if cultural exclusivity comes into play, the lacunae must be dealt with in the usual way, that is, by translation-definition: aoutien (August holiday-maker), pinch hitter (remplaçant d'urgence). Sometimes the source of difficulty is the fact that one language creates derivational forms unknown in the other language. Apart from monokiniste (topless bather) mentioned above, the following examples are typical of this state of affairs: cover-girlariat (the cover-girl syndrome), jusqu'au-boutisme (extremism), cruciverbiste (crossword puzzle buff), liftière (female elevator operator), artisane (female skilled worker), to tailgate (suivre une voiture de trop près), récréologie (the science of leisure-time activities), cinémomètre (radar speed detector), cinéroute (drive-in theatre).

6.3 Certain structural features of the lexicon lend themselves to analysis with a view to providing dictionary users with important information on relational characteristics, especially those which are recurrent and which therefore constitute vital elements of general vocabulary knowledge. In the sciences, the alternation -ics/-ique is useful data to enable a large group of words to be condensed to their main roots: physics/
linguistique, but exceptions must also be observed: statistics (= figures, results) is des statistiques; French logique is logic. Similar attention can be given to common suffix pairs like -al/-ique (peripheral/périphérique), -ist/-iste (pianist/pianiste), -ic/-ique (lactic/lactique), -ian/-ien (musician/musicien), -ness/-tude (uprightness/rectitude), -ful/-ble (hateful/haïssable), and many others. On pedagogical grounds these suffixes should be recorded in the bilingual dictionary, together with explanations and examples designed to suggest to the user relationships which may be transparent to an experienced lexicographer but which generally elude a young language learner.

The English suffix -less, for example, although it is not an independent lexical item, occurs with such frequency and generates such an abundance of derivational forms with more or less constant French equivalents that an entry is justified:

- less suffix [>] affixe qui permet d'établir une très grande série de mots dont l'équivalent en français peut être une circonlocution ou un mot plutôt savant avec préfixe en in-, im-, etc. sans . . . , dénué de . . . : countless, sans nombre, innombrable; lifeless, sans vie, inanimé.

Similar considerations apply to French compounds in porte-, coupe-, sèche-, aide-, and so on. A commentary
preceding the alphabetical grouping of a series of such compounds is not only useful as a means of pointing out interesting and beneficial relationships but also in order to mention a repetitive feature, such as the fact that the whole series is made up of masculine invariable nouns, and so on.

6.4 It is to be expected that the vocabulary of a science or technology which has developed contemporaneously along more or less parallel lines in two language areas will correspond fairly closely and this is usually the case. But the terminology of a well-established human endeavour may develop idiosyncratic tendencies like those which occur in other areas of the lexicon. For example, in mathematical terminology there are large numbers of word pairs of the type: \textit{concentric/concentrique}, \textit{function/fonction}, \textit{bilinear/bilinéaire}, \textit{angular/angulaire}, where equivalence is structurally as well as semantically near perfect. But there are also examples of differential semantic extension:

\begin{itemize}
  \item \textit{equivalent} n \ [\text{Math: vectors, matrices, elements}] équivalent; [sets] équipotent; [vectors] équipollent; [modules] semblable.
  \item \textit{graph} n \ [\text{Math: of two variables}] graphe n; [of a function] graphe m.\item \textit{excentre}
\end{itemize}

Sometimes a single word is lacking and a definition must be given: \textit{addend} (un des nombres additionnés);
(centre du cercle exinscrit); **completing the square** (méthode à faciliter la résolution d'une équation du second degré); **improper fraction** (fraction dont le numérateur est plus grand que le dénominateur; le mot *impropre* est réservé pour les sous-ensembles et les valeurs absolues); **latitude** (angle of elevation); **clock-wise** (dans le sens des aiguilles d'une montre).

Other mathematical information of interest to bilingual dictionary users is the fact that the English-speaking world employs a decimal point to mark off decimal fractions and a comma to indicate thousands, the opposite being the case in French, and the fact that, although French and North American English agree on *dimensional numbers*: **million**, **billion**, **trillion**, the latter two words mean something different to speakers of British English: **billion**: $= 10^9$ (France and North America), $= 10^{12}$ (Britain and Germany); **trillion**: $= 10^{12}$ (France and North America), $= 10^{18}$ (Britain and Germany).

6.5 Colour terminology, the preferential source of information used in ethnolinguistic studies of lexicocultural divergence, is no less interesting in French and English as a comparative study and as a lexicographical problem. The basic colours of the spectrum coincide more or less precisely between the two languages, although there are some notable differences. Vinay and Darbelnet, for example, demonstrate that French is more discriminat-
ing in various shades of brown: roux, brun, bistre, bis, marron, jaune, gris, as in yeux bruns (brown eyes), beurre roux (brown butter); crayon bistre (brown pencil), chaussures marron (brown shoes), pain bis (brown bread), papier gris (brown paper), cheveux châtains (brown hair). 37 While it is true that English may employ the more generic brown for these expressions, this is not to say that other nuances are not possible. "Brown eyes," in fact, are more often described as hazel; "butter" is usually described as having been "browned" in the process of cooking; "shoes" are often specified as being tan; "hair" may be chestnut, auburn or brunette. And we may call upon a number of other adjectives to depict varying hues of brown: bay, sorrel, bronze, coppery, cinnamon, russet, chocolate, walnut, ginger, mahogany, khaki, cocoa, coffee, camel, tobacco, mustard, beige, honey, sand, caramel, tortoise shell, almond. Virtually any naturally occurring element is susceptible to colour attribution in both languages: brun gingembre, vert sauge, aubergine, framboise, violette, gris argent, orange automne, fumée avocat soutenu, ébène, champagne, menthe, etc.

6.5.1 Even at the level of less common colours, those which are designated by metaphorical (and often highly exotic) terminology, equivalence is quite close: vert
bouteille (bottle green), vert cendré (sage green), vert chou (cabbage, lettuce green), vert pomme (apple green),
vert de mer, vert d'eau (sea green), vert jaune
(chatreuse green), vert avocat (avocado green); eggshell
(white) (coquille d'œuf), bone white (os), mushroom,
off-white (huitre), mousse jungle (jungle green), vin
(wine-coloured), tartan (écossais), salmon pink (saumon),
garnet red (grenat), snow white (blanc hiver).

6.5.2 In figurative applications, however, the two
languages are frequently widely divergent, even
diagonally opposed:

bleu = greenhorn; bruise (black and blue);
[steak] rare (and therefore red).
êt re bleu = to be green (i.e., inexperienced)

n'y voir que du bleu = to see through rose-
 coloured glasses

avoir une peur bleue = to be frozen (pale,
white) with fear
êt re dans une colère bleue = to be/purple, livid/
with rage

petit bleu = (light) red wine
le pays du bleu = the land of (golden) dreams
con te bleu tall story

the blues = idées noires

blue with cold = vert de froid

in a/blue fit, purple rage/ = dans une colère
jaune
blanchir = (hair) to turn grey, to grey
blanc-bec = greenhorn
pourpre = crimson, rich red
le pourpre lui monta au visage = he turned/crimson, purple/with rage
histoire verte = spicy story (blue)
	en raconter des vertes = to tell tales to make even a sailor blush
feu jaune = amber light

Organized in topical groupings of this kind, the relationships as well as the discrepancies come clearly into focus and suggest methods of arranging the material in a lexicographical context.

6.6 In describing language registers lexicographers must be extremely perceptive, as we have seen, in order to ensure that non-native readers of the dictionary are not placed in embarrassing predicaments through careless level designation. They must also, as some dictionary-makers have learned to their cost, avoid saying anything which is detrimental to a particular ethnic group. The repercussions, in this age of nationalistic sensitivity, are likely to be out of all proportion to the minor sin of omission--the failure to append the label "pejorative" to an ethnic term of abuse or derision.38

6.6.1 The lexicographer is in a particularly difficult
position. If he chooses to omit all derogatory ethnic terms he leaves himself open to censure for excessive purism and, worse still, the certain accusation that he has exercised his partial judgement on one small portion of the lexicon while leaving other equally abusive and defamatory elements unimpaired simply because the public reaction to them is less sensitive. If, on the other hand, he opts for global inclusion of such terminology, even though he may be scrupulously fair and provides both sides of every situation, he will have to be prepared to justify every entry and risk the consequences of a commercial boycott which any small but vociferous and articulate group is capable of mounting.

6.6.2 Those who will not tolerate inclusion of this kind of information in the description of a language are closing their eyes to reality; xenophobia is a fact of life and linguistic xenophobia is to be found in all languages. French and English have proved to be particularly creative in this regard. Sometimes the national terminology employed is equivalent in both languages: street Arab (arabe, gamin des rues), grec (card-sharp, Greek), Yank (Amerlo, Ricain, Amerloque), Jerry, Kraut (Fridolin, Fritz, Boche), Wop (Rital). In other cases the equivalence is provided in other terms: chingqiseries (red tape), to turn Turk (faire le méchant),
a young Turk (un enfant terrible), a real Turk (un tyran), ivre comme un Polonais (drunk as a lord, a Dutchman/), Dutch courage (courage/arrosé, puisé dans la bouteille/), Dutch comfort (piètre consolation), réponse normande (non-committal reply), parler français comme une vache espagnol (to murder French), boire le trou normand (take an alcoholic drink between courses to whet the appetite), German measles (rougeole bénigne), apache (hooligan), Scotch mist (bruine), to scotch (mettre hors de combat), to welsh (escroquer, lever le pied), Irish temper (caractère emporté), English Channel (la Manche), latrine à la turque (seatless toilet), c'est une américaine (she plans her pregnancies), travail de Romain (Herculean task), Roman nose (nez/quilin, busqué/), French heels (talons Louis XV), French curve (pistolet), French knitting (queue de rat).

6.6.3 In many cases the ethnic element is altered in the other language: French seam (couture anglaise), French wrestling (lutte gréco-romaine), French cleaning (nettoyage américain), Indian summer (été de la Saint-Martin, été des sauvages), Italian handwriting (écriture anglaise), Turk's head (tête de Maure), Ascot hat (chapeau Gainsborough), Epsom salts (sel/anglais, d'Angleterre/), it's/all Greek, Double Dutch/to me (c'est/de l'hébreu, du chinois/pour moi), French letter (capote
anglaise), filer à l'anglaise (to take French leave), French kiss (baiser à l'italienne). 39

6.6.4 Lexicographical treatment of this terminology entails the same considerations as any other lexical field with the important rider that, in the case of offensive materials, clearly-worded commentaries must indicate its derogatory nature. The physical location of phrasal examples, as we saw in Chapter 11 (Structural Methodology), is a function of the key words, which will invariably mean the ethnic element although repetition under some other focal word in the phrase is not precluded.

7. Conclusion

Differential lexicography is thus crucial to the whole question of bilingual dictionary-making. It is founded on the premise that the semantic structures of any given pair of languages are sufficiently divergent to warrant the entire attention of lexicographers in their attempts to establish the equivalence relation. In the last two chapters we have examined specific details in grammar and semantics and demonstrated the manner in which they must be treated in their dictionary context. In Chapter 14 we focus more closely on a particular differential problem—the existence of deceptive cognates which, for English and French at least, constitute one of
the most serious hazards in the transfer of meaning from one language to the other and therefore merit a separate discussion.
Notes to Chapter 13


2 See details in Chapter 11 above.

3 The last example is an extract from a lengthy explanation in both languages for the English prefix "un" in CD2.


5 Darbelnet (1970: 94) uses this term for divergent semantic extension. See also note 13 below.

6 Iannucci (1957), p. 278.

7 The question of lacunae is discussed in further detail below (3.3).

8 See the more detailed analysis below (5.0).

9 The question of deceptive cognates is treated in detail in Chapter 14.

10 See the detailed discussion in Chapter 15.

11 This is the well-known phenomenon of the positive ratio between frequency and number of senses as demonstrated by G. K. Zipf (1949).


13 Ibid., p. 312. Darbelnet (1970) uses the term "bifurcation," which is actually inappropriate for the description of this phenomenon since the word means division into two parts and the more usual situation is a division into more than two senses.


16 Ibid., p. 296.
18. Ibid., p. 69.
20. Ibid., pp. 60-66. See also the discussion on "coup" in Chapter 6, Part I.
23. Ibid., p. 177.
24. Ibid., p. 177.
25. Ibid., pp. 177-8.
26. Ibid., p. 178.
30. Ibid., p. 252.
32. Ibid., p. 33.
The Oxford Dictionary found itself involved in litigation over the promulgation of the word "jew" defined in a pejorative sense without indication of its restricted usage.

References for Chapter 13


CHAPTER 14

LEXICOGRAPHICAL PRESENTATION OF "FAUX AMIS"

1.0 A Deficiency in Bilingual Lexicography

The existence of deceptively similar word pairs in related languages has been a source of concern to generations of language teachers endeavouring to impress on their students the importance of being cautious about the too obvious translation equivalent. It was Koessler and Derocquigny (1928) who first sought to deal with the problem in a systematic manner by establishing a list of some 1,300 English words which they described as "trahisons du vocabulaire anglais" and which they dedicated to "ceux qui s'intéressent aux études anglaises et notamment aux traducteurs." Their lexicon was followed by a succession of studies: Boillot (1930), Derocquigny (1931), Las Vergnas (1947), Vinay et Darbelnet (1958), Kirk-Greene (1968), Rancoule et Péan (1971), the "fiches terminologiques" of Radio Canada, various publications of the Office de la langue française au Québec, not to mention passing reference and condensed lists in a variety of educational text books.
1.1 In spite of the lead provided by these studies, however, the bilingual lexicographers have, in general, failed to account for the phenomenon in their dictionaries. Among French and English bilingual dictionaries, only OXF and CD1 deal with the question by means of a system of warning symbols which draw the user's attention to the presence of an ambiguity.

OXF devotes two precious pages of preface space to the problem of "false friends," discussing their origin and characteristics, their treatment by Koessler and Derocquigny (1928) and the desirability of dealing with them in the context of a bilingual dictionary. But only about 600 "signs of danger" are to be found in this dictionary of some 40,000 entries, and this can hardly be considered an adequate treatment; a rapid consultation reveals a number of quite obvious omissions such as furnace/fournaise, to bomb/bomber, maternal/maternel. Those which are pointed out, however, are treated thoroughly and leave the reader in little doubt as to the correct translation to employ in a given circumstance. Moreover, it must be conceded that this was a first attempt to treat the phenomenon from a lexicographical point of view, and the authors admit that they have far from exhausted the subject. 3

1.2 The editors of CD1, equally sensitive to the
pedagogical aims of a translation dictionary, adopt similar methods to those of OXF in order to point to the presence of "faux amis." Although the treatment is somewhat more systematic than in OXF, there are still notable omissions, in particular the failure to mark the deception in both lexicons. For example, the French section marks *éventuellement* as a "faux ami" (by means of the symbol \([\times]\)), giving the correct renderings as:

(1) \[ \textit{éventuellement adv [\times] on occasion; if need be.} \]

But its English counterpart, in which the deception is equally apparent, is neglected:

(2) \[ \textit{eventually adv finalement, en fin de compte.} \]

Omissions of this kind may be attributed to a number of factors, chief among which is the actual size of both of these dictionaries which are concise editions and therefore could not hope to achieve an exhaustive coverage of all elements of confusion in the parallel lexicons. Thus, for example, OXF has the adjective *eventual* but lacks the derivative adverb *eventually*, assuming that the average dictionary user is capable of generating it for himself. On the other hand, some omissions can be traced to inexcusable error: OXF gives both *éventuel* and *éventuellement* in its French section rendered respective-
ly, if imprecisely, by eventual and eventually.

1.3 The apparent disregard shown by all but a few bilingual dictionaries for what must surely be considered a fundamental translation problem is a curious phenomenon of modern bilingual lexicography. Most editions rated above the level of the pocket dictionary recognize the necessity for discriminating between the different stylistic and semantic strata which dictate the term which is appropriate to a given context or situation. In the absence of appropriate sense indicators, the dictionary user would be prone to the same confusion which confronts him when, for example, he is led to believe that French trivial can be rendered by English trivial. Lexicographical procedure should thus be equally geared to making the user aware of the latter type of confusion.

2.0 Origins and Significance of Deceptive Cognates

Deceptive cognates are the result of the disparate development of a single word or word sense in two related languages. Hence, French volatil and English volatile are both inherited from Latin volatilis ("that which flies," "light," "ephemeral"), which in turn is derived from volare ("to fly"). However, although both words still share a common meaning in the chemical sense of
"evaporating rapidly," the English word has further developed the figurative senses "unstable," "capricious," "fickle," which may be variously rendered in French, but never by volatil. Conversely, the meaning "able to fly" and the nominal form "any flying creature" have disappeared from English while persisting in French.

Closely related languages like French and English are particularly abundant in word pairs of this kind. As Vinay and Darbelnet (1958) put it, "plus deux langues sont proches par la structure et la civilisation, et plus grand est le danger de confusion entre les valeurs de leurs lexiques respectifs, comme le montre, par exemple, la question des faux amis."

2.1 There is a wide range in the degree of resemblance and semantic proximity of these word pairs; at one end of the scale are those in which the identity is purely fortuitous and unrelated historically (which, moreover, should not be termed deceptive cognates but more accurately "faux amis"):  

(4) \( \text{jars nm gander} \)  
\( \text{jars npl pots, jarres.} \)  
\( \text{pour prép for.} \)  
\( \text{pour vtr verser.} \)  

At the other end of the scale are what might be termed
pure "faux amis"--cognates, etymologically speaking, but semantically independent in the modern idiom:

(5)  actuel, -le adj present.

actual adj réel.

trivial, -e adj common, trite;

vulgar.

trivial adj sans importance.

Within this latter group there is likewise a range of semantic distinction, from pairs like those in (5), where the sense relationship has been entirely broken, to those like (6) in which the confusion exists only at one semantic level in a multiple series of meanings:

(6)  idiom n language; dialect;

[langue] idiom.

idiome n idiotisme m; dialecte m;

[language] idiome m.

Examples of this last type are by far the most common; Vinay and Darbelnet label them "faux amis partiels" and they may be depicted schematically as partially overlapping semantic fields, as in Figure 10.

The source of confusion in this pair is two-fold: in the first place, English idiom possesses a sense lacked by French idiome--that of "idiotisme/expression of unpredictable meaning"; secondly, the frequency of the respective senses of these words is not parallel: the
Figure 10. Partially overlapping semantic fields.

sense "idiotisme" of English idiom is far more frequent than the generic senses it shares with idiome.

2.2 Vinay and Darbelnet, the only investigators to carry out a systematic analysis of the phenomenon of "fau amis," identify three "aspects" in their classification:

(a) **semantic** ("l'aspect sémantique"), where word pairs are distinguished by sense differences:

(8) **actuel, -le adj present.**

actual adj réel.

(b) **stylistic** ("l'aspect stylistique"), in which stylistic differences separate the pairs:

(9) **populace n foule f ("sens intellectuel").**

populace nf rabble ("sens affectif").
(c) **phrasal** ("l'aspect phraséblogique"), where the deception is related to the structure of the word or phrase; in this case, the individual elements of a compound, derivative or syntagm, are not "faux amis" but in combination they produce a global sense which cannot be paralleled in its opposite language counterpart:

(10) **cut-throat** n coupe-jarret m.

**coupe-gorge** nm death-trap; den of thieves.

**confidence man** n escroc m,

**chevalier** m d'industrie.

**homme de confiance** nm a man who inspires confidence.

2.3 This classification can be considerably enlarged upon (see below, 3.0), and any serious attempt to deal with the phenomenon in bilingual lexicography should, in fact, be based on a clear understanding of the kinds of deception likely to be encountered by the translator. Since translators use dictionaries habitually and specialized lists of "faux amis" rarely, the phenomenon should unquestionably be dealt with in the dictionary. As the editors of OXF suggest in their preface, it is far more appropriate for general dictionaries to signal the presence of "caltrops or **chausse-trappes,**" than for them to be relegated to the restricted domain of the
specialist dictionary. But the task must be carried out systematically and consistently, bearing in mind the needs of users in both languages. The basic problem is the identification of the word pairs which create confusion: they may be culled from existing specialized lists and educational texts, but even the most enterprising of these does not go beyond about 2,000 words whereas there are clearly far more than this. A truly adequate study would require a very methodical search, one which might best be accomplished by appealing to computational methods where a programme would be employed to generate lists of word pairs of identical or very similar appearance; having extracted manually from such a list those pairs which constitute a possible source of confusion, attention can be given to the manner in which they should be presented in a lexicographical framework. The basis for the latter would be a thorough classification of all possible varieties of "faux amis."
be discussed in section 3 below).

3.1 **Homographic ambiguity:** this category incorporates a large number of word pairs whose identity is the result of chance and is purely orthographical. Koessler and Derocquigny (1928) and Boillot (1930) included many of these in their lists, even though they are very often clearly identified by grammar, context or even orthography:

(11) **hurler** vintr to howl.
    **hurl** vtr lancer.
    **ail** nm garlic.
    **ail** vintr souffrir

(12) **bosse** nf hump.
    **boss** n patron m.
    **horse** nm pot, hash.
    **horse** n cheval m.

The examples in (11) are conclusively separated by their part of speech labels and should require no other discriminatory attachment to identify them to the normally careful dictionary user. Thus, in a sentence like:

(13) Les loups commencèrent à **hurler**.

the translation of **hurler** by to **hurl** is blocked because
the latter is a transitive verb for which an object is mandatory and which, moreover, requires an animate (and usually human) subject. On the other hand, the examples of (12) are not distinguished by grammar and confusion is possible; an English speaker, faced with the translation of:

(14) The file was blunt.

might be persuaded, in the absence of clear direction, to render English file by French classeur or dossier or file, rather than the contextually appropriate lime. This type of error is engendered and proliferated by the kind of abridged pocket dictionary which merely lists a series of possible equivalents for a word without regard to semantic or etymological distinction, as the following articles from the paperback edition of LAR1 illustrate:

(15) file s lime f; v limer.
    file s file f; classeur m;
    pl dossier m; v défile;
    classer; . . . . .

Lip service has been paid to etymology by separating two entry words for file (whereas in fact there are at least three: "classeur," "file" and "dossier"). But no attempt has been made to identify these three forms, and the confusion is compounded by the inadequately labelled verbal derivatives which make no distinction between the
transitive forms (limer, classer) and the intransitive form (défiler).

The ambiguity created by this type of "faux ami" can often be partly or completely removed by sound lexico-graphical techniques. The following articles from CD1 (16) and OXF (17) suggest two approaches to the problem of homographic forms:

(16) file¹ n 1. classeur m: spike ~, pique-notes minv; card index ~, fichier m. 2. dossier m.

file vtr classer, ranger.

file² n file f: in/single, Indian/ ~, en file indienne, à la queue leu leu; ~ -leader, chef m de file.

file³ n lime f: three-cornered ~, tiers-point m.

file vtr limer (le bois, le métal, etc.).

(17) file s (tool) lime f; (mil) file f; (long line) queue f; (papers) liasse f; (letters) classeur m in single ~, à la file (indienne).

A combination of these two systems (separation of semantically independent units as entries and adoption of subject labels), together with a symbol warning of the presence of "faux amis," would provide an even more ideal solution.
3.2 Semantic ambiguity: like the previous category, this encompasses most of the cases likely to be encountered; it may be subdivided into:

(a) those whose semantic fields never overlap:

(18) *grappe* nf bunch, cluster

    *grape* n (grain de) *raisin* m.

    *librairie* nf bookstore

    *library* n bibliothèque f.

They may be true deceptive cognates, related etymologically as in (18), or purely homographic forms, as in 3.1;

(b) those whose semantic fields overlap partially and in varying degrees:

(19) *pourpre* adj dark red, crimson, purple-red; [Fig] purple.

    *purple* adj violet; rouge violacé; [Fig] pourpre

    *comédien* nm actor; comedian

    *comedian* n acteur m comique;

As the majority of "faux amis" fall into this latter category, the ambiguity which they engender is only partial and therefore, from a lexicographical point of view, frequency of usage may dictate the manner in which
they are discriminated in the article. The word *sympathique*, for example, is not labelled as a "faux ami" in OXF:

(20). *sympathique* adj sympathetic; congenial, likeable.

This is essentially the structure and order employed by HNS. The word is likewise not signalled as deceptive by CDI:


However, the outdated, and therefore incorrect rendering of *sympathique* by *sympathetic* is avoided, and the rubrics further specify the accurate equivalents. The addition of the warning sign for a "faux ami" would make this article completely unequivocal. 8

3.3 **Structural ambiguity**: this term is employed by Vinay and Darbelnet (1958) to describe compounds, derivatives, phrases and syntagms, whose individual elements are unambiguous but in combination become "faux amis"; they fall into various sub-categories: 9

3.3.1 **Lexical structure**, which involves the combination of two or more lexical elements to form a virtually single element:
(a) composition:

(22) distaste n répugnance f
dégoût nm disgust

afterthought n réflexion f, après-coup m

arrière-pensée nf ulterior motive, mental reservation

(23) leap-frog n saute-mouton m

(24) figurehead n [Naut] buste m; figure f de proue; [Fig] homme m de paille.

Examples (22) and (23) illustrate the existence of literal translations whose total meaning is not equivalent; in (23) one element (leap/saute) retains the equivalent sense in the compound while the other does not; and in (24) the French equivalent employs entirely different elements which are also structured differently and, in the case of homme de paille, in a "locution figée."

(b) concatenation: in this case the elements of a word group retain their independence; individually the elements have unambiguous equivalents, but in combination literal translation is impossible:

(25) fresh water eau f douce
eau fraîche cool water
une voix rude  a gruff voice
a rude voice  un ton impoli

a wild man  un homme farouche
un homme sauvage  an introvert,
unsociable man

The difficulty in expressions like fresh water/eau douce can be traced to a discrepancy in the degree of semantic extension of the individual elements. The components water and eau have virtually identical extension, as OXF informs us:

(26) water s (gen., fig., tears, urine, of gems, expanse of water, of scents, etc.) eau f.

But this is far from the case with the elements fresh and frais, the English word possessing a much wider range of meaning:

(27) fresh water (a) (not salt) eau douce (b) (just drawn, cool) eau fraîche

fresh air (le) plein air, (le) grand air; air/pur,frais/; le frais

fresh/fruit, meat, butter/ /fruit, viande, beurre/frais, fraîche.

fresh (= green) vegetables légumes verts

fresh (= original) idea idée originale

fresh face (= new) nouveau visage

fresh horses (= relief) chevaux de relais
fresh news  nouvelles fraîches
fresh complexion 'teint/fleuri, frais/
fresh breeze  jolie brise
to get a fresh supply (of smthg)
se remonter (de qqc)
fresh supply (of troops, etc.)
reform (de troupes, etc.).

3.3.2 Syntagmatic structure, includes phrases made up of lexical and syntactic elements unable to be translated word for word. They are further subdivided into:

(a) phrasal structure: where the elements may be combined in a similar manner in the opposite language but with a change in meaning:

(28) habiter au premier  to live on the ground floor;  to live on the second floor.

avoir des relations  to have influential friends
to have relations  s'accointer (avec);  entretenir des relations (avec).

passer un examen  to/write, take/an exam

pass an exam  réussir, être reçu (dans un examen)

une partie de tennis  a tennis match

a tennis party  une/réunion, soirée, etc./de tennis
(b) **idiomatic structure:** where the combination of elements is peculiar to one of the languages and not the other:

(28) to take the chair *(at a meeting)*, présider *(une réunion)*

to burn the midnight oil, pâlir sous la lampadaire

scier le dos à qqn, to bore s.o. stiff.

(c) **proverbial structure:** like the preceding sub-category, the combination of elements is characteristic of the language and usually nonsensical in the other:

(30) A bon chat bon rat/Tit for tat

*You can't have your cake and eat it* /

*Il faut qu'une porte soit ouverte ou fermée.*

(d) **grammatical structure:** where the syntagmatic arrangement of elements produces different meanings in the two languages:

(31) *il nous manque* we miss him

he misses us nous lui manquons

*to take place* avoir lieu

*prendre place* to sit, stand, find a place
en vue de . . .  in order to
in view of . . .  étant donné

3.3.3 Paradigmatic structure: many morphological elements are parallelled in French and English; but sometimes the parallelism is interrupted, giving rise to series of "faux amis." Thus, the language student who has learned to recognize such "vrais amis du traducteur" as the pairs -ly/-ment, -ish/-âtre, dis-/dé-, -ness/-tude, -ity/-ité, and so on, should be equally alert to the possibility of deception as illustrated in the following examples:

(32) folâtre adj playful, full of fun
foolish adj bête, sot
display vtr exposer, étaler
déjouer vtr to foil, baffle, thwart
casualité nf casualness
casualty n accident m, désastre m.

In some instances the confusion may be produced by the lack of an equivalent morphological form:

(33) amical adj friendly
amicalement adv in a friendly manner
The English suffix -ly is almost invariably associated with adverbs and its appearance in adjectival form creates difficulties for the French-speaker.

3.4 Phonetic ambiguity: a lengthy study has been carried out by Las Vergnas (1947) on the "faux amis" of pronunciation; his investigation was intended as an aid to French speakers learning English, but his findings can be adapted to bilingual lexicography with some profit. In this category are the word pairs whose identical spelling is not accompanied by identical pronunciation:

<table>
<thead>
<tr>
<th>English</th>
<th>French</th>
</tr>
</thead>
<tbody>
<tr>
<td>apiculture</td>
<td>apikylty:r</td>
</tr>
<tr>
<td>item</td>
<td>item</td>
</tr>
<tr>
<td>cycle</td>
<td>sikl</td>
</tr>
<tr>
<td>archives</td>
<td>ariv</td>
</tr>
<tr>
<td>mètre/mètre</td>
<td>[méttr]</td>
</tr>
</tbody>
</table>

Homographs in the same language are also sometimes accompanied by non-identical pronunciation, usually, but not invariably, related to a change in grammatical category:

(35) rebel [ˈrebəl] n rebelle m
     rebel [ˈrebəl] vintr se révolter
     desert1 [ˈdɛzərt] n désert m
Certain repetitive elements, such as the enigmatic -ough series in English, may also be considered as "faux amis" under this heading:

(36) **bough** [bau] n rameau m
    **though** [ðou] conj quoique
    **thorough** [ˈθɔri] adj parfait
    **through** [ˈθruː] prep par
    **thought** [ˈθɔt] n pensée f
    **enough** [ɪˈnʌf] adj assez de
    **cough** [kɔf] n toux f

In some cases the deception applies to sense as well as to sound:

(37) **dais** [ˈdeɪs] nm canopy
    **dais** [ˈdeɪs] n estrade f

3.5 **Dialectal ambiguity:** some bilingual lexicographers have become increasingly aware of the necessity to record major dialect variations, both as entry words and as alternative translations. CASS2 incorporates a number of Canadian French terms supplied by Gaston Dulon and
acceptable to the Office de la Langue Française au Québec, although these are only to be found in the French-English section. HNS now recognizes the existence of a sizeable offshoot in North America of standard British English and labels (somewhat inaccurately) some of its translations as \( \text{US} \). One dictionary, CD2, uses four dialect labels: \( \text{Fr} \) (for European or "standard" French), \( \text{C} \) (for Canadian French or Canadian English), \( \text{Br} \) (for British English) and \( \text{US} \) (for American, as opposed to Canadian English). Far from being mere frills inserted at the whim of lexicographers, these "flags" are essential attributes of accurate translations and indispensable to the translator who deals with dialectal forms. There are some curious instances of "faux amis" between dialects which bilingual lexicography must take into account:

\[(38)\] IF chaussette (sock) = CF chausson

IF chausson (slipper) = CF chausette

BE inflammable (inflammable) = CE, AE flammable

BE non-inflammable (ininflammable) = CE, AE inflammable

CE, AE public school (école primaire) = BE state school

CE, AE private school (école privée) = BE public school
Examples like those in (38) are numerous and constitute a source of dialectal "faux amis" which must be identified and disambiguated in the dictionary. Thus, for example, the user must be directed to translate malle arrière by boot in a British English context and by trunk in a North American context; if the translator is confronted by the phrase habiter au premier, he must be informed that in a European French context this would mean to live on the second floor, while in a Canadian French context it would mean to live on the ground floor.

3.6 Register ambiguity: bilingual lexicography, as we observed in Chapters 11 and 13, has an obligation to provide translations which are equivalent in every sense, but one feature which is frequently neglected is the question of language level. Even larger editions can be shown to be less than adequate in this respect, as the following examples from HAP3 (39), HNS (40), CD1 (41) and OXF (42) indicate (the words in parentheses are more appropriate translations):

(39) baguenauder vintr to fool around, mooch about, loaf; to fiddle-faddle. ([Fam] fool around) ([Vulg] screw)
(vieux) birbe nm old dotard. ([Slang] old fuddy-duddy) ([Fam] old fart)

(40) cuisse nf : elle a la - /léger, hospitalière, facile/, she's a woman of easy morals. ([Fam] she screws around)

(41) violon nm . . . [Fam] prison cells. (jug, clink, lockup)


maint, -e adj many. (multitudinous)

(42) haw-haw vintr [Fam] rire bruyamment. (s'esclaffer)

nay adv (strong affirm) bien plus. (voire)

These samples exemplify a common error in bilingual lexicography, the failure to indicate the language level of the word in question and (even where the label is present) the inability to provide a translation of the same register. What dictionaries are doing, in effect, is introducing and perpetuating "faux amis" which should never have caused ambiguity in the first place.

3.7 Gender ambiguity: for the English-speaking learner of French (and other languages employing gender distinctions) the existence of pairs of words differentiated in meaning by gender alone constitutes a further source of "faux amis." Unaccustomed to gender discrimina-
tion, the English-speaker is apt to be insensitive to the
difference between such pairs as:

(43) moule nm mould, form
    moule nf [mollusc] mussel
    geste nm gesture
    geste nf deed, exploit

while epicenes like those in (44) are liable to cause him
confusion as to the sex of the individual concerned:

(44) personne nf person (male or female)
    sentinelle nf sentry (usually male).

3.8 Frequency ambiguity: a final category of "faux amis" has already been alluded to above (3.2); it
concerns the frequency of usage of the various senses of
a word and the inability of dictionaries to account for
meanings in current usage. Example (20), repeated here
for convenience, illustrated the word sympathique with
meanings which were not in accordance with attested
frequency:

(20) sympathique adj sympathetic;
    congenial, likeable

The same unspecified ordering is employed by LAR2:
sympathique adj sympathetic; attractive, pleasing, appealing.

HNS makes no attempt to identify the deception other than a more precise sense categorization:

1. sympathetic
2. likeable, attractive
3. encre, invisible ink
4. [Anat, Physiol] sympathetic (nerve, etc.)

Kirk-Greene (1968) lists sympathique among his "mots-amis" as follows:

1. sympathetic
2. nice, kind, likeable; que notre professeur est...: how nice our master is!
3. encre sympathique, invisible ink.
4. [Anat, Physiol] sympathetic (nerve, etc.)

But the word also appears in his "faux-amis" list, implying that he recognizes it as such (it merely contains a cross reference to the "mots-amis" list).

Although not indicating the presence of a "faux ami," CD1 (21) avoids the incorrect English cognate and further specifies the meanings by the use of rubrics:


The incorrectness of the sense orderings given in (20), (46) and (47) is borne out by reference to a monolingual dictionary. PROB indicates that the meaning "sympathetic" is archaic.
Almost fifty years earlier, Koessler and Derocquigny (1928) made manifestly plain to a French reading public the true meaning of sympathetic:

Sympathetic. -- A sympathetic man n'est pas un homme sympathique dans le sens d'un homme qui se concilie la sympathie, mais c'est un homme qui éprouve de la sympathie et même de la compassion, de la commisération.

Noter, d'autre part, que a sympathetic strike est une grève de solidarité et que a sympathetic landscape est un paysage qui émeut par les souvenirs qu'il évoque. . . .

Sympathetic a pour synonyme responsive (voir ce mot).12

The current meaning of sympathique is, according to PROB (II., 2) "qui inspire la sympathie"; to ensure what is meant by "sympathie," this word must also be consulted and it will be determined, from the same dictionary, that its modern and current meanings are "relations entre personnes qui, ayant des affinités, se conviennent, se plaisent spontanément; sentiment chaleureux et spontané,
qu'une personne éprouve (pour une autre).” By contrast, the first meaning given by WIII for sympathetic is "of, expressing, resulting from, feeling or showing sympathy"; this is followed by "in agreement with one's tastes, mood, feelings, disposition, etc.; congenial"; then "showing approval, favour or agreement"; and finally the special technical sense of Anat., Physiol. and Physics. The ordering of the non-technical senses is therefore the reverse of its French counterpart and this disparate frequency of usage must be reflected in the bilingual dictionary article.

4.0 Lexicographical Presentation

The preceding classification, which is not necessarily exhaustive, underscores the importance of the phenomenon of "deceptive cognates in translation. As far as their lexicographic presentation is concerned, a model is provided by the only two dictionaries to deal with the problem, OXF and CD1. In Section 1 it was suggested that a more thorough and systematic approach is required than that of either of these dictionaries; the chief question to be resolved is the matter of how far the lexicographer must go in commenting on "faux amis," for he could end up, in the extreme case, with commentaries on almost every word in the dictionary. In order to decide this question, consideration must be given to the kind of user
most likely to consult the dictionary and to his specific needs. If it could be determined that only competent linguists and fluently bilingual translators would consult the dictionary, it might well be possible to dispense altogether with warning signals and explanatory remarks. On the other hand, a dictionary placed largely in the hands of naive users—students and language learners in general—might be expected to provide considerable support and guidance; what may be a redundancy for experts is usually indispensable for learners, and since the majority of users fall into the latter category it seems clear that bilingual dictionaries should identify and explain deceptions systematically and fully.

4.1 Homographic ambiguity covers most cases of "faux amis," but every example will not need special treatment and each must be considered separately. Many of the pairs furnished by Koessler and Derocquigny (1928), which are clearly disambiguated by grammar and even spelling (see examples in (11) above), will not require any further identification. Even where the identity extends to both orthography and grammatical category (see examples in (12)), it may be possible to eradicate the ambiguity by sound lexicographical techniques (see examples (16) and (17) above). However, if the slightest
hint of ambiguity remains in the finished article, it should be eliminated by employing appropriate warning symbols accompanied by explanation or cross reference. The technique can be illustrated by making adjustments to (16) and (17) as follows:

(49) **file**<sup>1</sup> **n** 1. [X] [Stationery] classeur **m**: spike ~, piques-notes minv; card index ~, fichier **m**.

2. [Admin] dossier **m**: [≠ FILE, q.v.].

**file** **vtr** [X] [Admin] classer, ranger; [≠ FILER, q.v.].

**file**<sup>2</sup> **n** file **f**: in/single, Indian/-, en file indienne, à la queue leu leu; ---leader, chef **m** de file.

**file** **vintr** [X] marcher à la file, défiler; [≠ FILER, q.v.].

**file**<sup>3</sup> **n** [X] [Mech] lime **f**: three-cornered ~, tiers-point **m**: [≠ FILE, q.v.].

**file** **vtr** [X] [Mech] limer (le bois, le métal, etc.); [≠ FILER, q.v.].

It can be seen that the "faux ami" label [X] has been attached to every entry in the homographic series with the exception of **file**<sup>2</sup> for which the correct translation is **file**. Likewise the English verb **file** (transitive or intransitive) is not to be confused with French **filer** whose meanings are "spin," "pay out," "follow," and so on. The entire series is further disambiguated by the use of the subject labels.
[Stationery], [Admin(istration)] and [Mech(anics)], leaving the reasonably careful reader in little doubt as to which word to use in the context appropriate to his purpose. In order to complete the picture, the French words file and filer must be dealt with in a similar manner:

(50) file nf file, line; [Mil] file, rank: à la ~, in a row; se mettre à la ~, to line up; chef m de ~, leader; en ~ indienne, in Indian file; [cf. FILE¹, FILE³].

filer vtr [X] 1. [Cout, Ind, etc.] to spin; [Naut] to pay out (cable, etc.); [Fig] ~ une scène, to spin out a scene; ~ doux, to go easy.
2. [Police, Crim, etc.] to shadow, tail (s.o.); [≠ TO FILE, q.v.].

filer vintr [X] 1. to flow thickly; [lampe] to smoke. 2. [Fig] to fly, speed along: ~ à l'anglaise, to take French leave; ~ entre les mains (de qqn), to slip through s.o.'s fingers; [≠ TO FILE, q.v.].

In this homographic series (likewise adapted from CD1), only the verbal forms are labelled as "faux amis," there being no confusion in the noun forms going from French to English. To ensure that no user should be tempted to equate French file with the other English homographs, a cross reference is provided to file¹ and file³. In both series, the "faux ami" symbol is further reinforced by a note at the end of the article indicating the word to be avoided and suggesting that it
be consulted; this eliminates the necessity to indulge in lengthy essay-type explanations, as employed by OXF, which are very space-consuming.

4.2 Examples (49) and (50) are the models which should be followed wherever ambiguity is introduced. Every "faux ami" of sense or structure should be treated according to this pattern, whether the deception is full or partial:

(51) pourpre adj [X] dark red, crimson; purple-red; [Fig] purple; [≠ PURPLE, q.v.].

purple adj [X] violet, rouge, violacé; [Fig] pourpre; [≠ POURPRE, q.v.]

afterthought n [X] réflexion f, après-coup m; [≠ ARRIÈRE-PENSEÉ, q.v.].

arrière-pensée nf [X] ulterior motive, mental reservation; [≠ AFTERTHOUGHT, q.v.].

passer vtr 1. to get/through, across/. 2. to pass (across), hand (over). 3. to get/over, across/. 4. to pass, spend (time). 5. [X] [Educ] to take, write (an exam): il m'a fait le test, he made me take the test; [≠ TO PASS, q.v.].

pass vtr passer (le pain, un ballon, etc.); passer/devant, près de/(un endroit, etc.); croiser (qqn allant en sens contraire); [Road] dépasser (une voiture); [X] [Educ] être reçu, réussir a (un examen); [≠ PASSER UN EXAMEN, q.v.].

4.3 Grammatical "faux amis" will require varied treat-
ment according to circumstances. In the case of semantic ambiguity, the procedure will be as in the model:

(52) place n . . . . . . to take ~, [X] avoir lieu; [≠ PRENDRE PLACE, q.v.].

place nf . . . . . . prendre ~, [X] to sit, stand, find/a place; [≠ TAKE PLACE, q.v.].

Where semantic equivalence is unbroken but structure is the source of confusion, a cross reference replaces the inequality symbol:

(53) manquer vintr . . . . . . vous nous manquez, [X] we miss you; cf. TO MISS.

miss vtr . . . . . . we miss you, [X] vous nous manquez; cf. MANQUER.

4.4 Paradigmatic structure will only require attention where both morphological elements of the words could be misinterpreted:

(54) folâtre adj [X] playful, full of fun; [≠ FOOLISH, q.v.].

foolish adj [X] bête, sot, etc.; [≠ FOLATRE, q.v.].

In other cases, the lack of equivalence should be obvious; the words polish (vtr), fetish (n), for example, cannot be broken down into the components pol- + -ish, fet- + -ish, for which French equivalents in -âtret can be found.
4.5 Dialectal deceptions need careful handling because they can be very complicated:

(55) *chausette* nf sock; [X] C
slipper; [≠ CHAUSSON, q.v.].

*chausson* nm slipper; [X] C
sock; [≠ CHAUSSETTE, q.v.].

sock n *chausette* f; [X]
*chausson* m (C) [≠ SLIPPER, q.v.]

slipper n *chausson* m; [X]
*chausette* f (C) [≠ SOCK, q.v.].

(56) *flammable* adj (C US) [X]
flammable; Br inflammable cf.
ININFLAMMABLE.

*inflammable* adj (C US) [X]
inflammable, non-flammable, Br
non-inflammable; cf. INFLAMMABLE.

flammable adj (C US) [X]
inflammable: non-, inflammable;
[cf. INFLAMMABLE].

inflammable adj (C US) [X]
inflammable; [≠ French INFLAMMABLE,
q.v.]; Br inflammable.

4.6 Register ambiguities may normally be removed by the correct use of labels, with the assumption that consistency will be maintained in the level of the translation equivalent:

(57) *hock* vtr [Fam] mettre au clou.

*maint. -* e adj [Litt] multitudinous,
legion (in predicate use).
If, however, the entry word is prone to semantic or other
deception, the warning symbol might be introduced as
follows:

(58) fraise nf [Bot] strawberry; [Fig, Fam] [X] mug, kisser,
chops; [≠ STRAWBERRY, q.v.].

violon nm [Mus] violin; [Fig, Fam] [X] lockup, clink, (the) jug,
cooler; [≠ VIOLIN, q.v.].

4.7 Gender ambiguities should be recorded in the English
lexicon in the following manner:

(59) mould n [shape] moule m [X]; [≠ MOULE f, q.v.].

mussel n [mollusc] moule f [X];
[≠ MOULE m, q.v.].

This procedure draws the user's attention to the
possibility of error in gender and advises verification
of the other word.

4.8 Idiomatic and proverbial structures may, as we have
observed, engender "faux amis" if a word-for-word
equivalent is possible but with a different meaning;
these should be recorded as such in the appropriate
section of the dictionary:

(60) to take the chair (at a meeting), [X] présider (une réunion).

sucrer les fraises, [X] to have
the shakes.
But when the expression is so characteristic of the language that literal translation is non-sensical, the rubric alone ([Loc], [Prov]) should suffice to prevent error:

(61) \textipa{pipe nɪː} \ldots [Loc, Fam] \textipa{caʊsər sɑː}, to kick the bucket.

5.0 Although many lexicographers might object to the introduction of further typographical complexities in an already cluttered arrangement of dictionary information, the fact remains that the significant element of language interference represented by deceptive cognates must be recognized. The methodology proposed here complements the position taken in earlier chapters where the requirement of total disambiguation of differential structure was given prominence. The equivalence relation, as we have observed, is a complex and sometimes elusive ideal based on the abstract nature of lexicographical material. Thus, even if the result is considered somewhat cumbersome, it is essential for bilingual dictionary makers to account for the phenomena of deceptive cognates by drawing the reader's attention to them in just the same way that semantic ambiguities are resolved for him by the careful organization of labelled senses.
Notes to Chapter 14

1 Koessler and Derocquigny (1928), p. xiii.

2 Unlike the other works cited here, which are chiefly concerned with establishing lists of "faux amis," this study in comparative stylistics includes a chapter which classifies and analyses the phenomenon as a translation difficulty.

3 Considerations of space have compelled us to a mere signalization with a brief remark, a short commentary. We do not claim to have exhausted the subject. We have only paved the way, because we believe that no modern dictionary is complete and really useful if it does not warn readers and translators of the dangers and pitfalls offered by apparently too easy translations. Preface to OXF, p. x.

4 Vinay and Darbelnet (1958), p. 43.

5 Ibid., p. 71.

6 Ibid., pp. 71-73.

7 OXF, preface, p. x.

8 The same dictionary inserts the warning sign for the derivative verb sympathiser.

9 Referred to as "faux amis de structure," p. 73.

10 As illustrated by a Francophone correspondent who persists in terminating his letters with the salutation: "Friendly Didier," undoubtedly contaminated by his own "Amicalement." His state of confusion would not be reduced if he were to see the bumper sticker which invokes other road-users to "DRIVE FRIENDLY."

11 See "Préface sur les canadianismes" in CASS2; also Canadianismes de bon aloi, published by the Office de la Langue Française au Québec.

12 Koessler and Derocquigny (1928), pp. 363-4.

13 The etymological relationship between Old French filer ("to thread documents together") and English to file (documents, etc.) has been lost in the modern French word.
References for Chapter 14


1.0 Regional Variation in Bilingual Lexicography

The inclusion of regional forms in bilingual dictionaries is something of a rarity, for this kind of lexicographical undertaking normally purports to bring together pairs of national languages. Thus, for example, a dictionary containing French as the source language will represent what is commonly referred to as "international French"; if the target language is English the variety chiefly employed until comparatively recently is British English. In fact, before 1962, the year of publication of J-P. Vinay's *The Canadian Dictionary/ Dictionnaire canadien* (CD1), the field of bilingual lexicography for French and English was almost exclusively dominated by *Harrap's Standard French and English Dictionary* (HAP3) and by other European-published dictionaries. Canadians, and for that matter North Americans in general, were therefore obliged to use dictionaries produced in Europe, by Europeans and for Europeans. HAP3 extended what was referred to as a "wide hospitality" to
Americanisms, but no attempt was made to provide a
definition of the term and no reference whatever was made
to the question of Canadianisms in either language. Its
chief rival in Europe was Cassell's French-English English-
French Dictionary (CASS2) which, prior to the 1962 edition,
contained a small number of Americanisms but no
Canadianisms. The 1962 edition benefited from the
collaboration of Gaston Dulong who contributed some 200
Canadian French (CF) words, mainly terms which were at
that time the only ones acceptable to the Office de la
langue française au Québec (OLFQ), of which Dulong was a
consultant.

The most recent Harraps edition--Harrap's New Standard
French and English Dictionary (HNS)--exists so far only in a
two-volume French-to-English version which still gives
scant attention to CF and offers no explanation of the
source of canadianismes or of the terms of reference by
which they are defined. Canadian English (CE) continues
to be subsumed under the heading of "Americanisms" which
seem to have been considerably augmented; the question of
whether they are included as an aid to European French
users or to British users remains unanswered.

1.1 The first bilingual dictionary produced in Canada
for Canadians was CD1. But anyone who expected a totally
Canadian dictionary, depicting CF usage on an equal
footing with CE, would have been disappointed. Recognizing that members of both official language groups deserved a dictionary reflecting current linguistic usage in Canada, the editors established "international French" and "international English" as the structure on which to build a distinctly Canadian translation tool. They are not open to censure in this regard, because at the time the dictionary was being prepared the two languages did not enjoy equal freedom of movement. Although English lexicographers can select items and create dictionary entries for them with relative impunity, their French counterparts are answerable to a number of officially recognized linguistic regulatory bodies (such as the OLFQ, Observatoire de la Langue Française, Académie Française). The climate of linguistic opinion in Quebec regarding French canadianismes at that time was such that only a very limited number of words (about 100) enjoyed the sanction of the OLFQ and could be freely used in educational texts and dictionaries. In the face of severe regulation and the economic sanctions which non-conformity might entail, the presence of some 700 canadianismes in CD1 is surprising and a tribute to the foresight of its editors who perhaps predicted the change in attitudes towards CF which have emerged in the seventies.
1.2 But those 700 words represent a very small proportion of the 20,000 French entries in what is billed as a "Canadian" dictionary. In every other respect, the dictionary is manifestly another international French lexicon; the pronunciation and orthography are international; the examples and proverbs are almost exclusively European-oriented; even the reality depicted is more European than Canadian—to judge from the fact that Canadian examples usually follow the European, almost as though they were a tentatively appended afterthought. In contrast, the English side of the dictionary is almost totally Canadian—in pronunciation (although based on a limited area), orthography and content. Where there are marked differences between Canadian and British or between Canadian and American usage, these are quite carefully indicated. Items which are considered to be distinctly Canadian and not in general use elsewhere are further distinguished from the "common core" by the symbol indicating a Canadianism.

2. The Problem of Definition

If a dictionary purports to represent an international version of a language, then references to regional variations should be prefaced by a careful description of the editors' conception of the term "regionalism". In Canada the situation is comparatively
straightforward since the degree of regional variation in both languages is quite restricted by comparison with that of English in England (or the United States) and French in France. But regionalisms do exist in Canada and, as M. H. Scargill and H. J. Warkentyne (1973) demonstrated from the results of their nationwide survey, they exist in sufficient quantities and variety to warrant inclusion in lexicographical projects.

2.1 The term used to designate forms which are characteristically Canadian in origin and usage is Canadianism/canadianisme. There is not, however, a universally accepted definition of the term as a label for regionalisms in French and English. No general consensus has been reached by linguists or lexicographers, all of whom define the concept in their own terms (and interpret it according to what may well be the dictates of commercial exigencies); even the OLFQ has proved inconsistent in its application of the restrictive definition promulgated in the sixties.

For French, at least, the term canadienisme has been the object of vigorous debate since the early 19th century. The first mention of the notion is by Jacques Viger (1810) whose short manuscript refers to "néologie canadienne ou ( . . . ) mots créés en Canada et maintenant en vogue ( . . . )." But it was not long before the term
canadianisme came to be used as a catchall by outraged
purists polemicizing against the intrusion of
unacceptable forms in current language. J-P. Vinay
(1973b: 364) identifies them as "barbarismes, archaïsmes,
anglicismes, joulalismes, des faits de langue utilisés par
les Canadiens français pour traduire la réalité
canadienne ou nord-américaine, néologismes, et
amérindianismes." There seemed to be as many interpre-
tations as there were investigators. At one end of the
scale came the edict of the OLFQ, which in 1969
recognized only 62 "canadianismes de bon aloi"; at the
other was the penchant of some linguists to find a
canadianisme in every word which differed, however
slightly, from standard French.

2.2 In order to arrive at a satisfactory definition we
will now analyse the descriptions attempted over the past
century.

2.2.1 A. Buies (1888) in his little book Anglicismes et
canadianismes cries out against anything which is not
French:

. . . anglicismes et (. . .) canadianismes
dont notre langue et notre style
fourmillent au point d'en perdre
entièrement toute physionomie française
(p. 3).

Le spectacle des expressions, des phrases,
des paragraphes entiers, non seulement
anti-français et barbares, mais absolument, incompréhensibles et indéfinissables . . . (p. 4).

La campagne entreprise par moi dans un dernier chronique contre les barbarismes, les anglicismes et tous les autres -ismes dont le débordement, dans notre langue et dans notre style, menace de faire disparaître jusqu'au dernier vestige de français parmi nous, ( . . . ) est une reprise d'hostilités ( . . . ) contre le plus dangereux ennemi que nous avons (p. 16).

Buies was not a linguist and it is quite evident, from these citations, that his purist attitude prevented him from making an impartial judgement. For him, canadianismes constituted just one of the many "foreign" elements which were tainting the language he obviously adored and would defend with his life.

2.2.2 Somewhat later Adjutor Rivard (1914), in his Etudes sur les parlers français au Canada, summarized his interpretation of the term without actually giving a specific definition:

Le langage des Canadiens français n'est pas encore connu tellement qu'on puisse, l'ayant analysé, dire avec précision quels éléments l'ont formé, dans quelle exacte proportion chacun d'eux y a contribué; mais on aperçoit bien que ce n'est ni le français classique, ni un patois pur, ni un français corrompu, et que cependant il accuse des particularités assez saillantes et assez d'uniformité, sur toute l'étendue du territoire, pour constituer un parler régional ( . . . ) car on abuserait peut-être du langage en l'appelant un dialecte (pp. 39-40).
Rivard thus appears more definite about what Canadian French is not than about what it is. He considers it to be a relatively uniform regional form, without being homogeneous, characterized by an archaic flavour resulting from its isolated state and by significant numbers of borrowings from the languages with which it found itself in contact.

Dans son ensemble, le parler du peuple canadien n'est donc pas à proprement dire un patois; mais il est le résultat de la fusion de plusieurs patois différents . . . (p. 70).

C'est le vocabulaire français, mais pauvre et imprécis, à peu près pur du patois, mais assez fortement archaïque, et mêlé, hélas! d'anglicismes barbares (p. 73).

We can thus infer from Rivard's description that, for him, *canadianismes* are any elements which could not properly be labelled as "français de bon aloi": *anglicismes, archaïsmes, barbarismes* and *patois*.

2.2.3 Since Rivard was a principal editor of the *Glossaire du parler français au Canada* (GPFC), his conception of the term should be carefully considered by anyone using this dictionary and assuming that its contents are exclusively *canadianismes*. The preface to the GPFC (1931) makes the situation clear by specifying that the elements described by Rivard are those which contribute
to the existence of distinctively Canadian forms. It is notable that the editors restrict their observations to the province of Quebec and that, in the absence of input from Acadian, Ontarian and other regional varieties, their product should be described by the term "québécoisismes":

Il convient peut-être, ici, de remarquer que c'est, en effet, à ce territoire seulement que notre enquête se rapporte, de sorte que les mots "au Canada" (....) pourraient se lire: "au Bas-Canada." Nous n'avons relevé aucune des particularités qui peuvent s'étre introduites dans le langage des groupes situés à l'ouest de notre province, non plus que les formes caractéristiques du parler des Acadiens .... Nous considérons que le parler du groupe plus considérable des Canadiens français, habitants de la province du Québec, peut être dit proprement le français du Canada (pp. vi-vii).

Without actually giving a name to the concept, the GPFC editors describe its content thus:

Notre parler, on l'a souvent démontré, est un français régional, où se rencontrent, comme dans toute langue populaire et surtout dans toute langue transplantée, des archaïsmes, des formes dialectales et des néologismes ....; nous n'avons donc enregistré que les formes particulières qui, n'appartenant pas à la langue académique d'aujourd'hui, donnent au parler populaire et familier de chez nous son cachet particulier .... C'est la langue parlée (....) que nous étudions dans ce Glossaire; et nous y relevons seulement ce qui n'a pas été ou n'est plus admis dans la langue académique .... C'est la langue de l'Ile-de-France,
telle qu'elle était déjà répandue dans les provinces du centre, du nord et de l'ouest, lors des grandes émigrations en Amérique; nos pères, venus de ces régions, nous l'ont léguée, et nous l'avons conservée, avec certaines particularités provinciales, qui légites ou irrégulières, n'attestent pas moins que les archaïsmes la survie chez nous de la génie de l'idiome (pp. vii-viii).

A single phrase in the entire preface alludes to what might be considered truly "canadien":

... quelques termes pittoresques, qui ont de la naissance, et qui conviennent à l'expression de la vie canadienne (p. viii).

2.2.4 V. Barbeau (1963) used his Le français du Canada to plead for the acceptance of CF as a bona fide regional dialect; at the same time, he sought a middle road and gave a specific definition of the term canadienisme:

Le terme de canadienisme recouvre le meilleur et le pire. Il y a donc un partage entre l'ivraie et le froment... Le préjugé qu'on entretient généralement à l'endroit du parler populaire conduit souvent à une sévérité abusive. On oublie que bien des fautes d'hier sont devenues depuis des règles. Mais indépendamment de ces avatars dont plusieurs sont encore tout récents (j'è pense à réaliser et à ignorer condamnés hier et devant lesquels s'inclinent maintenant les bons auteurs), l'on s'expose, comme ce fut le cas au 17e siècle, à se priver d'un limon nécessaire à la croissance de la langue (pp. 175-6).
Que seraient nos hivers sans les bancs de neige, les bordées, ... le fraisil, ... la poudrière, ...? Que serait la forêt sans ses chantiers, ses érablières, ses cabanes à sucre, ...?

Il s'ensuit que nous nous méprendrions gravement sur la légitimité et la nécessité des canadianismes en les condamnant tous d'un bloc. Je concède qu'il y a beaucoup de lest à jeter. Mais outre qu'il n'y a pas, en cette matière, de dirigisme possible, que deviendrait le français privé de l'appoint de la création populaire? Une langue déséchée, ... ?

On comprendra, sans plus d'explications, que citer un canadianisme n'équivaut pas à le consacrer. Le passage de la langue vulgaire à la langue écrite échappe à toute règle. ... (pp. 176-7).

With Barbeau canadianismes are endowed, for the first time, with a measure of respectability; the whole language should not be relegated to the level of barbarisms and archaisms which are to be avoided because there is much that is good—picturesque, unique and highly descriptive of Canadian reality. This attitude, together with his list of 700 words and expressions, probably paved the way for the bold approach adopted by L-A. Belisle, the first major lexicographer in French Canada.

2.2.5 The 1959 publication of Louis-A. Belisle's *Dictionnaire général de la française au Canada* contains a fairly carefully worded description of the Canadian content of
the work; as much as anything this account was designed to stay the predicted criticisms of the promulgation of large numbers of "mots populaires." He was dubious about the practical value of the GPFC, in spite of the high quality of scholarship which went into its compilation, but he was anxious to establish a more precise framework for the definition of what was properly Canadian French as opposed to literary French:

Les mots essentiellement canadiens comme "maganer," bâdrer," "slaque" et "lousse" n'offraient à ce point de vue aucune difficulté. Là où la perplexité du rédacteur en alerte devenait grave, c'était devant les mots qui ont au Canada une acception autre qu'en France—comme, par exemple, "sucreerie," "sableusé," "four" et "fourneau," ( . . . ) etc. Où commence les canadianismes et où s'arrête le français littéraire? Rien, à cette époque, ne pouvait établir la ligne de démarcation si ce n'est pas un long travail de consultation et de comparaison des définitions européens et des définitions du Glossaire. Et puis les mots ne sont pas toute. Il faut tenir compte des tournures. Et je constatais chaque jour que quelques-unes de nos tournures populaires les plus savoureuses ne paraissaient guère usitées en France. La solution qui me semble la plus logique était la compilation d'un dictionnaire où l'on pouvait trouver indifféremment tout cela (pp. 15-16).

Belisle has here identified a very important characteristic of CF, the fact that many words shared commonly by both language groups have acquired, in Canada, a distinctive sense unknown in France; likewise, that many
expressions are exclusively Canadian and must be marked accordingly in the dictionary.

2.2.6 The editors of CD1 (1962) also adopted a bold and quite broad description of the term Canadianisme in the belief that Canadian readers should be sensitive to the existence of a considerable number of "indigenous" terms in both official languages. The title of the dictionary being "Canadian/canadien" we are not surprised to read the authors' explanation of the need for such a dictionary:

Il est curieux de constater que le Canada, pays bilingue, ne disposait pas jusqu'ici de l'instrument essentiel à toute communication linguistique: un dictionnaire bilingue spécialement conçu pour ses besoins. Dans l'espoir de remédier à cette situation, nous avons composé le Dictionnaire canadien, que nous dédions à tous ceux qui veulent connaître, apprécier et interpréter nos deux grandes cultures nationales (p. vii).

No controversy is aroused by the editors' pronouncement that the English in this dictionary is "Canadian," or at least "North American":

Tous les articles dont la langue de départ est l'anglais reflètent l'usage canadien et nord-américain à moins d'indication contraire (p. xxix).

But the translation of English entries is into international French, except where specific indications point
to characteristically CF usage:

Leur traduction dans la langue d'arrivée reflète l'usage commun au franco-canadien et au français international. Les signes © ou Ô précisent les mots ou expressions d'usage exclusivement canadien ou français (européen) (p. xxix).

The dictionary thus matches Canadian English with International French, thereby emphasizing the marked difference in the status which CE and CF enjoy in Canada.

2.2.7 G. Dulong's preface to the 1962 edition of CASS2 explains the choice of some 200 CF terms included in the dictionary:

La responsabilité de fournir des canadianismes à la nouvelle édition du Cassell's comportait certaines difficultés.

S'agissait-il d'incorporer des milliers de mots ou d'expressions dialectales ou patoisues qu'une enquête permet de relever dans le milieu canadien-français? Non, car ces mots et ces expressions n'ont leur place que dans les glossaires.

S'agissait-il d'incorporer les anglicismes plus ou moins nombreux qui corrompent le français-canadien, surtout dans certaines classes de la société, anglicismes que les enseignants s'efforcent de faire disparaître? Non, c'èut être faire trop d'honneur à des éléments sentis et considérés comme des fautes et des barbarismes.

Seuls ont été retenus les mots (substantifs, adjectifs ou verbes) que j'appellerai canadienisms de vocabulaire. Ce sont des mots concernant des réalités proprement
There is little difficulty in identifying, in this point of view, the prescriptive definition of the OLFQ (see below 2.2.9) for which Dulong acted as a consultant; the small number of words admitted to the dictionary emphasizes the restrictive conception of canadianisme adopted in official circles at this time.

2.2.8  G. Massignon (1962) completed a thorough study of Acadian French, the results of which have enabled other regional varieties of franco-canadien to be identified and given their rightful place. In establishing the difference between canadianisme and acadianisme, Massignon reminds us that the former term covers language features common to all French-speaking areas of Canada, not just those of the province of Quebec:

Tout en ayant un fonds commun (...), le "canadien" et l'"acadien" présentent des différences très nettes: (...). Ce sont des différences de phonétique (quelques traits, touchant un grand nombre de mots usuels), de morphologie (peu de faits, mais des faits changeant...
l'aspect de toute la conjugaison), et du vocabulaire; un certain nombre de mots, usuels chez les uns, sont totalement inconnus chez les autres.

Comme les canadianismes reflètent l'éclectisme des origines de la population française de la province de Québec, de même les acadianismes reflètent les origines relativement plus homogènes de la population française des provinces Maritimes (pp. 99-100).

2.2.9 In 1965 the OLFQ finally laid down its prescription for what constituted a genuine canadianisme. The "Cahier No 1" (Norme du français, écrit et parlé, au Québec) expressed the official view in Quebec thus:

Il faut éviter de considérer comme canadienismes de bon aloi les mots dialectaux ou patois conservés dans la langue du peuple. De tels vocables, même s'ils sont encore vivants dans telle ou telle province française d'où sont originaires les Canadiens français, doivent être remplacés dans l'enseignement par les mots de la langue commune, le but de l'enseignement n'étant pas d'entériner ou de répandre tel mot patois, mais de faire apprendre le mot en usage dans le français du vingtième siècle.

Les canadienismes de bon aloi à retenir dans la langue de l'enseignement et de l'administration se rapportent à des réalités canadiennes pour lesquelles le français commun n'a pas d'équivalents. Ces canadienismes ont trait, entre autres, aux domaines suivants: mesures, faune, flore, poissons, nourriture, politique, hiver, vêtements.

Il faut aussi tenir compte de certains cas où la variante canadienne-française se justifie autant que le mot ou expression du français commun et peut même parfois
leur être préférable. Exemples: fin de semaine, vivoir (p. 9).

It is clear that the "norm" being proposed is one which requires a literary variety of French for all education in the province; although a few "canadianismes de bon aloi" are worthy of retention, the bulk of CF is thus rejected.

In 1969, with its "Cahier No 4" (Canadianismes de bon aloi) the OLFO became even more specific and actually drew up a list of acceptable terms:

Les canadianismes admis par l'Office et sa Commission consultative sont des mots bien formés et de bon aloi, comme on pourra s'en rendre compte à la lecture des articles. Ils sont, en outre, si étroitement liés à la vie quotidienne du peuple québécois, à sa psychologie, à son histoire, à sa géographie, qu'il eût été impossible de les rejeter sans porter atteinte non seulement au patrimoine culturel mais aussi aux moyens d'expression des sujets parlants ( . . . ).

C'est en tenant compte de ces principes que l'OLF et sa Commission consultative ont dressé la présente liste de canadianismes ( . . . ).

Si nous examinons cette liste, nous constatons ce qui suit:

1. Vingt-sept canadianismes remontent au régime français, c'est-à-dire sont antérieurs à 1760: abatis, atoca, batture, bleuet, . . . . .

2. Huit de ces mots sont des amérindianismes, c'est-à-dire des emprunts aux langues des indigènes de l'Amérique: . . . cacaoui, maskinongé, ouananiche, savane, . . . . .
3. Treize d'entre eux, appartenant à l'ancienne langue française, nous sont absolument nécessaires parce que nous n'avons pas le système métrique: acre, arpent, livre, pouce. . . .

4. Sept se rattachent directement à l'hiver. Ex. banc de neige, bordages, frasil, souffleuse, . . . .

5. Treize servent à exprimer ou à décrire le milieu humain dans lequel nous vivons. Ex. abatis, bleutière, coureur des bois, transcanadienne. . . .

6. Enfin des mots d'emploi courant sont le signe évident de la lutte que nous avons à mener contre l'anglicisme: vivoir (living room ou living), verge (yard), oléoduc (pipeline), traversier (ferry). . . . .

Le dénominateur commun de ces canadianismes c'est qu'ils nous sont nécessaires pour décrire le milieu dans lequel nous vivons, qu'ils ont une graphie conforme au système graphique général du français et qu'ils ne constituent pas un obstacle à la communication entre les pays francophones.

Ce sont là, par conséquent, des mots que les francophones du Québec peuvent employer dans le langage soigné ou écrit.

Apart from the excessive rigidity of this definition, there is merit in the suggestion that words and expressions emanating from still-current dialects in France cannot be claimed as canadianismes. On the other hand, those which are no longer used in France, and therefore earn the title of archaism, should not be similarly labelled in Canada where their currency adds to the characteristic flavour of this regional variety of
French. The OLFQ justified their restrictive definition on the grounds that they were advocating the promulgation of French as a language of instruction, a literary vehicle in which there could be no place for patois words or expressions. In maintaining that dialect forms still extant in Europe should be excluded from consideration as canadienisms, they were undoubtedly correct; but in rejecting thousands of other regional forms which have acquired different connotations in Canada they were closing their eyes to linguistic reality.

Recent events in Quebec, especially with regard to language, indicate a marked change in attitudes both at the official level and at the level of the man in the street. The current spirit of nationalistic fervour (which echoes the Occitan movement in Southern France) has contributed to the drift towards a monolingual Quebec whose language, le québécois, instead of being an object of ridicule, has become something of a status symbol. In the face of mounting public opinion, and more recently as a direct result of political events in Quebec, the OLFQ appears to have softened its line on what constitutes a canadienisme. There is still a clearly-defined policy aimed at maintaining the "purity" of the language—in fact, the campaign against the inroads of anglicisms has probably never been more vigorous and more important to Quebeckers—but there is little doubt that dictionary
editors will not again be subjected to the bitter criticisms and recriminations of the fifties and sixties.

2.2.10 An English point of view of what constitutes a French can&ianisme is offered by M. Orkin (1967) in his popularized account Speaking Canadian French:

The distinctiveness of the Canadian French vocabulary derives from two opposing tendencies to which every language cut off from its origins is subject. One (...) is the attachment to "le parler des aîeux," which has resulted in the survival of words elsewhere archaic, as fiable (Fr. **digne de confiance**), or called "dialect" by Larousse and existing today only in provincial French, as **aveindre** "to take something out of a drawer," (...) The other is the coinage of new words, which is the life-blood of any living language. Words are born and words die, and it is in the manner of this birth that Canadian French has come to differ substantially from the French of France, for the conditions that inspire the creation of new words have been very different.

The most interesting class of Canadianisms is that which arises from the peculiar conditions of North American life. The best that classical French could do with Canadian weather was Voltaire's jibe about a few "arpents de neige," but the popular language of Quebec is rich in weather terms, among them **bordée**, a heavy fall of snow, **banc de neige**, a snow bank, (...) (pp. 56-7).

Orkin has correctly identified and emphasized the totally different natural environments of the two countries as being the greatest contributor to linguistic diversity; coupled with the survival of forms which expired in a different set of social and political circumstances, this
has created a distinctive CF dialect.

2.2.11 It is to be expected that, in the absence of official linguistic regulatory bodies, English linguists will have been able to arrive at more satisfying and unanimous definitions of the term **Canadianism**. But in a sense CE suffers from the same plight as CF, namely, the proximity and size of the neighbouring dialect of American English. The mass media has done much to deprive CE of its individuality, although it is at least still nourished by English-speaking immigrants from Britain, the United States and all the Commonwealth countries, whereas French-speaking immigrants dwindled to a mere handful after the English conquest of New France. As a result, it is not always clear whether a given term is exclusively Canadian or equally American or British; individual conceptions of the definition tend to make the distinction somewhat vague in some cases.

M. H. Scargill (1958) was one of the first to explain the problems of definition; either it will be so exclusive as to limit an investigation to a few hundred words, or else it may be so all-embracing that commentary would become superfluous:

This second course is the production of what we now call the *Dictionary of Canadian English on Historical Principles*. Canadian English as opposed to the English language in Canada. This ( . . . ) means an
emphasis on the investigation of the English language in Canada as it differs from as well as resembles British English. That is, our aim is to show what is distinctive about the English language in Canada (p. 115).

This book will probably be called "A Dictionary of Canadianisms" for convenience, although it will by definition not confine itself to words and usages indigenous to Canada. ...

[A] dictionary of Canadianisms, using the term strictly to mean "words originating in and used in Canada" could be produced. But it is doubtful if such a dictionary would be a valid contribution by itself to a history of Canadian culture. ...

I doubt if an intensive search for such indigenous "Canadianisms" would prove sensible. For, as Mr. Lovell himself has written, such an activity would limit our investigation to but a few hundred words.

To compile such a dictionary would not only require tremendous detective work, but it would fail to give anything like a fair and comprehensive picture of the distinctive features of Canadian English as a whole. Mr. Lovell and his publishers are therefore to be commended on their refusal to be strictly confined by a very narrow and exclusive use of the term "Canadianism" (pp. 117-118).

2.2.12 The existence of a "dictionary of canadianisms" would seem to be sufficient evidence of the existence of the phenomenon, but, as the editors of the Dictionary of Canadianisms on Historical Principles (DCHP, 1967) point out in their introduction, it was not a trivial matter to arrive at a definition for the very object of their study:
The purpose of *A Dictionary of Canadianisms on Historical Principles* is to provide a historical record of words and expressions characteristic of the various spheres of Canadian life during the almost four centuries that English has been used in Canada. . . . The problem of establishing a definition for Canadianism as used in the title of this dictionary was not easy to solve. The term is used in a less inclusive sense than the definition given in the body of the dictionary itself, for we are concerned primarily with vocabulary; on the other hand, it is used in a more inclusive sense than Americanism has been used by scholars in the United States. Mitford M. Matthews, for example, defines the term in his Preface to *A Dictionary of Americanisms* as "a word or expression which originated in the United States." Yet in his dictionary he uses a wide variety of Canadian source materials as evidence for a substantial number of "Americanisms," a practice which, to say the least, weakens his definition (p. xii).

The DCHP actually contains an entry "canadianism" which gives a bold and comprehensive definition of the term:

Any linguistics feature, as of pronunciation, morphology, syntax, vocabulary, orthography, that is characteristic of Canadian English (p. 118).

It is obviously important to understand exactly what is intended by the key words in this definition—"characteristic of Canadian English." Does it mean used by Canadians and unknown elsewhere? Or will a word or expression qualify if it is current in Canada but also known elsewhere in the same sense? Some answer is
provided to these questions when the term Canadianism is further particularized 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 (p. xiii).

2.2.13 Other general language dictionaries have also applied themselves to the task of defining the word Canadianism which appears as an entry in the body of the work. The Dictionary of Canadian English; The Senior Dictionary, W. Avis et al. (1967) is somewhat more restrained:

... a word or expression originating in or peculiar to Canada. The words "muskeg" and "caribou" are Canadianisms (p. 163).

The Winston Dictionary of Canadian English (Paikeday et al., 1969) contains a similar entry but also includes French in its definition:

... a word, phrase or expression peculiar to or originating in Canadian English or French (p. 187).

According to this conception, Canadianisms originate in the language and not in the country or in the particular conditions of Canadian life; this is too imprecise for a scientific definition.

2.3 One fact which clearly emerges from this analysis is
the lack of agreement among scholars on a suitable
definition for lexical items peculiar to Canada.
According to J-P. Vinay (1973b: 364-5) three tendencies
erange from reviewing various opinions on the subject:
(a) only new words and expressions, created and used in
   Canada, can be considered Canadianisms; (b) only terms
   relating to Canadian reality, whether they are lexical
   creations or meaning extensions, can be considered;
   (c) anything which differs in any way from Standard
   French (or from American and British English) is a
   Canadianism.

2.3.1 A more precise breakdown of the various conceptions
of this phenomenon, however, yields four distinct
categories. The purist attitude, represented by Buies
(1888), admitted nothing which was not academic European
French; it was an attitude engendered by an intolerable
situation for Francophones, but which was taken to
extremes ("ces choses informes, monstrueuses, innommables,
... "). The French purists had their counterpart among
Britons who, at that period in Canada's history, perhaps
had equal cause to bewail the ever-increasing American
influence on their transplanted language:

It is really melancholy to traverse the
province and go into one of the common
schools; you find a herd of children
instructed by some anti-British
adventurer, instilling into the young
and tender minds sentiments hostile to the parent state (...); and American spelling books, dictionaries and grammars, teaching them anti-British dialect and idiom.

Such emotive judgements, expressing an excess of partiality, have no place in scientific description and are therefore eliminated from consideration.

2.3.2 The second category is one degree removed from the first in terms of its restrictive definition, illustrated by the OLFO for CF but also considered appropriate for CE by those who view it as a subgroup of American English. In its Cahier No 1 the OLFO explains the restraint imposed on its definition of "canadianismes de bon aloi," emphasizing that the norm is applicable, in particular, to three spheres—educational, administrative and judicial:

Parmi les divers usages linguistiques d'une même communauté, il y a toujours un qui prend le pas sur les autres dans certains domaines, tels que l'enseignement, l'administration, la justice, et c'est à partir de cet usage que se constitue la norme (p. 7).

Although the OLFO recognized the existence of language levels, they were clearly at pains to specify that, for them, canadianismes belong to special kinds of language and are admitted for special purposes. Their definition is thus extra-linguistic.
As no official body attempts to regulate English usage in Canada, we have only the point of view of those who, like the OLFQ, recognize a very small number of acceptable terms created and used in Canada in response to Canadian reality. There is no direct equivalent in CE to the official position expressed by the OLFQ for CF. If there were, no dictionaries of Canadian English like those already in existence could have been published.

2.3.3 A third group comprises all terms which have either been created in Canada or which have developed senses in Canada unknown in the original language. The former are lexical neologisms which may be totally new elements, new combinations of existing elements and borrowings from indigenous languages.

2.3.4 The fourth and final category enlarges upon the third by embracing any form which differs in any way from the mother language. This group can be further subdivided into those composed of lexical and syntactic elements only and those which include all kinds of linguistic phenomena (lexical, morphological, syntactic, phonological, orthographic, etymological). For English lexicographers there is little difficulty in introducing even the most extremely inclusive definition of Canadianism; in the absence of official linguistic prescription editors can include almost anything in their
dictionaries and have only the buying public to answer to. Thus the DCHP characterized *Canadianisms* as any linguistic feature peculiar to the English language in Canada; it was not even necessary that the words or expressions originate in Canada or that they conform to some standard of "bon aloi." But French lexicographers operated with much less freedom and their conception of *canadianisme* has undoubtedly been coloured by the requirement that they conform to official norms or publish at their own peril. J-P. Vinay (1973b) sums up the position thus:

Il est dangereux, à l'heure de dirigisme, de parler de canadianismes qui n'ont pas "reçu le feu vert" de l'Office; il est difficile, dans ces conditions, de donner une véritable image du CF (canadian français), à moins qu'on ne possède des moyens d'existence indépendants (p. 365).

2.3.4 Between the two extremes categorized above, there are a number of lexicographers and linguists who have attempted to produce extended inventories of *Canadianisms/canadianismes*. However, the existence of these dictionaries does not necessarily bring us any closer to a satisfactory definition. V. Barbeau (1963), for example, has provided large numbers of popular words and expressions for which he gives no reference to language level; nor is there any etymological indication, which prevents any judgement from being made on whether
the *canadianisme* in question is really a distinctively Canadian form or whether it is merely a dialect form transplanted from France where it is still in current usage. Since it would be useful to be able to distinguish between diachronic and synchronic facts, such gaps in the documentation reduce the effectiveness of otherwise valuable materials.

2.4 In order to reach a satisfactory conclusion on the definition of *Canadianism*, the object of regional illustration in a bilingual dictionary in Canada, we must take into account the preceding discussion. But if the definition is to be purely linguistic, there can be no appeal to motivations arising from pedagogical, administrative, political and other extralinguistic considerations; we must, therefore, confine our discussion to language features. The first problem to be resolved is what kinds of linguistic phenomena are to be considered. We have observed that most of the attempted definitions confine themselves to lexical facts, which are the most readily observable, the most fully documented and easiest to analyse. Phonological features might be included in the analysis, but since almost every element of CE and CF differs in pronunciation from the original language, such divergences would mean labelling every lexical unit (or at least its phonetic form) given.
in the dictionary as a Canadianism. Lexicographers usually prefer to discuss pronunciation in their preface and omit regional variation of this kind from the body of the work. Phonetic divergence is certainly a significant linguistic feature, but one which seems an unlikely parameter for use in the present discussion.

2.4.2 Syntactic divergences, although far less common than phonetic, are of considerable importance and must be considered in the analysis. The problem with this kind of material is that it usually implies a value judgement on the part of the linguist. A lexicographer must decide, for example, whether the following expressions are to be proscribed as ungrammatical, included as illustrative material with an appropriate cautionary remark, or promulgated with the status of acceptability and as characteristic of the regional dialect: c'est propre effrayant (= c'est extrêmement propre); cette salle est grande en masse (= ... est extrêmement grande); le plus beau qu'il y a pas (= le plus beau qu'il y ait); it's real nice (= it's really nice); to rest a bit (= a while); to be some hot (= very hot); snuck (= sneaked).

The question of acceptability in such cases is largely related to frequency, which eventually tends to mask the underlying "grammatical" structure. At some point in language development, perhaps when someone is
bold enough to commit them to the published page in unquoted form, these new structures become the norm for a given speech community. This evolutionary process may well be a kind of universal within languages, for as J-P. Vinay (1973b) points out:

La liste des "fautes" relevées à ce sujet par Aubert (Ier Congrès LFC, 1914) comporte un certain nombre de formes maintenant courantes en FS, ce qui illustre encore une fois le parallélisme des évolutions CF/et FS: "Bon pour manger; attachez ce cheval après le poteau; que ferai-je avec tout cela? Une médaille en or. . . . (p. 361)."

2.5 The lexicon is undoubtedly the language feature most appropriate for inclusion in the definition of Canadian-ism. This is not simply because most investigators, particularly the purists in French Canada, have restricted their consideration to "words" (c'est le seul domaine où il soit possible d'accepter des divergences), but because it is a phenomenon that is neither open-ended, like phonology, nor difficult to analyse and categorize, like syntax. Lexical features may be studied according to their form or their content and regional forms may be categorized according to a number of parameters.

2.5.1 Etymology: the origins of a word may be used as evidence of its synchronic status. For example, CF or CE words may be traced to English or French terms respectively, or to other contiguous or indigenous languages; in
order to establish them as **Canadianisms**:

- CF **achigan** (fish) < Amerindian
- CF **conseiller de la Reine** < CE Queen's Counsel
- CE **voyageur** < CF voyageur
- CE **chowder** < CF chaudière

There are various reasons, however, for regarding etymologies with care. As the editors of DCHP explain, words which have entered the language through Amerindian must be treated with great circumspection since the documentation on them usually dates back to sources compiled by "linguistically naive writers" (p. xvii).

CF terms documented in the GPFC must be treated with equal caution. Published in 1930 by the *Société du Parler Français au Canada*, the *Glossaire* claimed a dual role:

"l'étude scientifique du langage de notre peuple, et la correction des fautes qui s'y trouvent" (p. vi). Taking this statement together with the subtitle of the work—"Mots et locutions en usage dans le parler de la province de Québec et qui ne sont pas admis dans le français d'école," it is not difficult to predict the boundaries of its application: it is a descriptive investigation with its own interpretation of the status of the term **canadianisme**.
The research for the publication began about the turn of the century and it therefore could not draw on good research materials such as the *Atlas Linguistique de la France* and W. von Wartburg's *Französisches Etymologisches Wörterbuch* (FEW). The GPFC also limits itself to CF of the province of Quebec, its authors having judged that only "le parler du groupe plus considérable des Canadiens français, habitants de la province de Québec, peut être dit proprement le français du Canada" (p. vii), and designates Acadian French as an object for independent study since it constitutes a special linguistic zone.

Entries in the GPFC which are considered to be *canadianismes* are marked "Can." but, as the following example illustrates, it is not always clear whether it refers to the entry as a whole or to the last sense since the indication appears at the end of the article; (normally it is placed after the sense(s) to which it refers):

**cage** (kà:j) s.f.

1. Train de bois, raideau, assemblage formé de pièces de charpente ou autres, liées ensemble pour leur faire descendre le courant d'une rivière sans les charger sur le bateau.

2. Pile de branches, de morceaux de bois en échiquier.

3. Espèce de quai construit au milieu d'une rivière pour y attacher des estacades.

**Can.** - Dame m.s.
Although the indication of *canadianisme* seems to apply only to the third sense, etymological research will reveal that none of these meanings appears in modern French dialects. The entry *cavea* of FEW provides examples of some regional EF forms which also convey the notion of "treillis" (or construction of some kind) inherent in the CF senses. For example, in Modern French *cage* is defined as "treillis dans lequel un orfèvre étale sa marchandise," in Savoyard the word means "buffet d'osier."

Let us now pass to an entry where etymological information is supplied in some detail. The word *cheniquer* is listed as a *canadianisme* with the definition "renoncer à une entreprise, céder, se dérober" and the following details are provided on its origins:

**Dial.**  --*cheniquer = dérober, boire,*  
Bas-Maine;  
*cagner = céder,* Aunis.

**Fr.**  --*Pop., cheniquer = dérober, boire.*

**Etym.**  --*Cf. ang. to sneak,* m.s. prononcé 
*shneak* par les Irlandais.

Once again the information is far from clear; the dialect variations are confused and the reference to Aunis "*cagner*" is of doubtful relevance. From the rubric "*Etymology,*" the reader is led to believe that CF
cheniquer is derived from English sneak (m.s. = même signification); or else that the CF sense "renoncer à une entreprise," being unknown in France, is a canadianisme de sens and is the result of CE influence in Canada.

The explanation provided by the GPFC is the one which has found most universal acceptance, in spite of lengthy debate which has centred around this particular etymological problem. According to the FEW, the origin of the verb cheniquer in the sense "boire" is the German schnick (brandy) which gave *schnicker (to quaff brandy) in the Manceau patois. Rivard (1914: 198), in discussing this etymology, rejects the notion that the German word could have exerted an influence on the particular sense found in CF; he suggests that one school of investigators considers the CF word to be derived from CE to sneak, while another has compiled a derivation which claims Latin *canicare as the origin, giving cheniquer in various French dialects. But, as Rivard maintained, this is an unlikely development, since such a form has not been attested; even if it had existed, it would have given *chenier or *chener because of the well-known tendency of isolated medial consonants to be dropped in the transformation from Latin to French, for example: judicare > juger; mendicare > mendier; manicare > manier.

This leaves the derivation from English sneak. For
this route, says Rivard (1914: 202) English [*s*] would have to be transformed into French [*ʃ*]; but such a transformation would have had to occur before the word became Gallicized, and since no such examples have been attested, he rejects this hypothesis.

Having eliminated all the hypothetical etymologies, Rivard and a number of others (including, presumably, the editors of the GPFC) concluded that [*s*] was in fact already [*ʃ*] in CE before the word passed into CF. Their solution which they propose, more as a conjecture than a hypothesis based on fact, is that Irish immigrants in Canada pronounced words like sneak, slide, bliss, etc. as [ʃniːk], [ʃlaːd], [blɪʃ]. Rivard concluded: "nous pouvons donc écrire: ang. sneak → irl. shneak → can. cheniquer" (p. 203).

The validity of this proposed etymology could be established only if sufficient information were available on the concentration of Irish settlers in Quebec relative to the English population. The question could then be asked why they were able to exert influence only on the word sneak and not on many other CE words of similar phonetic configuration which passed unaltered into CF.

As none of the suggested solutions seem acceptable, we are obliged to revert to the FEW derivation, from German *schnick* to French *cheniquer*. In Bas Maine and in popular EF, as the GPFC confirms, the word has two
meanings: "boire" and "dérober"; the first is easy to comprehend in view of the sense of the German etymon; the second, for which FEW gives no assistance, is much less transparent and is the reason why Rivard and others sought different answers to the problem. One zone gives "s'enivrer" as a sense; from this it would be possible to postulate "dérober" as a metaphorical derivation drawn from the notion of the eventual flight of the drunken person—often with shamefaced stealth. Whatever the origin of the sense "dérober," the CF meanings "céder, renoncer à une entreprise" are much closer to it than to English sneak which means "to slink, go furtively, make off with, steal" (OED). The linguistic area in France where schnick is believed to have given cheniquer, Bas Maine, happens to be one which provided large numbers of colonists to Canada. It seems probable, therefore, that they imported their word cheniquer ("boire, se dérober") and that, in contact with CE sneak (which in the 17th and 18th centuries also had the meaning "to lack courage, independence or straightforwardness"), it acquired the sense "céder, renoncer à une entreprise." 

This lengthy digression into the etymology of a single word serves to illustrate the degree to which etymologies must be treated with caution when attempting to determine the status of regional elements. In dealing with such forms, the synchronic situation in the mother
language must be taken into account; if a word or expression employed in CF or CE can be traced to a dialect in European French or English where it is still current, it cannot be claimed as a Canadianism. For example CF crémer (to freeze) comes from Bourgogne; CF barrée de glace (ice-jam) from Anjou; CE yelk from Lincolnshire, Gloucestershire and Cumberland; CE some + adj (= very) from Cornwall, Lincolnshire and Lancashire. On the other hand, CF and CE terms which came from European dialects which no longer employ them, may justifiably be claimed as Canadianisms. Thus the CF measures of capacity gallon, chopine, demiard were left in Canada after the Conquest but were eventually replaced in Europe by the metric system.

2.5.2 Linguistic processes: the second parameter to be considered involves the linguistic processes applied to the development of the Canadianism as a distinct form. The simplest process is meaning change, where a lexical unit enlarges its semantic field, usually by association, in order to embrace a new sense: CF poudrerie (blizzard), bleu (conservative, in politics); CE reeve (councillor), tickle (channel), bluff (clump of trees). Under this heading could be included a number of borrowings which have acquired a different meaning in their new environment: CF smart, smat (< CE smart) is given by the GPFC.
the meanings "alert, supple, lively, skillful, active, intriguing, cunning, elegant, nice, friendly, distinguished," whereas its English counterpart means "severe, painful, vigorous, lively, skillful, intelligent, elegant," and in CE also "impudent." Thus, in the CF expression il est bien smart, the frequently used meaning "nice, kind, distinguished" must be considered a canadianisme.

A more complex process is lexical creation; it may take the form of new combinations of existing words: CF bordée de neige (snowstorm), gomme à mâcher (chewing gum); CE dew-worm, concession road; morphological extensions of existing words: CF magasiner (to go shopping); CE longliner; or totally new creations: CF berlot (winter vehicle); CE nunny bag. Morphological derivation may occur with a change of meaning: CF traversier (ferry); érablière (maple grove); CE wintering; or with extended meaning: CF raquetteur (wearer of raquettes); CE (forestry) driver; poudrerie (blizzard).

Probably the most frequently-used device is borrowing from contact languages: CF achigan < Amerindian, couque < CE cook; CE portage < CF portage, oomiak < Eskimo. In this process there is always a phonetic adaptation to the borrowing language and usually morphological assimilation as well. The transformation sometimes all but obscures the origin of the word: CF
chasse-pinte (saucepan), cotte (cut); CE snye (chenail), Siwash (sauvage); in some cases there has been confusion with regard to the meaning of the original: Picketwire (purgatoire); snicrouëte (snake root), mush! (marche!).

CF makes frequent use of calques to fill lexical gaps, usually ones which are created by the existence of Anglo Saxon institutions to which the French population has been obliged, federally at least, to adapt:

Conseiller de la Reine (Queen's Counsel); l'honorable X (the Honorable Minister . . . X); la Colline parlementaire (Parliament Hill). But CF is equally capable of creating its own equivalents which are not "transliterations": Editeur Officiel du Québec (= Queen's Printer); Assemblée Nationale (= Legislative Assembly); and even improvises where SF accepts English words: traversier (ferry); oléoduc (pipeline); fin de semaine (weekend); vivoir (living, living-room).

2.5.3 North American reality: the third parameter to be considered in attempting to establish the status of Canadian regionalisms is the reality which transplanted languages are called upon to describe. Colonists arriving in a vast new land of boundless natural wealth had to use a remarkable degree of linguistic ingenuity in order to give names to totally new referents. We have already observed a number of the processes used by the
two language groups which concern us here to deal with this problem: borrowings from indigenous languages if they already possessed a term, borrowings from each other, semantic extension of existing terms by analogy or specialization and finally lexical creation or neology. The resulting lexical units, contrived by Canadians of both language groups to depict Canadian reality, can scarcely be considered anything but Canadianisms.

2.5.4 In order to observe how the processes of linguistic creativity operate and the manner in which it is applied to Canadian reality, we will now examine a particular sphere which has produced a wealth of new or changed terminology. It will be no surprise to anyone to be told that the first and most devastating natural phenomenon faced by new arrivals in Canada was the climate. Accustomed to generally milder conditions in Europe, English and French-speaking colonists were shattered by the harshness of the Canadian winter and eventually sought to describe it in a colourful range of terminology. The linguistic resources available in order to generate this terminology are as follows:

(a) **Analogy** -- adaptation of existing words or phrases to new concepts, sometimes by a process of commutation: CF banc de neige (snowbank) by association with EF banc de sable and banc de glace;
CF bordée de neige (heavy snowfall) by analogy with Saintonge bordée de pluie, bordée de grêle; CE white-out (cf. blackout), glare ice, rubber ice, whirling snow.

(b) Change of meaning -- expansion of the semantic field, often by analogy: CF poudrerie/CE powdering; balise (Old French = road markers or posts to indicate the way for travellers; Saintonge = tree planted to mark a boundary; CF = winter road markers); but most often as a simple augmentation of related senses: CE frazil/CF frasil (= ice crystals in rivers, etc.) < EF frasil, fraisil (= clinkers) < Old French freser, fraser (= crush, grind into fragments); CE jam (of ice), hummock (of ice), glitter or glidder (white frost).

(c) Morphological derivation -- CF raquetter, raquetteur < raquette; CF brunante (dusk) < brune (à la brune); CF neigeasser, neigeailler, neigeotter < neiger

(d) Compounding -- whereas CF most often uses morphological variations, CE resorts to the Germanic process of compounding words to create new lexical items: CE break-up, drift-ice, winterkill, snow fence, silverthaw. (Three stages of development are indicated by the presence or absence of hyphens and
the continued separation of some compounds.) CF does so to a lesser extent: Boueige (= boue + neige).

(e) Metaphor -- the creative and artistic qualities of language are brought to bear in likening winter phenomena to other properties: CE rubber ice, blue ice, black blizzard (actually a dust storm); CF chasse-neige (snowplough) gazon de glace (flbating ice), roulis de neige (wind-blown waves of snow).

(f) Neology -- CE brash (ice fragments), permafrost, dwye (snowstorm); CF tuque (hat), doux-temps (thaw).

(g) Borrowing -- CE chinook/CF chinook, chinouque < Amerindian; CE cahot < CF cahot; CF claireur < CE clearer; CE sastrugi < Russian. 11

This brief summary of some of the procedures employed by the two languages to describe meteorological phenomena serves to illustrate the general development of regional variation and provides further evidence for the establishment of a satisfactory frame of reference for the labelling of regional entries in a dictionary. (A complete inventory of CF and CE winter terminology is included in Appendix VI).

2.6 The definition: Our analysis of the various
interpretations given by a number of scholars and authorities and of the naming procedures employed by two national languages isolated from their country of origin provides an adequate framework for the identification of Canadianisms. We have observed that the definition of regionalisms in Canada varies widely—there is considerable disparity of opinion within each language group as well as from one language to the other. But lexicographers are obliged to adopt an interpretation which is neither excessively restricted and prescriptive nor overly inclusive; either of these viewpoints would upset the balance of a dictionary purporting to be a translation tool for the speakers of each language.

2.6.1 One practical approach to the problem is to arrange the various kinds of regional forms into categories according to source, physical structure and linguistic function, and then, from a synthesis of the results, to derive an appropriate definition. From the foregoing discussion we may classify regional forms in Canada as follows:

A. STRUCTURE

1) Phonetic variation -- petit: IF [petɪ], CF [pɛtsɪ]; production: BE [pæˈdækən], CE [ˌpjuˈdækən].

2) Orthographic variation -- BE tyre, CE tire; IF plumer, CF pleumer.
(3) **Morphological variation** -- étang (IF masculine, CF feminine); IF pêcheur, CF (Acadian) pêcheux; BE lake boat, CE laker.

(4) **Syntactic variation** -- IF être en train de faire, CF être après faire; BE at the back of, CE in back of.

(5) **Compounding** -- CE anchor ice (ground ice).

(6) **Frequency** -- CF cédule (schedule) for IF horaire, etc.

**B. ORIGIN**

(1) **Archaïsms** -- dépareillé: CF = priceless; IF = odd, unmatched.

(2) **Dialect** -- CF berlander (= IF flâner); CE riding (= BE electoral district).

(3) **Borrowings** -- CE oomiak < Eskimo umiak; CE snye < CF chenail; CF micouenne < Algonquin emihkwen.

(4) **Anglicisms** -- CF mitaine < meeting place.

**C. FUNCTION**

(1) **Semantic change** -- poudrerie: CF = blizzard; IF = powder factory; cookery: CE = camp kitchen, rendering plant of a whaling station; BE = art of cooking.

(2) **Metaphor** -- CF bordée de neige (heavy snowfall); CE corduroy road (made of transversely laid logs).

(3) **Neology** -- CF patinoire (skating rink); traversier (ferry); CE brash (ice fragments), snowmobile.

(4) **Calques** -- CF gomme à mâcher (chewing gum); CE normal school, separate school.

2.6.2 Taken in isolation many of these features might not
be sufficient to justify their consideration as manifestations of a regional dialect. For example, pronunciation differences alone, however distinctive, are unlikely to influence editors that they should merit a place in the dictionary. In the case of CF and CE such decisions are invariably governed by non-linguistic factors, such as the establishment of pedagogical norms and the vague notion of acceptability. Therefore, although in purely linguistic terms the distinctive pronunciations of both dialects are highly relevant, from a lexicographical point of view they will demand independent treatment. The pattern used in CD2, for example, is to give all pronunciations of English words in a generalized form which was a supposed extrapolation from the variations observable throughout the country. This is the most that can be done. In a bilingual dictionary, after all, the pronunciations are supplied as an aid to the non-native speaker not as an attempt to exemplify the synchronic state of affairs with regard to the phonology of CE as a whole. The phonetic transcription of the French entries in CD2, on the other hand, is manifestly IF, which has its own problems with regard to the establishment of a norm for the Francophone world. If all users of that dictionary acquired the pronunciations given, a theoretical possibility, but one which is highly improbable since most second language learners
imitate the accent of their teachers or acquaintances, they would be in a position to use it anywhere in the English or French-speaking worlds without facing ridicule. The preface of the same dictionary provides a succinct outline of CF and CE phonology and any highly motivated reader of the work is able to gain insights into regional variation in this area. One refinement which could profitably be applied to French entries in a bilingual dictionary is the provision of CF pronunciation for explicitly CF words like achigan and brunante. Thus, while we may speak of phonetic Canadianisms, they will be considered as a separate phenomenon.

Similar remarks apply to orthography. In CE, the variation from BE is slight and revolves around a few words like tire and jail, influenced by AE and highly fluctuating pairs such as -our/-or, -ise/-ize, -re/-er, -ce/-se and -ll/-l-. The orthographic differences between CE and BE are in fact so slight that they can scarcely be considered as constituting an aspect of regional variation. Orthographic differences between CF and IF are related to phonetic distinctions: cuive/cuivre, crever/crever and so on. The GPFC treats these deviations very fully giving, for example, as many as nine variants for the word "chaudière": chaquiére, chaquère, chaudiére, chauquérée, chayère, chauvère, chayéré and cheillère. If a decision were made to
incorporate such variants in a bilingual dictionary, they would presumably be entered as *canadianismes*, with appropriate pronunciation, and cross-referenced to the relevant entry for full treatment.

2.6.3 Because the basic task of lexicography is to deal with lexical units, it is therefore lexical features which will be the prime consideration in an attempt to establish a working definition of regionalisms in the dictionary. Furthermore, since the linguistic processes called upon to create new lexical items conform to the normal procedures of word creation they should not occasion any controversy. This leaves us with the origins of the words as the single most important feature in determining their acceptability as bona fide regional forms.

2.6.4 The first subcategory in this classification relates to *archaisms* which are a frequent manifestation in CF. As a consequence of historical factors, the language used by a majority of French colonists, following their separation from the mother country, persisted in French Canada in the absence of the diversity of factors which influenced the development of the language in France and saw the dialect of the Ile de France emerge as the dominant idiom. What is to be the status of lexical items which, from the point of view of modern Standard
French, are archaic? No global statement can be applied to them as a whole because different circumstances apply to different cases. For example, terms relating to weights and measures under French rule continued to be employed after 1760 with appropriate adaptation to the English system. In the meantime, France converted to the metric system and the old measures degenerated into the realm of "literary archaisms." But the terms which persisted in French Canada—acre, arpent, chopine, demiare, gallon, pinte, pouce, etc.—had official status and reflected Canadian reality; they should therefore be designated as canadianismes. In the light of the current conversion of Canadian measures to the metric system, however, these words will once again undergo a process of degeneration and pass from current usage.

It would be tempting to suggest that lexical items, such as measure terms, which have survived linguistic degeneration should be the only ones accorded the status of regionalism, while all others are simply archaic forms just as they are in Standard French. This is in fact the position taken by the OLFQ which admitted only the small handful of measurements because they were "absolutely necessary" in the absence of the metric system. Yet these archaic forms, sometimes existing side by side with the modern word, often used with greater frequency and frequently acquiring new senses in their new environment,
are precisely the elements which bestow upon CF its distinctive flavour. It seems reasonable, therefore, to consider such words as likely to conform to the requirements of the definition in most cases; each should be treated on its individual merits. Moreover, words which are labelled "archaic" in one synchronic period may undergo a regeneration in another. Such is the case for the word fiable, abandoned in Canada with its seventeenth century meaning of "reliable" and consequently discarded in France in favour of sûr digne de confiance and various other circumlocutions. In this century the word has experienced a rebirth, first in technical senses referring to the "reliability" of machines and consequently, by extension, with reference to people and institutions. The word has, so to speak, turned the full circle in European French while remaining constant in CF.

2.6.5 The second subcategory of our classification of the origins of regionalisms in Canada refers to dialectal forms; these are words which can be traced to a particular dialect in the European mother language. Again, the most obvious criterion is that of currency in the contributing locality; if crémer in the sense of "freeze" can be traced to Bourgogne, is it legitimate to claim the word as a canadianisme? It would seem that a
distinction must be drawn between diachrony and synchrony and that a word which is still in current use in an identical sense in its place of origin should not be labelled as a regional form in Canada. Yet even this formula cannot be applied indiscriminately because it is highly improbable that the "freezing" experienced by Burgundians bears any similarity to the ferocious cold experienced by their descendants in Canada. Each item must therefore be subjected to careful study before it is labelled either "co-occurrent dialect" or "Canadian regionalism" because, except where there is a significant concentration of speakers of a given dialect, the usual tendency on the part of composite groups is to employ the common core of their language and to relinquish the more specific elements of their own idiom. Therefore, those dialect forms which did survive must have either fulfilled a particular need or have developed new meanings. The word riding, for example, is first attested in CE in 1792 in the sense of "electoral district"; in BE it is an "administrative district" restricted to the county of Yorkshire and has never been used in the electoral sense. 14

2.6.6 Borrowings, the third category of origin under consideration, are less likely to be a cause of contention in the debate on definition; a language borrows
items from other languages for the sole purpose of filling the lacunae in its own inventory. We have observed that CF and CE borrowed from each other, and from other contiguous languages, in order to account for the phenomena of North American reality. Such borrowings are usually adapted to the structural and syntactic restraints of the host language and as a legitimate linguistic process cannot be censured or relegated to the status of unacceptable items. The only reservation which might be made is with regard to the special case of anglicismes in CF. These can be classified into various levels of acceptability: (a) those which conform to the criteria just described with reference to the motivations for borrowing: for example, words like drave, draver, draveur (< CE drive) relate to Canadian reality and have been suitably adapted; under this heading also would appear calques of the type Chambre des communes; (b) those which were not needed to replace lacunae but whose frequency of use has conferred upon them an element of respectability: bàdrer (< bother), bécosses (< back-house); (c) words which acquired extended meanings in CF: mitaine = church (< meeting); (d) a number of phonetic deviations whose Gallicized form bears little resemblance to the original and seem to constitute independent forms: arlérose (< early rose), snicroûte (< snakeroott); (e) large numbers of words whose presence cannot be
justified: cotte (< cut), tchipe (< cheap).

It is probable that items belonging to (a), (b) and (c) can with justification be labelled as **Canadianismes** and incorporated in the dictionary as full-status entries. Those which belong to (d) will require special care before a decision is made, while the members of (e) will in general be treated as unacceptable or entered with special conditions. 15

2.6.7 In the light of the foregoing discussion we are now in a position to postulate the definition of **Canadianism**/canadianisme appropriate for lexicographical purposes:

A Canadianism is any lexical element (or group of elements) in current use, regardless of origin, which answers the linguistic requirements of its users and is peculiar to the geographical area of Canada. Excluded from consideration are phonetic and orthographic variants, dialect forms which imitate European usage and non-relevant **anglicismes**.

3. Regionalisms in the Dictionary

Once the identification and classification of regional forms have been appropriately studied, the lexicographer can proceed to the somewhat less demanding (although no less controversial) question of entering them in the dictionary. It will be clear, from the preceding discussion, that this task cannot be attempted until the status of each lexical unit to be processed is
clearly described and the sociolinguistic context in which it occurs is fully comprehended. We will continue to illustrate this approach to regional variation in lexicography using English and French spoken in Canada.

3.1 The linguistic situation of French in Canada has been depicted by Vinay (1973a: 184) as shown in Figure 11.

The overlapping area \(ab\) of the Venn diagram represents the zone of mutual intercomprehension, the so-called "International French" (IF), while \(a\) and \(b\) are the zones of regional forms peculiar to Canada and France respectively and not generally comprehended beyond those geographical limits. Examples are: achigan, bourgot, tire (a); toit, rue, train (ab); chai, lamanage,
dépotayer (b). From this we conclude that the zone ab represents the "common core" of French and will constitute the major part of the lexicon of target language in one side of the bilingual dictionary and of source language in the other. A dictionary presenting the French used in Canada will also incorporate large quantities of material belonging to zone a; the amount which is included will depend on considerations such as the definition agreed upon for canadianismes, the scope and destination of the dictionary (educational, encyclopedic, technical, etc.) and the degree to which these terms are acceptable to linguistic regulatory bodies. Elements from zone b will probably be selected on the basis of frequency or literary relevance, again depending on the size and objectives of the dictionary.

3.2 A similar schema, shown in Figure 12, can be employed to depict the situation for English in Canada. At first sight this diagram appears to describe identical phenomena to those of Figure 11, but this is not entirely correct as we have seen from the discussion above. To begin with, there is no universally accepted version of literary or scholastic English equivalent to IF (instructed as such in educational institutions and fostered by organizations like the Alliance Française). The most we can say about the differences between CE and
British English (BE) is that they are two variants of the same idiom; by far the most common variations are phonetic and orthographic and in general there is total mutual intercomprehension on both sides of the dividing-line, to a far greater degree than would be true of CF and EF. The area depicted by $xy$ in a Canadian dictionary will thus comprise a version of English which is fully intelligible on both sides of $xy$ but spoken and written only in $x$. It is to be expected, then, that the components of $x$ and $y$ will be restricted to relatively small numbers of lexical units exclusive to the two linguistic zones. The status of $y$ is, in fact, identical to that of $b$, and only those items of great frequency or whose use may lead to ambiguity would be included. For
example BE boot (trunk), cot (crib), public school (independent school), inflammable (flammable), all of which have quite different connotations in CE.

3.3 The editors of CD1 took this situation into account in producing their "Canadian" dictionary with the result that, whereas the French lexicon represented the zone ab (of Figure 11) with a few hundred references to regionalisms in a and b, the English lexicon described x and xy (of Figure 12) with occasional marking of forms from y. Thus, the source language for English is composed of [x + xy] and for French of ab and the respective target languages are [ab + a] and [x + xy], where the ordering of the symbols reflects the relative importance of the linguistic zones which they represent. In other words, French entries are provided with translations in CE only, with occasional references to BE, while English entries are directed towards IF with a significant number of CF lexical units.

Therefore, what appears, in algebraical terms, to be an imbalance is merely a reflection of the linguistic reality in Canada. At the time of publication of CD1 (1962), CE enjoyed complete autonomy at all levels of usage, whereas CF was restricted to non-academic or descriptive levels. The situation has changed remarkably since 1962; the current version of the same dictionary,
CD2, will provide a much expanded representation of CF, illustrating the degree to which the dialect has been obliged to adapt to North American reality and technology. It may well be that a future CD3 will emerge in which the a of Figure 11 enjoys equality with x in Figure 12, both as a source language and target language.

3.4 The division of the two national languages into two overlapping regional forms for each is, however, an oversimplification. CE, for example, is greatly influenced by American English (AE) and large numbers of words and expressions in daily use in Canada properly belong to a third zone which might be depicted as in Figure 13.

Figure 13. French and English in North America.
In this schema, the zone \( xyz \) (combined horizontal; vertical and diagonal shading) is the "common core" of English employed in all three countries. The horizontal shading on the outside of \( xyz \) represents the area of forms common to CE and AE; these would be either unmarked in the dictionary, indicating that they are words in general use and probably comprehended and even used in \( y \), or else marked as exclusive to CE and AE. Certain elements of \( z \) would appear in the dictionary, marked as Americanisms, indicating that they-are terms of American origin but of significant frequency and importance to warrant their knowledge by Canadians.

Further categorization of the variation in the English used in Canada is possible in zone \( x \) which contains lexical units marked as exclusively Canadian. While these items may in general have wide territorial extension, it is evident that some will be restricted to certain geographical areas, thus constituting regionalisms within a regional dialect. Research materials are now available which permit more accurate specification of the regions covered by such forms; the DCHP is the most authoritative source and further information can be gleaned from Scargill and Warkentyne (1973) who present the results of a nation-wide survey on lexical and phonetic variation in CE. While the editors of CD1 were prevented from advancing beyond the global marking of
Canadianisms by a lack of such materials at the time when their research was carried out and eventually published, current lexicographical undertakings can correct this deficiency and enhance the value and accuracy of their work by specifying the areal restraints imposed on regionalisms in the manner described below.

3.5 The situation for French in North America is no less intricate, but once again of a totally different kind of complexity. In the first place, the "international" variety of French contains a small number of belgicismes, suissismes, algérianismes and other forms borrowed from Francophone territories in Africa, North America and the Pacific. A reformulation of Figure 10 could be carried out to reflect this state of affairs. But like CE, CF has its own regional variation stemming from the original separation of French since 1608 into québécois and acadien. The former can be further subdivided into varieties spoken in Montreal, Quebec City, various parts of the province of Quebec and further west into predominantly English-speaking provinces. Historical conditions have spread pockets of Acadian French throughout the Maritime provinces and into Louisiana and Pennsylvania in the United States.

Once again, research materials now enable more accurate labelling of the internal regional variation of
CF; notable among the sources is the Centre de Dialectologie at the Université Laval (directed by G. Dulong) which will eventually contribute to the proposed Atlas linguistique du Canada français.

3.6 The concept of translation equivalence is the cornerstone of bilingual lexicography; editors endeavour to provide translations which match the source word in every linguistic feature, and one of their chief tasks is the disambiguation of equivalents by means of various discriminatory labels. The incorporation of indicators to distinguish between regional forms therefore seems only a natural extension of what has proved to be an adequate and workable system. It is just as important, for example, to inform the reader that glidder (hoar frost) is used only in Newfoundland as it is to prevent him from assuming that the word versatile is identical in French and English.

The manner in which these regionalisms are marked in the dictionary is a matter of personal preference and typographical aesthetics. Obviously it is insufficient merely to append the label (for Canadianism) to our example glidder, since this would imply to French readers that the translation given, gélée blanche, could invariably be rendered by glidder in Canada. Further precision of the territorial extension of the word may
well be required also by most English readers west of Newfoundland who would probably consider it to be a misprint for *glitter*; (it is actually a phonetic variant of the same term).

One means of showing the regional boundaries is to add to the label a symbol which would warn the user of a restriction in its dialectal extension:

\[ \text{glidder} \quad n \quad \text{[Meteor] gelée f blanche.} \]

But although this may sound a mild note of caution for the reader, it really does little else and he would be obliged to consult another source in order to determine whether he is justified in using the word in a given context or not. Since such sources are often not readily available, and since dictionaries are supposed to cater to user convenience, it would seem appropriate to expend a small quantity of typographical space in order to provide the relevant information thus:

\[ \text{glidder} \quad n \quad \text{[Meteor] gelée f blanche.} \]

3.7 The introduction of regionalisms in bilingual lexicography calls into question the integrity of the translation equation upon which the effectiveness of the dictionary depends. There is some justification in posing the question whether the balance of the equation
can be maintained without rendering regionalisms in one language by regionalisms in the other. To what extent, also, should regional forms be offered as equivalents for source language words in the "common" idiom? Ideally, it would seem that a term \( a \) (in Figure 10) should always be equated with a term \( x \) (in Figure 11), but this is clearly impossible since regional variation is manifestly language-specific. In general we will have \([a + xy]\) or \([a + ab]\), although it is not entirely inconceivable that \([a + x]\) would exist or that multiple equivalents would give \([a + xy + x]\) and other possible combinations.

The following (hypothetical) entries serve to illustrate some of the problems involved in labelling regionalisms in a bilingual dictionary. (The label \( C \) immediately following the grammar indication refers to the entry word; if the translation is a regional form, the label follows it):

- **maller** vrt \( C \) [Poste] to mail (a letter).

This is an example of \([a + xy]\) in which a CF entry is translated by a general English term; the reverse of this case is \([x + ab]\), in which a CE term is rendered by IF:

- **clinkerbell** n \([C, \circ Nfld]\) [Meteor]

- **glaçon** m

The reasons for labelling regionalisms in a bilingual
dictionary are clearly justified by the next entry, 
[xy + a + b], in which no "common" term exists; the two translations are clearly marked to indicate the geographical areas in which they would be comprehensible:

\[
\text{collect call} \quad \text{loc n} \ [\text{Tel}] \ \text{appel h/à} \\
\text{frais virés C, en P.C.V. Fr}.
\]

In some cases the lexical unit is a "common" term in one sense and a regionalism in another, giving a compound example of [ab + xy] and [a + xy]:

\[
\text{malle n} \ 1. \ [\text{Poste, Transp}] \ \text{mailboat, packetboat.} \\
\text{2. } \subset [\text{Poste}] \ \text{mail.}
\]

On lexicographical grounds, and perhaps also on semantic grounds, a case can be made for separating these two senses into two distinct entries malle\textsuperscript{1} and malle\textsuperscript{2}. In a few instances the ideal equivalence [a + x] or [x + a] will exist:

\[
\text{moose n} \ \subset [\text{Zool}] \ \text{original m} \ \subset \\
\text{frost boil} \ \text{loc n} \ \subset [\text{Meteor}] \ \text{ventre-}
\text{de-boeuf m} \ \subset .
\]

3.8 One difficult problem which remains to be resolved is the question of how to deal with large numbers of anglicismes in the CF lexicon. We have already observed that in French Canada, a linguistic island in an English-speaking ocean, there is even more vigorous opposition to
English infiltration than in France itself and many lexical items generated in Quebec have found acceptance in IF to replace English borrowings. The fact remains, however, that large numbers of *anglicismes* continue to penetrate CF and are in everyday use, particularly in the spoken language. They may enter the language and remain structurally unchanged, or else undergo progressive Gallicization to the point where they are barely recognizable as borrowings. Faced with a decision on the status of such words, the lexicographer can choose between a number of options. He can omit them altogether; but in so doing he has virtually passed judgement on them and condemned them to obscurity as corrupt, adulterated language. Alternatively, he could enter them indiscriminately and prepare to receive the censure of normative organizations which will undoubtedly have opinions on the subject. A compromise would be to admit them accompanied by a register indicator, such as "abusive" which would enable the dictionary user to form his own judgement even though it would be already coloured by the lexicographer's prescriptive approach.

None of these choices is ideal and the final decision will be certain to attract censure from various quarters; but a reasonable policy would be to include a number of commonly employed *anglicismes* whose frequency of occurrence and appearance in written texts demands
some kind of explanation. At the risk of drawing criticism for prescriptiveness, the lexicographer may enter such words with full relevant details (grammar, pronunciation, field) but without supplying a translation. The latter can be achieved by giving, instead, a cross-reference to the normally acceptable term in IF or CF, as the case may be. This approach has the merit of enabling a reader, who may have been puzzled by finding a word in a context which prevents him from guessing its meaning and which he cannot find in standard dictionaries, to infer its meaning from the synonym suggested as a cross-reference. At the same time, the lexicographer can claim to have maintained a descriptive policy with regard to anglicismes without, so to speak, having passed final judgement on them. Such words might be entered thus:

\[ \text{créne} \text{ nf } \odot \text{ [Méc]} = \text{GRUE}. \]

The reader may not need to look further than the label [Méc] to comprehend the sense of the word, or he can refer to the entry grue to find its meaning (or to confirm a suspicion which context may have already suggested to him). Thus, while the lack of a direct translation may be interpreted as a kind of silent judgement on the word, its presence in the dictionary in some form at least acknowledges its existence.

A higher level of anglicismes are calques which in
some cases are totally acceptable because they reflect Canadian reality which cannot be catered for by IF. For example, Conseiller de la Reine (Queen's Counsel), la Colline parlementaire (Parliament Hill). Other calques have legitimate claims to recognition on the basis of frequency and long-standing usage; since they are neither total anglicismes nor yet full status canadianismes, they will require a slightly different lexicographical treatment. Their level of acceptability suggests that a translation may be given; at the same time, a cross-reference to the more usual IF term provides the reader with the means of forming his own opinion on the word which suits his particular purposes:

valise nf (Auto) trunk; cf. COFFRE, MALLE.

4. Conclusion

Such are the problems involved in the incorporation of regionalisms in a dictionary for Canadian users; similar situations would undoubtedly occur in other pairs of languages and require similar solutions. But there is no global answer to the question of regional variation in bilingual dictionaries and the phenomena must be treated as circumstances demand. The basic philosophy outlined above will probably prove satisfactory to the greatest number of dictionary users, but will probably fail to
appeal either to normative organizations or to linguists with extreme views on the purely descriptive goals of lexicography. In the long run, it is the convenience of his readers which must exert the greatest influence on the dictionary maker.
Notes to Chapter 15

1. The revised version of this dictionary (CD2) is currently being prepared at the University of Victoria's Canadian Bilingual Dictionary Project under the direction of J-P. Vinay. It will be three times the size of its more modest predecessor and has incorporated large numbers of CF lexical units and examples in an attempt to correct the kind of imbalance which characterized CD1.

2. See G. Dulong (1966), entry No. 12.

3. H. P. Angus (ed.), Canada and Her Great Neighbor, Toronto: 1938; refers to the travels of an Englishman, Thomas Ralph between 1832 and 1833; p. 88.

4. The GPFC, however, includes large numbers of words which are phonetic or orthographic variants of a single lexical unit. The effect of this is to distort the true statistical situation with respect to the number of Canadianismes it contains. See examples in 2.6.2 for the word chaudière.


6. English words of the configuration [sCV-] invariably pass into French without undergoing palatalization: snack, smart, switch, snob.

7. OED gives as citations: "He sneaked like a cock that hangs down his head when he's beaten" (1699); "Law . . . may allow . . . them to sneak--but law cannot wipe away the reproach of sneaking" (1845).


10. Presumably these words will eventually become archaismes in Canada also as the metric system replaces the old measures.

11. For a full discussion of borrowing in CE, see B. P. Harris (1975).
The examples are for the most part taken from Vinay (1973a) and Harris (1975).

Petit Robert, entry for "fiable."

B. P. Harris (1975), p. 221.

See below 3.8.
References for Chapter 15


1.0 Data processing in linguistic research is now an established fact, having found applications in fields as varied as stylistic analysis (shedding light on disputed authorship) and statistical linguistics (providing new insights into the phenomena of lexical distribution). Automated lexicography is an important part of this new approach to linguistic analysis and is epitomized by the colossal undertaking of the *Trésor de la langue française*¹ (TLF) whose computers were nourished on a diet of some 250 million words from literary and non-literary texts of the nineteenth and twentieth centuries, resulting in the compilation of the most complete lexicon ever produced. The consummation of this project will be a dictionary whose proportions dwarf even the monumental creation of Emile Littré: tome I absorbs 1,000 pages in advancing from *a* to *affiner*; the entry *abandonner* (v.tr.) alone embraces 8 pages, 16 columns and 1500 lines of single-spaced complex typography.

More significantly, the sophisticated electronic equipment of the TLF has enabled a variety of hitherto
impossible manipulations to be carried out. For example, statistical facts have been established on the basis of the analysis of large quantities of contemporary text and expressed in the form of relative and absolute frequency for each of the two centuries covered by the investigation; the description of morphological features has been systematized by the use of affix families; the true semantic impact of the word in context has been emphasized by the richness of the accompanying illustrative material. Yet the planning and execution of the entire project has been accomplished in a fraction of the time taken to produce the Oxford English Dictionary.

Two dictionary projects in the United States have also made extensive use of automated editing. The Rand Corporation’s *Random House Dictionary of the English Language* (RHD) was a pioneer in the field, using sophisticated programmes to sort and organize the source material and manipulate it up to its final form. The *Dictionary of American Regional English* (DARE) is a current venture which has improved on its predecessors by the flexibility of its retrieval mechanisms, which permit a variety of internal element combinations to be effected, and in the employment of on-line editing procedures.

Many other smaller projects have employed electronic means to carry out certain parts of the work which would otherwise remain in the arduous domain of manual labour.
For the most part, these have been monolingual undertakings all of which have their own specific problems. In the area of bilingual lexicography data processing seems to have made little impact, probably because no lead has been given by any major dictionary; there are numerous bilingual lexicons which are being compiled with some recourse to computers, but these are generally of limited scope and application and probably have little to offer in the way of a model to larger dictionaries involving pairs of national languages with more complex traditional material.

Bilingual lexicography differs from monolingual lexicography in a number of obvious but nevertheless significant ways. If it were merely a matter of using the computer as a gigantic storage and reproduction device, these differences would be of little consequence because it would only be necessary to examine the question of suitable coding procedures for input and of suitable format for output. Most lexicographers would not view this situation as an important advance over traditional methods of storing and publishing their material; in fact, they would probably refuse its viability on the grounds of expense and an undeniable hazardous element attendant on all electronically-controlled materials. On the other hand, the computer is potentially capable of speeding up production, adding new
dimensions of accuracy and completeness and rendering correction and future revisions simple in the extreme by comparison with manual methods. With such possibilities in mind, the unique kind of material of a bilingual dictionary has to be studied very carefully in order to determine the best methods of storing and manipulating the data.

In addition to the kind of material to be processed, a study must be made of the kinds of operations which it is hoped the machine will accomplish. Computer manipulation on the large scale demanded by lexicographical projects is very expensive and the prospective user must be fully aware of what he is getting involved with before launching the venture. Decisions must be made on the desired operations well in advance of the recording of the material in machine-readable form since once it has been committed to storage in a given structure it will be very difficult, if not impossible, and certainly very costly in terms of time and money, to make major alterations in order to render it compatible with a given kind of machine function. The input of many thousands of entry words, even by comparatively effort- less means on a sophisticated terminal, together with a complex network of hierarchically-organized data which belongs to each one, is an arduous and time-consuming task, and the sometimes burdensome features of an
expensive tool must be carefully weighed in the balance against the obvious advantages to be gained—speed, accuracy and manoeuvrability.

2.0 Input Hardware

Allied to the problems relating to the material itself is the question of the best hardware medium available to record and manipulate it. Not more than a few years ago the only hardware available was the key-punch machine or paper tape, both of them devices which were contrived basically for the input of numeric data and therefore ill-suited to complex linguistic material. A restricted keyboard of some 70 alphanumeric characters meant that language-oriented projects were obliged to work out numerous involved coding arrangements to account for diacritics and other symbols which were unavailable in standard form. Further restrictions imposed by the kinds of print chains then in use also meant that output was in the form of an aesthetically unattractive, capitals-only impression which carries with it a measure of confusion and a good deal of inconvenience. Linguistic applications, whether they be large-scale dictionaries or text-editing of extensive concordances, usually involve vast quantities of material and the choice was thus between hollerith cards, with the storage and cataloguing problems associated with the production
of literally millions of them, and paper tape, which is considerably more convenient to use and has a keyboard which allows a greater degree of flexibility, but which is notoriously error-prone and difficult to correct.

2.1 Another input device considered by some users is the optical character reader. Initially they were available in a form which permitted the scanning of material typed in a special script only, and since at some stage in operations a manuscript would have to be typed anyway, there seemed little to be gained by submitting the material to a further procedure which was not notable for its accuracy. More recently, OCR's have been developed which will interpret handwritten material placed within certain clearly defined limits, but these have proved to be even less accurate than the earlier versions.

2.2 Within the last five years a new development has come to the aid of all computer users in the form of the interactive typewriter terminal. Linguists who could see the benefits to be gained by a powerful tool but despaired of ever being able to make efficient use of it, called upon the industry to improve their technology in the interests of a potentially large group. The plea came first from a gathering of eminent scholars from around the world at the Colloque international sur la mécanisation des recherches lexicologiques held in Besançon in
1961, and more recently at the Conference on Lexicography in English which took place in New York in 1972. To judge from the recent proliferation of these devices it would appear that their request did not fall on deaf ears. Early models were like highly sophisticated typewriters, with a similar range of characters but permitting certain operations such as overstriking to correct errors. The development of replaceable elements meant that the range of characters could be extended to embrace, for example, the special writing systems created for exotic and unwritten languages. The most recent addition to the devices available is the interactive CRT video screen terminal which combines all the advantages of its predecessors with extraordinary speed and flexibility. Some models, like the Hewlett-Packard 2640, have achieved a degree of sophistication unheard of only three or four years ago. The HP2640 has a built-in memory of up to 8k, meaning that considerable quantities of material can be recorded and corrected in "local mode" without using costly machine-time which is invoked only when the data is considered to be in its final form; it also permits a number of operations to be programmed directly in the terminal and has scrolling and paging functions to enable movement within the data stored in its local memory. The latest version of this terminal, the HP2644A is available with twin cassette decks each of
which is capable of storing, locally, up to 120k (120,000 bytes) of information in addition to the normal memory. The tapes can be programmed from the terminal, in much the same way as regular computer tapes, with file marks, erasures and so on. It is even envisaged that the same kind of terminal will soon be available with a so-called "floppy disk" which will emulate all the functions of the regular computer disk.

2.3 Perhaps the most important feature of these terminals is their system of plug-in modules each of which corresponds to various kinds of enhancement, such as alternative character sets and line-drawing sets. The HP2644A, for example, is available with up to four character sets of 120 characters each which is more than most linguists are ever likely to require. There are standard character sets provided by the manufacturer (roman, mathematical, logical, etc.), or the user may generate his own using a dot matrix chart which is eventually programmed and electronically wired into the terminal. At the terminal the operator can change from one character set to another as required by a single key movement; similar keying enables strings of material to be underlined, written in inverse video (black on white), flashed on and off or any combination of these enhancements.
Perhaps the only disadvantage of this kind of terminal is the lack of hard copy which typewriter terminals of the kind produced by IBM and Deckwriter do have. A thermal printer to accompany the Hewlett-Packard models is available, but in the present stage of development it will only reproduce characters corresponding to the manufacturer's standard sets. However, an alternative to computer printout with its restricted range of characters is now possible with the advent of laser printers which can be programmed to accept any characters generated by similar methods to those used in the HP terminals. The quality of laser output and the incredible speed with which it is produced make its combination with CRT terminals of the Hewlett-Packard variety an unrivalled tool for linguistic research.

3.0 The Structure of Lexicographical Material

Since the lexicographer of the 1970's no longer has to concern himself with the means of recording and manipulating his data, he can turn his attention directly to the structuring of the material. L. Urdang (1966) emphasized the complexity of lexicographical material and the difficulties which face its conversion to machine-readable form:

Although one of the early problems in any computer program is getting the information into the computer, dictionaries create
special problems. They are, both in information content and typography, probably the most complicated books in the world (p. 31).

His remarks referred to monolingual lexicography, in particular, to the RHD project already mentioned, but they are no less applicable to bilingual lexicography. It may fairly be claimed that bilingual dictionaries include most of the essential material of a monolingual dictionary and a good deal more. With the exception of etymological information, which lies beyond the sphere of the translator, all the common features of a monolingual production are at least implied, if not treated in some detail, in any bilingual dictionary which purports to be more than a mere pocket edition. Accurate translations presuppose an intimate acquaintance with precise definitions in both languages. But the responsibilities of the bilingual lexicographer, as we observed in Part II, do not end there; he must ensure that the suggested equivalent form in the target language conforms in every respect to the source language original—with regard to semantic content, level of formality, field of reference and dialect type.

3.1 The modern bilingual dictionary is, in name at least, a two-way translation device, with each section constituting an independent and complete dictionary in
its own right. Yet the two sections are inextricably linked by an intricate network of translations and cross-references, each entry-word appearing as a translation equivalent in the opposite language and vice versa. The basic aim is to provide a framework within which the dictionary user is prevented from faltering when semantic equivalence diverges abruptly or when the graphic image creates ambiguity and deception. Thus, for example, there must be a system of explanatory rubrics to account for semantic extension into varied fields, warning devices to signal the presence of "faux amis" and an economic convention to describe morphological paradigms.

3.2 The unit of manipulation, in automated lexicography is the entry word. The material which it identifies and delineates is known as the "file," and the sum of these files constitutes the corpus of the dictionary. The arrangement of files within each section of a bilingual dictionary is identical, consisting of an alphabetical ordering determined by the shape of the entry word. This sequence is, in some dictionaries, interrupted by the insertion of morphologically dependent words whose status is signalled by their incorporation within the file of a main entry word. In the case of homographs, the usual conventions for marking semantic differences are observed (bat\textsuperscript{1}, bat\textsuperscript{2}). When French homographs are distinguished
only by punctuation, ordering appears to be arbitrary and constitutes a difficulty for automatic processing. For example, in CD1 the graphic series [PECHE] (and its derivatives) has the following order:

pêche\(^1\) (peach)
pêche\(^2\) (fishing)
péché (sin)
pécher\(^1\) (to sin)
pêcher\(^2\) (peach tree)
pêcher (to fish)
pêcheresse (female sinner)
pêcherie (fishery)
pêcheur, -ressse (sinner)
pêcheur, -se (fisherman, -woman)

The order established by the first three words (peach-fish-sin) is broken in the next group of their derivatives (sin-peach-fish) by the attachment of pêcher as a sub-entry of péché. We are led to conclude that it is the two accents of péché which relegate it to a position below pêche. Strict alphabetical ordering follows for all the remaining derivatives of the series, with pêcheur/pêcheur apparently decided by their feminine forms, in spite of the fact that this is a reversal of the base word sequence. Where two words are otherwise identical, therefore, their feminine forms, if any,
decide the order adopted. This cannot be taken as a strict rule, however, for we also find mat, -e (adj.) followed by mât (n.m.), and franc1, franche (adj.) before franc2 (n.m.). In fact, the sequence of this homographic series [PECHE] is that adopted by the Petit Robert (PR) and also by Harrap's New Standard French and English Dictionary (HAP3).

3.3 One manner in which the problem of entry ordering has been dealt with recently is morphological grouping of "word families." By this convention the words above are ordered as in Fig. 14.

```
pêche
péché

pécher
péchez, pêcheresse

pêcher (n.)
pêcher (v.)

pêche2

garde-pêche

pêcherie
pêcheur, -euse

martin-pêcheur
repêcher
repêchage
```

Fig. 14. Topical ordering of senses.
While this arrangement is orderly and reasonably logical, the dislocation of normal alphabetical sequence could constitute a look-up difficulty for the novice or non-native user.

From the point of view of automatic processing, a more formalized ordering system must be employed, because if the topical convention were adopted special programming procedures would have to be contrived. Existing sort programmes alphabetize on the basis of strict letter sequence and the systems designs must be devised with this in mind.

3.4 Of even greater significance for computer adaptation is the internal structure of the dictionary articles. In bilingual dictionaries this structure varies from the completely ad hoc, unexplained and therefore confusing classification of pocket-sized editions, to the well-illustrated, exhaustive and sometimes equally confusing arrangement of HAP3. Once again, the model appears to be PR which uses historical ordering, occasionally interrupted by logical sequence if this creates a more comprehensible file. In Table XIV the respective structures in PR, HAP3 and CD1 of the article [PEINE] are shown in "exploded" and abbreviated form. It can be seen that the PR monolingual model is followed by both bilingual dictionaries, with varying degrees of inclusiveness.
<table>
<thead>
<tr>
<th>TABLE XIV CORRESPONDENCE BETWEEN FILE STRUCTURES</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. sanction (châtiment, condamnation, pénalité).</td>
</tr>
<tr>
<td>2. [Jurid.] sanction de loi (blâme, confiscation,</td>
</tr>
<tr>
<td>amende, réprimande).</td>
</tr>
<tr>
<td>3. [Loc.] sous peine de.</td>
</tr>
<tr>
<td>4. [Rel.] peines de l'enfer (damnation).</td>
</tr>
<tr>
<td>1. souffrance morale (chagrin, tourment).</td>
</tr>
<tr>
<td>2. état psychologique (détresse, misère).</td>
</tr>
<tr>
<td>3. [Vieilli] en peine (inquiet); errer comme une</td>
</tr>
<tr>
<td>âme en peine.</td>
</tr>
<tr>
<td>1. activité qui fatigue (effort).</td>
</tr>
<tr>
<td>2. difficulté qui gêne (embarras, mal).</td>
</tr>
<tr>
<td>3. avec, sans/peine (difficilement/facilement.</td>
</tr>
<tr>
<td>4. à peine (presque pas).</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>1</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>2</td>
</tr>
<tr>
<td>3</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>4</td>
</tr>
</tbody>
</table>

**TABLE XIV Continued**
(Although CD1 seems to suffer somewhat by comparison with HAP3, it should be borne in mind that it is essentially a desk dictionary--a "concise edition.") The computer data file must reflect this structure, using similar conventions to indicate semantic demarcation, extension, subject field, syntagmatic expressions and so on.

3.5 To the uninitiated, and perhaps even to the sophisticated user, the structure of bilingual dictionary articles may appear to be arbitrary and often (as in the case of the long and rambling entries of HAP2) unnecessarily confusing. Each entry comprises a variety of information, depending on the kind and function of the word, and since no two entries seem alike the task of compiling a file structure which will be unique for each entry seems to be well-nigh impossible. In some dictionaries the arrangement of translation equivalents follows no set pattern, and even in those dictionaries which make semantic divisions there is often a lack of consistency in its treatment. Added to these difficulties are the continually varying type fonts and frequent occurrences of special lexicographical symbols.

3.6 On closer analysis, however, the material which makes up the entries of a bilingual dictionary can be broken down into a relatively simple and repetitive framework. The general features of each entry are as
follows:

1. **Entry word:** this is the canonical form of the abstraction which represents all the possible paradigmatic realizations of the word in context. It must be a readily identifiable unit, unique and unlike any other entry, for it is the address for all the information provided about the word.

2. **Morphological variations:** these are usually provided only in cases where irregularities occur or where there is some variation from normal features (such as verb paradigms) usually described in the preface.

3. **Pronunciation:** appears as the phonetic or phonemic representation of the canonical form of the entry, but sometimes includes variations under 2.

4. **Grammar code:** normally an abbreviated description of the part of speech.

5. **Semantic level:** numerical subdivisions marking off independent senses of the same word.

6. **Field labels:** these are of three basic kinds—dialect labels, domain labels and register labels.

7. **Translation equivalents:** one or more lexical units linked directly to the entry word through 5 and 6.

8. **Examples:** syntagms, "locutions figées" and contextual phrases or sentences containing variable forms of the entry word.
9. **Usage notes**: parenthetical information which may relate to any of the preceding elements and be placed accordingly.

10. **Cross-references**: providing semantic links with other files or assisting the user to specify the equivalence of the word he is looking up.

3.7 Some of these elements have fixed positions within the file. For example, the entry word is always first and always followed by morphological run-ons, pronunciation and grammar, in that order. Semantic subdivisions may or may not be present, depending on the complexity of the entry; they always follow grammar and precede field labels if present. Field labels, if present, always immediately precede translations; when all three kinds are present they normally have the order dialect-domain-register. Translations are the last element of the basic structure; anything which follows is consequent on the preceding structure. Thus, for example, there may be a cross-reference drawing attention to another lexical unit related to the entry word and its translation; this is one of the variable elements because it may be placed immediately after the translation or after the translation of an example, depending on which of these it is related to. Examples follow immediately after translations of the entry word, although an example may
follow grammar or field labels, if present, in cases where the entry word has no existence as an independent unit but occurs only in set phrases; for example:

Canossa n [Loc]: venir à C., to eat crow.

Usage notes may appear in any one of a number of places: directly following part of speech indication, preceding or following a translation or examples, and so on; its position clearly depends upon what element the usage refers to. Finally, the entire range of possible orderings may be repeated one or more times under separate semantic subdivisions which, from one point of view, constitute separate "words."

4.0 File Structure Analysis

The breakdown of the material of bilingual dictionary entries into the ten basic elements we have just discussed provides us with a working model on which to base the computer files. It is at this point that decisions relating to the kinds of operation likely to be required must be made by the lexicographer. If he wishes, for example, to have access to all ten kinds of information which may appear within a dictionary entry, then the files must be structured in such a way that these elements have a retrievable address. Let us now examine in greater detail each of these ten elements in
order to determine the likely uses to which they could be put in a machine-readable corpus.

4.1 The **entry word** needs little discussion in terms of accessibility, since it is self-evident that it is the flag marking the contents of the entire article. In computer files it will likely be the unique address in storage of a self-contained unit of data.

4.2 **Morphological variants** seem logically to belong to the entry word; in isolation they have little significance since they are usually not applicable to other words; they may take the form of irregularities of the type *mouse/mice, bite/bit/bitten, faux/fausse*, or merely constitute added information about the entry: *Seychelles (the), Epiphanie (l').* The only decision that would have to be made about this kind of material is whether it should be distinguished from other entries made up of multiple elements: *stock car, pomme de terre, dungeness crab.* One reason for making such a distinction might be a statistical study in which the computer is called upon to count as "words" examples of the kind just quoted as opposed to those in which the added material does not form part of the real or abstract graphic image. In an investigation on word length, for example, it would be misleading if the indication of a morphological variation such as that of the following entry were counted in the
calculation:

hit [-tt-] vtr frapper, ...

A very simple procedure, the variation of spacing between entry and run-on material, can account for such distinctions and at the same time render morphological information accessible to the researcher who requires it.

4.3 **Pronunciation** might also be considered an offshoot of the entry word with no call for independent access, but a number of objectives might persuade the lexicographer to bestow a separate address even upon this element. One of the more obvious motives would be ability to obtain printouts of entry word and pronunciation for proofreading purposes or for use in statistical studies on morphophonology. Another less evident objective would be as a part of an automatic phonetic generation programme of the type discussed below (1.5).

4.4 **Grammar codes** should be considered a basic addressable unit which would permit studies of word groups according to part of speech function, gender or number. Depending on the degree to which exponents are employed to distinguish between homographs, identification of pairs or groups which differ only with regard to grammatical function would depend on this feature remaining accessible.
4.5 Semantic level indicators serve as partitions which separate discrete senses of the same word into individual units each of which is equivalent to a simple entry file, repeating some of the data relating to the entry word and introducing new material requiring different definition or translation. Programmes which must search the entire contents of a complex dictionary article comprised of many different senses may well hinge on these indicators. They have no intrinsic value of their own as retrievable elements, but will be essential to the efficient structuring of the data base files.

4.6 Field labels: from the point of view of recoverability these elements are probably the most important in the file. One of the most useful applications of a machine-readable corpus of lexicographical material is the interrogation of files in order to extract batches of related items. For example, suitably structured files can be scanned by a programme in order to retrieve predetermined information such as chemical terminology, vulgarisms or combinations of elements like informal legal language. The programme can be designed to produce the entire contents of a file containing one or more of the interrogation elements or to furnish only the relevant information for that particular field. This kind of flexibility has a number of important byproducts:
one is the ability to maintain control over the semantic content of the dictionary and to prevent loss of internal equilibrium due to excessive emphasis having been placed on one subject area at the expense of another; a more distant goal might be the eventual publication of mini-glossaries for specialist translators; another is the organization of the entire contents of the dictionary on an analogical basis in order to facilitate rapid access to the material as a "Word Bank."

4.7 **Translations** are the culmination of the internal ramifications of dictionary entries; in the physical dictionary they are the solution to a translation riddle which the user resolves by following a number of fixed conventional paths beginning with the entry word. They must be uniquely accessible in the computer files for a number of important reasons. To begin with, as the logical conclusion of a linear progression through the file they constitute an element of equal importance and status to the entry word itself. In a programme of automatic inversion, such as that outlined below (8.0), their unique identity is vital to the success of the operation. For statistical studies comparing words in one language with their translation in another it will be essential to be able to gain rapid and easy access to them.
4.8 **Examples** are always directly related to the entry word and sometimes also to one or more of the preceding translations; in addition, they descend from field labels of various kinds, either as an extension to the subject or register or dialect label already indicated for the translations or else as entirely new headings. Accordingly, it will be important for them to be retrievable elements, both as part of an interrogation for related materials and also as important constituents of an inversion programme.

4.9 **Usage notes** are probably of limited immediate use to a lexicographer as retrievable elements, although it is not inconceivable that an investigator would want to study relationships between words with similar grammatical or other peculiarities as described under this heading.

4.10 **Cross-references** in a bilingual dictionary usually serve a more important task than the mere indication of synonyms. Reference to other terms may be designed to further elucidate a translation difficulty, to provide information on a grammatical peculiarity and so on. Access to these elements will permit their accuracy to be verified quickly and may shed light on otherwise concealed relationships of importance to comparative research.
5.0 **Computer Data**

**File Structure**

When the dictionary-maker has applied his knowledge of the content of lexicography in order to produce a summary of the kind we have just outlined, including a statement of the manner in which he hopes to manipulate his materials, the next step in his endeavour to come to grips with computer technology is to consult with a programmer, preferably one whose skills include some acquaintance with linguistic applications. Together they will study the analysis of the lexicographical material and agree on the most appropriate methods of attacking the various difficulties which invariably surface.

5.1 Assuming that the question of the best input hardware for the project has been resolved, attention must now be given to the software questions relating to the internal structure of the lexicographical information. We have seen that the contents of a dictionary can be subdivided into relatively simple and repetitive categories. Each entry constitutes a separate file of information which is both a self-contained unit and also enters into certain relations of a semantic nature with other files. The totality of these files makes up the machine-readable corpus which in turn has a structure permitting rapid access and the application of a number of sorting routines. This corpus will usually, although
not necessarily, reflect the structure of the dictionary: strictly alphabetical, quasi-alphabetical, alphabetical-morphological, analogical, etc., and while any of these arrangements can be catered for it is preferable to opt for one which will require the fewest special programmes for sorting and merging operations. If the lexicographer is prepared to sacrifice what he probably considers to be the most pedagogically sound organization of his entries—alphabetical order interrupted by morphologically-related entries (usually referred to as "word family" grouping)—to a strictly alphabetical arrangement, he will be able to capitalize on existing computer programmes which sort, merge and alphabetize on this basis. We have seen (Part II, Chapter 11) that there are good reasons for maintaining rigidly alphabetical order over other arrangements, and that modern practice, especially in monolingual English lexicography, seems to favour this approach. If he decides, in spite of the advantages, to use some other system, he must be prepared to add a measure of extra complexity to the overall programming of his materials.

5.2 The structure of each entry, which we will refer to as the data file in order to distinguish it from the physical dictionary entry, will be based on the analysis given in 1.3. Various approaches to the data file
construction are possible, but whatever system is chosen, it must be designed in such a way that it will facilitate the manipulations which the lexicographer has in mind for his material. One of the most efficient structures which has been designed so far (and the only one which has been described in literature generally available to linguists) is the "catalog" system conceived by the Rand Corporation and used in the automatic processing of the RHD. Each data file was organized as a hierarchical tree-like structure with the entry word as the highest level node of the tree and the various elements of the file arranged as dependent data classes. The dictionary material was broken down into seven categories corresponding to the seven main elements of information in a monolingual dictionary: entry, definition, variants, etymology, run-on entries, illustrations and other information (such as usage notes, synonyms, cross-references, etc.); grammar and pronunciation formed part of the entry word material.

At the time when the Rand system was developed, software limitations dictated that a fixed format be used in order to permit access to the various levels of the data file. The main disadvantage of this arrangement is that all files will have to reserve space equivalent to the requirements of the most complicated of them, even though the majority would probably be of quite modest proportions. The alternative to this is to have
programmes which can account for the individuality of each file, a physical impossibility in a corpus containing some 100,000 data files. Systems engineers at the Rand corporation solved this problem by means of a compromise which consisted of a coded description of the contents of each file and which could be interpreted before the file is read. The description was referred to as a "map" and consisted of a miniature representation of the data file in the form of a key to its contents and an indication of the size of each of the discrete data classes; this arrangement allows gaps in the file where information for a given data class is not present, but at the same time space is not reserved for the non-existent material.

5.3 Fortunately, recent developments in software techniques have eliminated these difficulties and the comparatively cumbersome solutions which they engendered. Built-in cataloguing systems and variable-length format enable data files to be constructed without the requirement that their individuality be specially accounted for. One approach is to devise a simple coding system for each unique component of the data file to which access is required; for a bilingual dictionary the coding might be organized as in Table XV where the sequential numbering corresponds to a linear order which
<table>
<thead>
<tr>
<th>Field</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>main entry word</td>
</tr>
<tr>
<td>11</td>
<td>sub-entry</td>
</tr>
<tr>
<td>12</td>
<td>extra information pertaining to 10 or 11</td>
</tr>
<tr>
<td>15</td>
<td>pronunciation</td>
</tr>
<tr>
<td>17</td>
<td>grammar code</td>
</tr>
<tr>
<td>18</td>
<td>usage note (relating to entire entry)</td>
</tr>
<tr>
<td>21</td>
<td>semantic level indicator</td>
</tr>
<tr>
<td>28</td>
<td>usage note (relating to this semantic level only)</td>
</tr>
<tr>
<td>31</td>
<td>dialect label</td>
</tr>
<tr>
<td>32</td>
<td>field label (always relating to 10 or 11)</td>
</tr>
<tr>
<td>33</td>
<td>register label</td>
</tr>
<tr>
<td>35</td>
<td>translation(s) of 10 (or 11)</td>
</tr>
<tr>
<td>37</td>
<td>cross-reference (referring to 10, 11 or 35)</td>
</tr>
<tr>
<td>38</td>
<td>usage note (relating to 35)</td>
</tr>
<tr>
<td>41</td>
<td>dialect label</td>
</tr>
<tr>
<td>42</td>
<td>field label (always relating to following example)</td>
</tr>
<tr>
<td>43</td>
<td>register label</td>
</tr>
<tr>
<td>44</td>
<td>example containing a form of 10 or 11</td>
</tr>
<tr>
<td>45</td>
<td>translation of the example</td>
</tr>
<tr>
<td>47</td>
<td>cross-reference (referring to 44 or 45)</td>
</tr>
<tr>
<td>48</td>
<td>usage note (referring to 44 or 45)</td>
</tr>
<tr>
<td>49</td>
<td>problem zone (non-printed area for recording queries, etc.)</td>
</tr>
<tr>
<td>50</td>
<td>end of file</td>
</tr>
</tbody>
</table>
is always maintained but which permits looping so that
certain kinds of information can be repeated any number
of times depending on the complexity of the article for
that entry. Each data file will contain any or all of
the coded elements with the minimum requirement being
that there be a 10 (or 11) followed by a 17 and either a
34, 35, 37, or 38 and concluding with a 50.

The application of this coding system to actual
dictionary entries is illustrated in Table XVI which
compares the typographical entry with the computer data
file of a very simple entry and a moderately complex one.9

It can be seen that the linear order corresponding to the
first entry is coded as 10-15-17-37-50 while that of the
entry "nature" is 10-15-17-21-32-35-44-45-45-43-44-45-
41-44-45-21-32-35-44-45-44-45-50. The latter contains a
number of loops illustrating the repetition of various
kinds of elements in the data file.

5.4 This particular structure has a number of important
features, one of which is the ability of the system to
act as a generator of varying type fonts, a procedure
which can be built in with a view to its eventual
employment in automatic photocomposition for the final
publication of the dictionary. Each of the element codes
may be organized to assume a dual role: the identifica-
tion of data file elements and the flagging of a variety
TABLE XVI  CORRESPONDENCE BETWEEN

TYPOGRAPHICAL ENTRY AND COMPUTER DATA FILE

<table>
<thead>
<tr>
<th>Typographical Entry</th>
<th>Computer Data File</th>
</tr>
</thead>
<tbody>
<tr>
<td>sat [sat] v &amp; pp cf sit</td>
<td>10 sat</td>
</tr>
<tr>
<td>15 [sat]</td>
<td>17 v &amp; pp</td>
</tr>
<tr>
<td>37 cf sit</td>
<td>50</td>
</tr>
<tr>
<td>nature ['nertʃə] n.l. [Bot, Zool, etc] nature f, la Nature; the laws of ~, les lois de la Nature; the ~ of things, l'ordre m des choses; [Fam] to answer the call of ~, faire ses nécessités; [Psych] nature f, naturel m: by ~, par tempérament; good ~, bon naturel.</td>
<td>10 nature</td>
</tr>
<tr>
<td>15 ['nertʃə]</td>
<td>17 n</td>
</tr>
<tr>
<td>21 l.</td>
<td>32 [Bot, Zool, etc]</td>
</tr>
<tr>
<td>35 nature f, la Nature: the laws of ~, les lois de la Nature; the ~ of things, l'ordre des choses; [Fam] to answer the call of ~, faire ses nécessités; [Psych] nature f, naturel m: by ~, par tempérament; good ~, bon naturel.</td>
<td>44 the laws of ~, les lois de la Nature; the ~ of things, l'ordre des choses; [Fam] to answer the call of ~, faire ses nécessités; [Psych] nature f, naturel m: by ~, par tempérament; good ~, bon naturel.</td>
</tr>
</tbody>
</table>
of typographical fonts. For example, 10 and 11 would indicate boldface for entry words, 17 italics for grammar codes, 21 boldface for semantic level numbers, 37 and 47 small caps for cross-references, and so on.

5.5 Another feature is the flexibility which it allows in a number of important editorial tasks such as addition of new material or correction of existing material, automatic generation of phonetic transcription (6.0), retrieval of designated kinds of elements from the files (7.0), merging of certain kinds of information; automatic inversion of entries (8.0), automatic photocomposition, and so on. If it were discovered, for example, that the entry "nature" (described in Table 3) lacked the example nature study (étude des sciences naturelles), this could be added to the existing entry by calling up the file in which it should appear and merging it as a 44-45 pair following the element les lois de la nature. Using a CRT interactive video terminal of the type described above, this kind of operation is accomplished easily and rapidly. In a case which involved a change in field label the appropriate modifications would have to be carried out to the entire data file. Let us suppose, for example, that it was found necessary to add another translation to the entry "nature" invoking the field label [Food]; the editor would first have to decide whether this would
constitute a new semantic level, in which case there would be an additional 21 to be incorporated; next he would have to determine whether this semantic division should follow, precede or interrupt the existing ones; finally, he would have to call up the file and record the new material in its designated position. Complicated as the operation may seem, it is no more so than it would be under manual conditions where a search would have to be made by hand among the file cards or in the typed manuscript and the additions made by hand. There is a further advantage to the alteration or addition of computer-based material: it can be carried out right up to the very moment of publication (in the case of a photocomposed edition) and does not appear as a corrected or overstruck item on a manuscript which adds to illegibility and confusion; a computerized alteration simply removes the incorrect string and replaces it by the correct one.

6.0 Automatic Generation of Phonetic Transcription

One of the most difficult tasks in lexicography is the typing or typesetting of the complex typographical material, of which phonetic transcription poses the greatest problem for untrained typists or typographers. It would therefore be of immense practical value, to say nothing of theoretical interest, if it were possible to
develop a computer programme which would take graphic input and transform it automatically to phonetic output. A number of studies have been done in this area, for the most part aiming to facilitate phonological research. One practical application of orthographic to phonetic transformation is speech synthesis from the written word, a procedure which requires a phonetic input in order to make the transformation to the spoken word; the most obvious use for a device which will accomplish this is in providing "talking books" for the blind. In spite of a good deal of research on the question of speech synthesizers, however, the problems involved in the transformation process from the graphic to the phonetic image have remained largely unresolved. Among various investigations seeking to shed light on phonological problems by this method, the most highly developed is the procedure described by B. Pratt and G. Silva (1967) for the transcription of French poetry.

6.1 None of the results of this kind of research have yet been applied to lexicography even though it is clear that the use of an automatic transformation system would save many thousands of man-hours in the arduous task of transcribing and typing by manual methods. The reason is probably that no lexicographical undertaking has ventured far enough into the somewhat uncertain sea of computer
technology to take advantage of the situation, and also that the kind of hardware needed to accomplish automatic transcription was, until comparatively recently, unavailable. For bilingual lexicography there is an added difficulty: given the differences in the phonological systems of any pair of languages, it is unlikely that a single transcription algorithm could be evolved which would work for both languages; even with modifications; it will therefore be necessary, in bilingual lexicography, to have access to two separate transcription procedures.

Lexicography operates at the level of the discrete "word," or more precisely, the lexical unit since entries may comprise more than a single "word." Therefore, the phonetic or phonemic representation in a dictionary, as we have already observed (Part II, Chapter 11), is essentially an abstraction which is seldom realized in context where other factors, such as prosodic features and elision, will come into play. This is fortunate for a project aiming to employ a system of automatic transcription because the simpler rules of abstract phonology mean the programming required will be accordingly less complicated. In a dictionary of French and English the two procedures will be of uneven complexity because the accentual phenomena of English are difficult to reduce to a reasonably compact set of
consistently applied rules, whereas those of French are so regular that it is probably not even necessary to indicate them in the phonetic transcription. Thus, while it is not inconceivable that an automatic transcription algorithm could be written for English, including those prosodic features applicable at the level of the non-contextual lexical unit, the discussion which follows centres round the simpler algorithm for French.

6.2 The main problem of an orthographic-to-phonetic conversion for any language is the fact that a number of different spellings may give rise to a single sound or group of sounds, and that conversely the same spelling may be pronounced in a variety of ways; one needs only to quote the *-ough* words of English for evidence of the latter. Similarly, some individual letters will almost invariably be represented by the same phonetic symbol (*f*, *l*, *q*, *r*, *v* and *z* in French) while others will have a number of environmentally-controlled realizations (*m* may give [m] or a nasal vowel in French; *s* may give [s] or [z], and so on). For French at least, even though it may appear unbelievable to a non-native learner of the language, the various possible transformations can be reduced to a closed series which can be tabulated and coded for automatic processing.

6.3 To a very large extent the phonetic realizations of
French words are dependent on syllable structure and therefore syllabification is the first part of the algorithm for French transcription. Fortunately, this aspect of the problem has been largely resolved by studies on automatic hyphenation and researchers need only apply their efforts to tabulating the outcome of a syllable breakdown. The rules for French syllabification are summarized by Silva (1974) as follows:

if n = number of consonants between vowels (in the orthographic sense), then n - 1 belong to the preceding syllable, the nth consonant introducing a new syllable, with the exception of some digraphs such as bl, cl, fl, gl, gn, th which are inseparable and always introduce a new syllable (p. 259).

There are, of course, other exceptions to this general statement and some allowance has to be made for differences between poetry and prose and between continuous context and individual word.14

6.4 The flow-chart in Table XVII gives an example of the mechanical procedures through which a word goes in order to generate a phonetic output from a graphic input. This is basically the programme devised by Pratt and Silva (1967) for French poetry with adaptations for the particular needs of CD2. The more accurate and fully documented syllabification procedures developed by Moreau (1964) replaced the somewhat restricted programme put
TABLE XVII

Flowchart for Automatic Transcription Cycles

1. copy lexical units from all 10's, 11's into selected file
2. submit data to syllabification programme
3. test word for 4-character syllables
   - YES: process next word
   - NO: test word for 3-character syllables
4. test word for 3-character syllables
   - YES: match against acceptable combinations in catalogue
   - NO: test word for 2-character syllables
5. test word for 2-character syllables
   - YES: apply codes for special conditions
   - NO: test word for context-sensitive constraints
6. test word for context-sensitive constraints
   - YES: apply codes and convert to IPA
   - NO:
6.5 The general algorithm uses the data file structure described above (1.4) in order to obtain input from 10 or 11 entries and generate output in the empty 15 slot. It can be seen that there are several cycles through which each lexical unit proceeds. The first is the syllabification cycle which reads the lexical unit and applies the rules to break it down into the requisite divisions. In the second cycle, the syllable elements are compared to the tables in order to determine whether the particular consonant distribution in each case occasions any special treatment or whether the programme can proceed to the final cycle which is the assignment of the predetermined codes corresponding to phonetic items in IPA. The ingenious programme devised by Pratt and Silva calls for a five-phase scan of the syllables, beginning with groups of four letters, then three, two and finally one. During each phase the groups of characters of the lexical unit in question are compared with the tabulated permissible combinations and if a match is found the appropriate numerical code is assigned. If no match is possible, smaller combinations are attempted until all possibilities have been accounted for. The final cycle takes the coded transcription and
6.6  The CD2 adaptation of Pratt and Silva's algorithm for automatic phonetic transcription, like its predecessor, is far from a perfect product. There are significant numbers of exceptions which cannot be handled by the existing programme which would become unwieldy if altered in order to cover these cases. Some typical errors are oignon [waːnɔ], instead of [ɔ̃pɔ], fils [fil] instead of [fis], us [y] instead of [ys], gentil [kɑ̃til] instead of [kɑ̃ti], vis [vi] instead of [vis], ville [vij] instead of [vil], tabac [tabak] instead of [taba], événement [evenmɔ̃], sourcil [sursil] instead of [sursi].

In common with Pratt and Silva, the CD2 programme cannot predict the aspirate of words with initial ħ, but the occurrence of these is restricted enough that it does not constitute a serious setback; the small number which must be aspirated can be easily so marked by post-editing. The CD2 programme likewise could not handle the numerous non-French words, such as those of place names, which are given a transcription according to the rules of French orthoepy. Thus la baie James is rendered [la bɛ ʒam] instead of [la bɛ dʒeːmz], l'Iowa as [liˈoːva] instead of [laɪˈoːwa], and so on. Similarly, the CD2 programme encounters considerable numbers of canadienismes which are transcribed as if they were standard French words.
(since the overall programme was designed for International French and based on Pierre Léon's *Aide-Mémoire d'orthographe*, Besançon, 1963). These lexical units, as well as the exceptions just described, must be treated by post-editing.

6.7 Although at the present stage of development the CD2 automatic phonetic transcription programmes necessitates a considerable amount of post-editing in order to correct these anomalies, the general procedure is efficient enough to generate thousands of completely correct pronunciations in a minute fraction of the time taken to go through the manual procedures of writing them by hand on file cards and then having them typed in the manuscript and eventually typeset for publication. Even the post-editing can be carried out quickly on a CRT video display terminal, so that the overall result is the automatic generation of material at the cost of a few minutes' worth of CPU time and with a saving of hundreds of man-hours and the human error which goes with this kind of manual labour. The arguments in favour of using such a procedure seem irrefutable.

7.0 **Data Retrieval**

One of the motivations for a carefully structured file system, as we observed above, is the need to maintain control over the contents of the computer-based
data files. If the compilation of the physical "fiches" for the dictionary was organized around subject headings (Biology, Sociology, Measurement, etc.), a procedure which, as we have seen in Chapter 10 of Part II, permits a logical and balanced progression towards semantic closure, then this element of control will eventually have to be surrendered in favour of merging and alphabetizing all the material in preparation for final editing of the manuscript. Once alphabetized and edited by manual methods, lexicographical materials are no longer analysable even by the somewhat crude methods of inspection or counting of file cards. By contrast, a file structure composed with retrieval mechanisms in mind will permit a variety of scanning operations to be carried out which will provide the editor with constant information on the internal constitution of his dictionary. Not only will this information be available to him at any time, and literally in seconds, but it will be of a kind impossible to obtain by any other procedure.

7.1 The file structure described in 1.4 lends itself well to programmes for retrieval of selected elements. From the schematic view (presented in Table XVIII) of a typical data file, it can be seen that the kinds of item likely to be of interest to a dictionary editor have been located at strategic points in the hierarchically-
TABLE XVIII. Schematic view of data file containing 49 elements of accessible information.
organized structure; the most obvious elements, apart from the entry word itself, are the various field labels and the data dependent upon them. A scanning operation can be designed to take the most direct path through each file in order to reach the particular constituent sought by the retrieval programme and to report on its contents. Thus, for example, a command to extract from the files all information on ornithological terminology would read all 32's and 42's, which are the elements containing this kind of information in the form of an abbreviated string roughly corresponding to the labels of a conventional dictionary, and, in the case of a positive result, would proceed to collect all the appropriate data from the file and store it for eventual printout listing. Each information file (as opposed to the data file which represents just one dictionary-type entry) contains about 3,000 lines of text, or approximately 350 individual data files, stored in disk packs available for continuous and instantaneous consultation. The entire retrieval operation, scanning, identification, transfer to storage and printout of the results from one of these information files takes only seconds to accomplish; it would take an experienced reader several hours to carry out the same task.

As an experiment, and in order to obtain materials used in Chapter 15 ("Regionalisms in Bilingual Lexico-
graphy"), the nine information files of CD2 English Canadianisms, stored as a separate unit on a temporary basis, were submitted to a retrieval programme known as "Keylabel" in order to obtain Canadianisms relating to winter phenomena and filed under the general heading of "meteorology" (abbreviated as Meteor). The scanning cycle read the 2,566 entries, comprising 22,910 lines of data, and stored the requisite information ready for printout with a consumption of 16.77 minutes of CPU time. The output consisted of 72 complete data files with a total of 606 lines which took 1 min. 42:21 secs. and cost 27 cents to print. These statistics speak for themselves.

7.3 Such a programme is not, however, obtained without cost because there are a number of thorny problems to be solved if the procedure is to function efficiently (in terms of time taken to execute the operation), and accurately (with regard to the fitness of the data which it produces). There is a corollary to any computer programme which states, in effect, that material can only be processed effectively if it is compiled in such a way that the computer can analyse it in a complete absence of ambiguity. In the case of the retrieval programme "Keylabel," it is not sufficient that the lexicographical information be carefully stored in the discrete levels of the file structure which will ensure its accessibility; a
good deal of attention must be given to the seemingly simple matter of the typographical layout of that information, in particular to its punctuation logic.

7.4 Published dictionaries adopt certain conventions with regard to punctuation: commas mark the end of an example and the beginning of its translation or mark off synonymous translations; colons introduce examples; semi-colons mark perceptible semantic divisions between otherwise synonymous translations or separate one translation of an example from a following example, and so on. The list is quite long and not necessarily consistent, although it is possible to reduce it to manageable proportions and apply the consistency essential to automatic processing. It is also possible to maintain the kind of punctuation usual in the typographical form and yet at the same time bend to the demands of the machine that there be absolutely no ambiguity. Furthermore, what may at first sight appear to be excessive servitude imposed on dictionary editors turns out in practice to be a salutary lesson in consistency and accuracy.

The following are some punctuation conventions which have been found appropriate for machine processing and which at the same time do not violate the usual typographical requirements; the only stipulation is that
they be employed with scrupulous consistency to prevent breakdowns in the retrieval programme (and later, also, the inversion programme). The punctuation period (as opposed to the abbreviation period) is reserved for end of file and for marking separations between major semantic subdivisions within the data file; it will therefore be limited to use after 35-elements, when these are not followed by other information, or after 37's, 38's, 45's, 47's and 48's. Commas indicate divisions between synonymous translations (35's and 45's) and have the important function of separating examples (44's) from their translations (45's). Semi-colons separate translations which are related but not synonymous, a condition which is sometimes indicated by preceding explanatory rubrics. They also mark the end of a translation which may be followed by a related cross-reference or usage note and, in the case of 45's only, may signal the end of that example and its translation(s) and the beginning of a new example and its translation(s). Colons are used to signal the introduction of examples; if they are related to a preceding translation or series of translations, the colon is placed immediately after the 35 elements; if there is no preceding 35, as is the case for some nonce words and others which only occur in set phrases not deemed to be eligible for entry status, the colon occurs after the immediately preceding element
(which may be a field label, a semantic division number or a grammar label). If a string ends with an exclamation mark or question mark, it must be terminated by the punctuation appropriate to the element according to the foregoing rules.

7.5 Another important item in a file structure which is intended to make certain elements uniquely accessible in retrieval operations is the physical placement of labels for dialect, subject field and register. The retrieval programme must be furnished with unequivocal instructions with regard to the sphere of influence of these labels so that only the requisite information is extracted. This requirement again imposes certain restraints on the format of the lexicographical material, but once again the limitations only serve to enhance the final product by removing ambiguity and enforcing consistency. In many dictionaries a specifying label is introduced and followed by translations which relate directly to it; yet when further translations or examples follow, without a new label being introduced, it is only the inconsistently applied punctuation which informs the reader that the application of the first label has been cancelled. In the computer-based data file only the period, as we have seen, terminates the semantic implications and the context of what has preceded. However, the period can
only be introduced at the end of major semantic subdivisions or, at the end of the entire entry and this means that more attention has to be given to the logical presentation of the material than is often the case in dictionaries produced by conventional methods. By way of example, let us now examine an actual dictionary entry (from OXF) in its usual typographical form and determine the modifications which will have to be made:

noce s.f. 1. (sing. or pl.) Wedding, nuptials, wedding-feast; 2. (sing.) drinking-bout, revelry, carousal, orgy; gâteau de ~ s, wedding cake; il n'avait été à pareille ~, he had never had such a time of it; elle avait épousé X en premières ~ s, she was X's wife by a first marriage; ~ s d'argent, silver wedding; ~ s d'or, de diamant, golden, diamond wedding; n'être pas à la ~, to have had a bad time of it; faire la ~, to go on the spree.

7.6 There are a number of serious flaws in the actual layout of this entry, but since they have already been discussed in some detail (in Part II, Chapter 11), we shall devote our attention here only to those aspects of immediate concern to a retrieval programme. Although they are not specifically named, there are two major semantic subdivisions (indicated by the numerals): the religious and sociological sense (1) and the informal extensions (2). If the entry ended at the translation "orgy" there would be no difficulty, but the examples which follow (and which logically belong to sense (2) are
a hopeless mixture of both senses; we may represent the muddled arrangement thus, using the digits referring to the correct sense: 1-2-1-1-1-2-2. If specifying labels had been employed in the article and the material recorded in this order, a retrieval programme would produce some laughable results. This penchant of many lexicographers to leave all the thinking to the dictionary user (and it is by no means certain that all users are sufficiently sophisticated in the linguistic sense to make all the decisions thus left to them) cannot be emulated by their data-oriented counterparts, because the machine can only "think" along the somewhat narrow lines dictated to it by the programmer. Yet, in the process of reformulating the above material to make it compatible with the exigencies of the computer, a result is obtained which is nothing if not vastly superior to the original product:

noce s.f. 1. [Rel, Soc] (sing. or pl.) wedding, nuptials, wedding-feast; gâteau de ~ s, wedding-cake; elle avait épousé X en premières ~ s, she was X's wife by a first marriage; ~ s d'argent, silver wedding; ~ s/d'or, de diamant, /golden, diamond/wedding. 2. [Fam] (sing.) drinking-bout, revelry, carousel, orgy: il n'avait jamais été à pareille ~, he had never had such a time, of it; n'être pas à la ~, to have had a bad time of it; faire la ~, to go on the spree.

Various other refinements could still be applied to this
entry, and some will be discussed below under the heading of inversion, but even as it stands it will comply to the logic of the retrieval programme and at the same time make the consultation by even the most naive dictionary user a far simpler and surer operation.

7.7 The flowchart in Table XIX depicts the main stages in the progressive scanning of data files in order to retrieve designated materials. For example, the procedure referred to above which obtained printouts of all meteorological information was accomplished by calling up the "Keylabel" programme which asks first for the name of the information file to be scanned and then for the abbreviated name by which the designated field label is known in the stored data (in this case Meteor). The programme proceeds on an incremental basis, making certain tests as it goes along in order to determine whether the zone embraced by the required label has been terminated by a punctuation period or by the addition of a new label. It continues to accumulate consecutive records, in accordance with the punctuation rules, until one of the terminating conditions is met, at which point a loop is made to the START position and the entire operation repeated for the next successfully identified label. The operation continues in this manner until the end of the file is reached when the terminal operator is
TABLE XIX
Procedure for Retrieval of Designated Field Labels

- apply retrieval programme to files
- designated field label identified
- stack preceding 10 or 11 and other data
- test for 31 or 34
- stack preceding dialect label
- stack designated field label
- stack following records
- test for
- test for new 32,42
given the option of calling up another information file for similar processing or terminating the procedure at that point.

7.8

Similar tactics are employed to devise programmes which will extract all kinds of retrievable information and provide printouts in whatever form the user requires. As a short-term operation, the retrieval of selected kinds of information provides the editor with an efficient means of maintaining control over the internal and somewhat illusive contents of his dictionary. A simple statistical comparison, for example, will enable him to determine whether there is a reasonable balance between the main subject fields covered by the dictionary; if this approach proves inadequate he can inspect the entire contents of selected subject fields, submit them to specialist translators for opinions, make selections on the basis of usage frequency and carry out whatever changes or additions his microcosmic study suggests to him. In the long term, this flexibility of operation will facilitate future revisions, enable the publication of mini-glossaries for the use of translators and serve as the basis for a rapid-access automatic dictionary. Equally important, as we have seen, is the perhaps unexpected byproduct of this kind of work—a finished result with a high level of accuracy and consistency in its physical layout.
8.0 Automatic Inversion of Entries

One of the most arduous and repetitive tasks in bilingual lexicography is the manual production of materials for one side of the dictionary based on entries in the other side. This kind of operation is usually an on-going process which must be carried out conscientiously if useful material is to be transferred from one section to the other and, in the case of bipartite dictionaries, if the essential relationship and balance between the two parts is to be successfully maintained. The sad facts are that the important activity of manual inversion is rarely undertaken on a systematic basis, either because it is a time-consuming and therefore expensive operation or because the editor in question has an inadequate understanding of its importance. The first of these probable reasons can be eliminated by a computer programme of the kind described below; the second, as we have suggested in various chapters of Parts I and II, is a matter of promulgating among bilingual lexicographers the theoretical and pedagogical basis of their work.

8.1 Automatic inversion of bilingual dictionary entries has been attempted on a small scale by one or two very specialized research projects; the lack of documentation on the subject seems to indicate that attempts have so far met with very little success. The French-Flemish
Index of the vocabulary of N. de Berlaimont was probably the first serious use of mechanization in bilingual lexicography and one of its outcomes was the automatic transformation of Flemish-French entries into French-Flemish entries. Since the source material was a simple vocabulary list with single-word translations, it cannot be said that this project advanced far into the labyrinthine intricacies of a true inversion procedure. Yet it was a beginning and achieved at a time when mechanical resolution of linguistic problems was very much in its infancy. All other ventures into automatic inversion, as far as it is possible to ascertain from the generally cursory accounts available in the literature, involve the so-called automatic dictionaries—data banks designed for rapid consultation and lookup—which usually build in a procedure for creating the opposite language "fiche" as soon as new material is added to the bank. Although automatic dictionaries are usually multilingual, their primary goal is usually definition and for this reason they operate on words or word groups in contexts which tend to define them. Inversion therefore consists simply in making the bottom half of the "fiche," which contains the translation and its context, into the top half of a "fiche" in the other language; the contexts bear no relationship to each other except insofar as they contain the equivalent lexical unit in the respective
languages. But the creation of new entries in the other language from the complex materials of the original entries in a normal bilingual dictionary is a much more difficult undertaking and calls upon all the ingenuity of lexicographer and programmer alike.

§1.1 Mechanical inversion of bilingual dictionary entries is based upon the simple conception that if an entry word a of one language has translations x, y and z in the other, then it should be possible to programme a computer to create three separate entries x, y and z each translated by a. A "flipping" routine to cover elementary dictionary entries containing nothing but the entry word and its translation(s) is an elementary programming task, but if more complex entries are to be processed, and if all the various elements described above are to be taken into consideration, the programme becomes proportionately complicated. To be an efficient editing adjunct it will have to make a number of significant changes in the restructured entry so that it becomes the accurate counterpart of the original. Among these changes are transformations involving grammar codes: if the original entry is an English noun (n) the inverted material will produce one or more French noun entries whose grammar notation in the dictionary is n plus the appropriate gender, which can be obtained from the
original translations; similarly, French reflexive verb entries, marked *vpron* will produce intransitive English verbs which must be marked *vintr*. Another important transformation involves all *labels* and *rubrics*: these materials appear in the language appropriate to the side of the dictionary in which they appear and must be accordingly altered in the process of inversion; those which belong to accessible elements of the file structure can be transformed automatically by a sub-routine which would match them against an internal "dictionary"; those which are inaccessible will have to be submitted to post-editing procedures.

Probably the most difficult part of an inversion algorithm is the need to deal effectively with contextual information following translations. If the entry word and any one of its translations appear in the example and its following translation respectively, the inverted examples can be correctly assigned; if they have undergone any morphological transformation in context, however, it is almost impossible to assign them to the correct headword after inversion. The situation becomes even more complex if the translation of the example does not contain any of the preceding translations in any form; only manual intervention will permit such inverted materials to be correctly located.
8.1.2 Once again, the success of an inversion programme depends largely on the efficiency of the file structure, although now a number of new dimensions are introduced which make it even more imperative that the consistency and accuracy referred to above should be given special emphasis. To take the simplest example, when a source language lexical unit has multiple translations which, after inversion, should become individual new entries, the means of separating these elements is of prime importance in order for them to be correctly identified as disparate lexical units. As with the retrieval programme, it is punctuation, consistently and accurately applied, which determines the appropriate divisions. Three kinds of punctuation are permissible at the 35 (translation of headword) level: commas and semi-colons separate synonyms or near synonyms and, from the point of view of the programme, are indistinguishable; colons signify the end of the translations and announce the beginning of an example containing the headword; if no examples follow, a period signals the end of the information at this particular semantic level, and if immediately followed by a 50 flag the end of the data file is indicated. If interrogation or exclamation marks are a part of the translation, they must still be followed by the appropriate punctuation:
The other reversible elements are 45's (translations of examples) which, in the case of multiples, are separated by commas and ended by semi-colons, if another example follows, or by periods if that is the end of the semantic level.

8.1.3 The employment of diagonal slashes to avoid repetitions is a common typographical expedient to conserve space. For example, Il est une heure/et quart, un quart/ avoids writing the two phrases Il est une heure et quart and Il est une heure un quart in their entirety. Strict rules must apply to the use of slashes if they are to be correctly interpreted by the inversion programme, especially when they are used at the 35 level where the constituents must be correctly separated and brought together again as new headwords. For alternations of two or more elements with a single one, the slashes must be contiguous with the alternate units:

empoigner vtr to/catch, take, lay/
hold of.

When double alternations occur, two sets of slashes must be used, each set abutting the materials it encloses:
/rentrer, se coucher/ /de bonne heure, tard/ which is interpreted as: rentrer de bonne heure, rentrer tard, se coucher de bonne heure, se coucher tard. The correct
spacing and contiguity of the slashes is crucial to accurate interpretation by the programme (just as it is to correct reading by the dictionary-user).

8.2 Identification of translations is the first cycle of the inversion programme. The second cycle isolates the associated elements and assigns them to their correct location in the newly created entry. In this cycle, the first transformation to be made is the conversion of grammar categories where required using whatever information is already available in the original dictionary data file and applying rules for adding data which is not there. Probably the majority of the grammar categories in the two sides of a dictionary of a pair of related languages will require no transformations, but significant numbers of equivalents are different enough to warrant building in a transformation procedure.

French and English, for example, differ in that the nouns of the former must have the gender appended and that many intransitive English verbs are rendered by French reflexive verbs. Gender information is obtainable from French translations (when transforming English entries) and from the grammar code (when transforming French entries); a sub-routine can be built into the overall programme to assign the correct grammar code to reflexive verb translations of English intransitive verb entries.
and to convert this code \( (\text{vpron}) \) to the appropriate form when making transformations in the other direction. Instances of non-equivalence in grammatical category are impossible to account for in an automatic procedure of this kind and must be dealt with by post-editing.

8.2.1 The other important ancilliary data to be treated in the second transformation cycle are the various labels for dialect, subject field and register which are retrievable elements of the data file described above. Subject labels are the simplest to process because they necessarily apply to both headword (or example) and translation(s) and will therefore occupy the same position in the newly-generated data file as they did in the original. The only modification which has to be carried out is translation of the abbreviated form to the appropriate form in the other language, if in fact they are different; this can be accomplished by a separate procedure applied to the entire inverted information file, either immediately following inversion or else when post-editing has been completed and the materials are considered to be ready for merging and contraction operations (described below, 8.6).

8.2.2 Dialect labels are a little more complicated because they apply only to the source language (headword or example, 31 and 41 respectively); if the target language
translation is also a regionalism, this is indicated by an appended label which has similar properties (in terms of data processing) to the gender markings described above, that is, it may be recognized, extracted and replaced in a 31 or 41 position in the new data file. Conversely, a source language entry or example marked regional at the 31 or 41 level may be converted to a translation with the regional label attached. The use of symbolic-type labels which are interlingual dispenses with the need for a translation cycle for regionalisms.

8.2.3 Like subject labels, register labels tend to be associated with source language as well as translations, although for a variety of reasons lexicographers sometimes see fit to provide a translation which is in a different register from the original; in the latter case, the label immediately precedes the translation and forms part of the 35 or 45 element:

\[
\text{narcotique nm [Drug] narcotic; [Slang] dope.}
\]

After inversion, "dope" will be a new entry word with a 33 label "Slang":

\[
\text{narcotic n [Drug] narcotique m.}
\]
\[
\text{dope n [Drug] [Slang] narcotique.}
\]

The eventual inversion of other French "Drug" terms will
provide equivalents of the same register as "dope" and which may either supplement or replace the more formal "narcotique."

8.2.4 There is another important series of labels used in bilingual dictionaries, the so-called explanatory rubrics, which may act as subsets of subject fields with more specific implications: [Sport, golf], [Ind, pêche], or as interpretive and disambiguating material:

\[
\text{dry adj sec, sèche; [country] aride; [well] tari; [course, etc.] abstrus.}
\]

As these labels clearly belong to an open set, it would be impracticable to provide a separate "dictionary" of them for later translation and they must be relegated to post-editing. There is also a good reason why they should not be automatically transformed in the inversion cycles and that is because there are numerous situations where they are inappropriate in the opposite language. For example, in the entry "dry" above the inversion programme would generate five new entries (one of which, "sèche," would eventually be subsumed under "sec"):

\[
\text{sec, sèche adj dry,}
\text{aride adj [country] dry,}
\text{tari adj [well] dry,}
\text{abstrus adj [course, etc.] dry.}
\]
The three rubrics "country," "well" and "course" are appropriate in the English-French versions because they illustrate an example of differential lexicology in which a single lexical unit of one language has a variety of equivalent realizations in the other language, according to context. By contrast, the inverted entries "aride," "tari" and "abstrus" are invariably rendered by "dry" and do not require the specifying labels; they must therefore be eliminated during post-editing.

8.2.5 In the course of the second cycle of inversion, various other elements are repeated or eliminated in the newly created entries. Pronunciations are deleted and a blank left for later processing by automatic phonetic transformation. Usage notes (18, 28, 38 and 48) are duplicated in the new data file so that the editor can make his own decision whether to retain them unaltered, modify and translate them or reject them altogether. Cross-references (17, 27, 37 and 47) are likewise duplicated because they often serve as extra translations or may be employed in usage notes. No attempt is made to imitate the numbered semantic subdivisions of the original since at this point in the process entries are monvalent (in the sense that they relate a single translation to a single headword); later procedures attempt to merge related materials and create meaningful,
although purely logical, semantic subdivisions.

§.3 One thorny question remains to be resolved. What is
to become of original translations of examples which have
no direct relationship with the preceding 35-level
translations? That is to say, if an entry $a$ is trans-
lated by $x$, $y$ and $z$, and a series of examples containing
$a$ are translated by strings which do not contain $x$, $y$ or
$z$. Let us now examine a real case to see what methods
might be adopted to keep track of this important
material:

- Inversion of this data file provides the following new
entries:

1. *nuisance* n [» no single equivalent in Fr.] *ennui* m, *désagrément* m; [Fig] *fléau* m: *What a ~!, Quel ennui!, Comme c'est ennuyeux!; Il gêne tout le monde, Il est assommant; [Pop] C'est un casse-pieds.*

This leaves four example translations unplaced because
they do not contain recognizable elements of the
preceding 35's:
The new entries (1), (2) and (3) require only to have the usage note removed since it is no longer appropriate. It is clear that the examples (4) to (7), which are useful and should be retained, can only be located under the correct headword by the intervention of the editor. This intervention should ideally occur during the recording of the original entries when a flag could be prefixed to the element of the string which should be the headword. In the case of these examples, the choice is obvious: ennuyeux, gêner, assommant and casse-pieds. Occasionally there may be two or even three words under which the editor would want the example to appear; he will be guided by user convenience and pedagogical motives of reinforcement, notions which cannot, unfortunately, be built into any computer programme. The alternative to pre-editing dislocated translations of this kind is to treat them by post-editing, a procedure chosen by the CD2 project which did not have the inversion programme prepared and operating in time to intervene at the recording stage. Post-
editing involves modifying the programme to take the "orphans" and dump them in a separate file from the regular inversions where they can be independently processed.

8.4 The functioning of the three inversion cycles is outlined in the flowchart sketches of Tables XX, XXI and XXII. The first cycle (Table XX) emphasizes the importance of punctuation in the correct identification of the elements of the file which are to be treated as independent new entries; thus, the programme first isolates, and then stacks in new data files as 10 elements, the former 35 elements. Unless a colon (signalling the introduction of examples) or a period (specifying the end of semantic operations for that particular sense) is encountered, the programme keeps looping to stack successive translations as new headwords. If a semi-colon (normally followed by another translation element) is immediately succeeded by a 37 or 38 (which is quite legal), the programme stacks these elements and searches for a colon or period after them.

8.4.1 The second cycle adds the associated labels to the appropriate elements as described above and then proceeds to make whatever modifications are required by differences in grammatical function. Table XXI describes an abbreviated design of these procedures indicating how
TABLE XX

First Inversion Cycle: Identification of 35's

select file for inversion processing

required 35 element identified

read to first punctuation

preceding element to 10 in new data file

original 10 stacked as 35 in new data file

next punctuation: or

NO

YES
TABLE XXI
Second Inversion Cycle: Label and Grammar Adjustments

- associated data assigned to new data file

- is 17 code n?

- adjust gender marks

- is 17 code vpron?
  - YES: make new 17 code vi
  - NO: is 10 se, s (se)?
    - make new 17 code vpron
TABLE XXII
Third Inversion Cycle: Identification of 45's

- Read following 45 string to first punctuation.
- Stack as new 44 with old 44 as new 45 - empty data file.
- Match with new 10?
  - Yes: Add 45 to new data file as 44 with old 44 as new 45.
  - No: Next punctuation?
    - Yes: Assign associated labels.
    - No: Stack as new 44 with old 44 as new 45.
alterations are made for gender and for verb discrepancies.

8.4.2 The third cycle (Table XXII) is a repetition of the first cycle but at level 4 of the data file. If a colon has previously been encountered, there must be following examples and their translations together with related labels. A period signifies the end of the operation and a return to the first part of the programme for new information. Each 45 element is read and matched against the newly created 10 headwords; if a word in the string is identical to any of these 10's it is stacked in that data file as a new 44 with the original 44 as its 45. If no successful match is made the original 44-45 pair are stacked as a new 44-45 pair (flipped over) in an empty data file. Loops to the beginning of the cycle are repeated until a period is encountered at which point the associated labels are added to the new data files in their requisite positions and a return to the first cycle is made in order to process new information or pass on to the next data file to be inverted.

8.5 The completed inversion of an entire information file is automatically stored on disk and labelled in such a way that it can be distinguished from original materials. For example, the first file of the letter G (called G01) would generate an inversion file called
INVGOl or some other appropriate name. It will by now be apparent that the automatically inverted dictionary is only a first approximation to the desired final form of the other half. Large numbers of the computer-generated entries will actually be almost perfect specimens requiring at most the addition of a punctuation mark or the removal of some extraneous element. These are the monovalent entries which make up a considerable portion of any dictionary and whose successful inversion by automatic procedures alone justifies the effort required to create a successful programme. An actual example follows:

(1) diamond n ◊, ◊ [Sport] [baseball] losange m ◊ .

(2) losange nm ◊ [Sport] [baseball] diamond ◊, ◊ .

The conversion of the original entry (1) to the new entry (2) was effected with but a single flaw—the lack of the final period—and even this could be added; but it is omitted in order to enable the possible addition of other translations during the merging and consolidation process when final punctuation can be carried out.

8.5.1 Approximately half of the inverted material, however, must be subjected to post-editing in order to correct errors and make whatever modifications cannot be
accomplished automatically. In order to avoid expensive on-line editing, the new entries can be obtained in hard copy printout as well as being transferred from disk to cassette; the editor simply does his corrections on the printout and the typist makes the changes in local mode on a terminal of the HP2644 variety. When all the required modifications have been carried out, the second version is returned to disk to replace the first version and await a merging and consolidation programme.

8.6 Automatic inversion carries with it a lot of duplication. If an entry a has translations x, y and z and produces three entries each translated by a, it is evident that this will not be the only occurrence of these new entries, although it is the only time they will appear with the translation a. Therefore they are not, strictly speaking, duplications but rather synonymous translations. Eventual alphabetization will place together all the occurrences of x translated by a, b, c, ... n, but before an attempt is made to merge them by automatic means it will be necessary for the editor to inspect every x in order to determine the possibility of homographs which can be appropriately numbered. A simple merging programme would take all the unlabelled translations and string them together as one 35, eliminating all of their data files. In the absence of any other
indication or intervention, the programme would assume all these translations to belong to the same sense of the word, and as this may not be accurate they will have to be submitted to further manual editing. If the same word \( x \) also had translations under a given label, these would be gathered in a single 35 and assigned an arbitrary semantic level number (probably based on the alphabetical order of the labels) starting with the next highest digit above the unlabelled translations.

8.6.1 For example, a series of inverted entries \( x \) might appear thus after alphabetization:

\[
\begin{align*}
\text{x} & \quad \text{p.} \\
\text{x} & \quad \text{b.} \\
\text{x} & \quad \text{g.} \\
\text{x} & \quad \text{d.} \\
\text{x} & \quad \text{[Archit] t.} \\
\text{x} & \quad \text{[Bot] m.} \\
\text{x} & \quad \text{[Fig] j.} \\
\text{x} & \quad \text{[Mil] a.}
\end{align*}
\]

After the second merging cycle, these various forms would be concatenated thus:

\[
\text{x \ n (1) p, b, g, d. (2) [Archit] t. (3) [Bot] m. (4) [Fig] j. (5) [Mil] a.}
\]

8.6.2 The arbitrary alphabetical ordering of the labelled
translation will in all probability offend traditional lexicography with its insistence on formal arrangements according to historical and statistical parameters. Yet these criteria of monolingual lexicography, as we have seen (above Chapter 11, Section II), are not necessarily appropriate to bilingual lexicography; nor are they the most efficient basis for user convenience which is probably best served by a repetitive arbitrary system which reflects the arrangement of the entries themselves. The appearance of figurative senses at differing and unpredictable positions in the data file, however, is one feature of an arbitrary ordering system which will require attention.

8.7 One improvement which can be brought to bear is a programme which retains alphabetical order for subject fields only (32's and 42's), while senses which relate to register (figurative, familiar, slang, etc.) could follow in whatever order is required; this would enable a logical concrete-to-abstract progression which, if consistently maintained, would only serve to improve ease of consultation. In fact, almost any particular sense ordering could be applied within certain limits.

9.0 Conclusion

A number of theoretical and pedagogical motives can be cited in support of a rigorously implemented inversion
procedure, whether achieved by automatic means or by traditional manual methods. Continuous cycles of inversion will ensure a balanced final product with a network of semantic relationships linking every part of the whole. In fact, this internal semantic equilibrium, which may be termed a kind of "semantic closure," is probably the most important real outcome of inversion. The other important goal, the certainty impressed upon the user that he may verify any element he finds in the dictionary by referring to the opposite section, is likewise achieved by inversion. Both goals are more readily realized by automatic procedures which ensure that inversion has in fact taken place on a systematic and global basis.

Like other computer operations, the inversion programme is accompanied by some significant by-products. In common with retrieval programmes, the success of automatic inversion is largely governed by the efficiency of the general file structure and by the same conditions of accuracy and consistency. Most of the problems encountered by the CD2 venture into automatic inversion were related to inconsistency in the recording of entries, particularly with regard to punctuation and the placement of slashes indicating alternations. But perhaps the most important side-effect of the inversion programme is the exposure of errors, not just typographical mistakes,
which are highlighted when they are transposed from the relative obscurity of translations to the readily inspected position of headword, but also the quality of the translations themselves which sometimes appear inadequate or inappropriate when viewed in isolation. Finally, automatic inversion, when carried to its limits, is the means of producing large numbers of entries with a very small expenditure of human effort. Probably 60% of the entries of a bilingual dictionary are monovalent ones, that is, there is a one-to-one correspondence between the respective signs of the two languages; even if only these entries are successfully inverted, a very significant proportion of the opposite side of the dictionary can be produced by the computer with minimal intervention on the part of the lexicographer. Combined with a phonetic generation programme this means, in effect, that almost two thirds of one side of a dictionary may be generated automatically.

A machine-readable corpus represents an incredible advance over totally manual methods of dictionary compilation. Although the computer is no substitute for human intellect and creativity, it does possess a number of advantages related to the rapid processing of carefully structured materials. It can read a text comprised of many millions of characters in a few short minutes and manipulate pieces of information with unbelievable speed.
and accuracy. But the automatic editing of lexicographical data is only one aspect in a whole series of potential applications, including automatic photocomposition of the final product and the ability to revise and update an existing "machine corpus" on a cumulative basis so that new editions could be produced at very short notice. The CD2 project is even planning the conversion of its materials to a form which will make them compatible with computerized word banks with which it may form a network for instant consultation on translation problems. Finally, as the CD2 experience has clearly demonstrated, the computer is a complex device which can only digest accurate and consistently organized information, and these are requirements which impose on lexicographers the need for an extra measure of care and precision in the formulation of their dictionary entries.
Notes to Chapter 16

1. Trésor de la langue française: dictionnaire alphabétique de la langue des XIXᵉ siècles (1789-1960), published by the Centre National de la Recherche Scientifique under the direction of Paul Imbs.


4. The Dictionnaire Bordas (1972), M. Cohen et al. (eds.) uses this kind of ordering arrangement.

5. There seems to be no good reason, however, for pêcher (noun) to be separated from pêché by the intervention of pêché.

6. In most cases, in fact, there is little confusion because of a thorough system of cross-references which ensures that the desired word is eventually tracked down.

7. For a complete description see Hays (1967), pp. 50 ff.

8. Ibid., pp. 52-3.

9. Adapted from CD2 entries.


12. See the investigation of an automatic transcription system, including prosody, described by Vanderslice (1968).


14. Details are supplied by Pratt and Silva (1967).

15. For example, the "context-sensitive" rules described by Silva (1974), pp. 262-3.
H. Josselson (1966) mentions, in particular, the experiments carried out by the Bunker-Ramo Corporation and described by G. Martins and G. Reitz in "Automatic Thesaurus Compilation," Progress Report No. 11, National Science Foundation, Bunker-Ramo Corporation, C 157-3011, October, 1963: "Taking what they term a 'reciprocal bilingual dictionary approach' to the problem of thesaurus compilation, the authors process dictionary entries through cycles of inversion, i.e., Russian-English-Russian. By this process the Russian "headwords" of a Russian-English dictionary are listed, along with the English equivalent word(s). The English words are then listed as the "headwords" of a reciprocal English-Russian dictionary, along with the Russian equivalent word(s), which are also listed. From these data, frequency-of-occurrence indices are constructed for the purposes of evaluation of the semantic proximity of a given word to its corresponding headword" (p. 77.).

Described by B. Quemada (1961).
References for Chapter 16


CHAPTER 17

GENERAL CONCLUSION

In the introductory chapter of Part I an allusion was made to the fact that lexicography has been practised for centuries and yet has neither defined the limits of its application nor succeeded in developing a theoretical framework for its performance. This was not intended as a disparagement of the generations of "harmless drudges" whose labours are a memorial to their scholarship and assiduity (for lexicography is not an idle diversion for the faint-hearted). When necessity calls for the production of a tool or weapon, artisans do not hesitate to develop an appropriate solution without reflecting on considerations other than the obvious utilitarian ones. So, too, early dictionary-makers responded to practical necessity and devoted their energies to the task in hand.

Present-day lexicographers tend to be drawn from the ranks of trained linguists whose motivations transcend purely commercial goals; therefore, it is only natural that they have begun to question traditional methodology and endeavoured to find linguistic justifications for the scientific precision which they would want to apply to their
craft. Projects like the *Trésor de la langue française* and the *Dictionnaire du français contemporain* bear witness to a new approach to lexicography in which the findings of modern linguistics are an integral part of every phase in the compilation of the dictionary; monolingual lexicography is thus emerging as a recognized and scientifically-based discipline within linguistics whose practitioners are leading scholars in the field. But one of the flaws in this movement is the tendency for bilingual lexicography to be subsumed under the same heading—a logical consequence, but one which has prevented it from establishing its own frame of reference, even though the two activities are demonstrably different.

The object of this study, as we asserted in the introductory chapter, was to attempt to prove that bilingual lexicography merits separate consideration in view of its distinctive characteristics and different requirements. In the historical and typological analyses of Chapters 2 and 3 we illustrated the particular features of bilingual dictionaries which distinguish them from their monolingual counterparts; this descriptive introduction was undertaken explicitly in order to define the limits of the study, but at the same time it revealed a number of significant structural details and indicated the factors which should motivate bilingual lexicographical ventures. A further specification of the boundaries of the field was made
possible by separating bilingual lexicography from translation which is erroneously considered to be more or less the same thing.

Since the purpose of a bilingual dictionary is to constitute the "equivalence relation," a good deal of attention was devoted to the properties of semantic structure and the manner in which the lexical inventories of two languages might be juxtaposed. Lexicographers must carefully weigh every pair of equated terms in their dictionaries, for those who consult them usually believe that the dictionary is the final authority on meaning (or equivalence). The establishment of the "equivalence relation" is not an elementary matter. In fact, many linguists have cast doubt on the theoretical possibility of translating from one language to another, and if this holds true for messages and entire texts, how much less viable is word-for-word equivalence?

There seems to be sufficient evidence, however, that the disparate lexical structures of different languages are not an intolerable impediment to communication across languages. A microcosmic view of the semantic organization of certain French and English lexical fields, for example, reveals that, although the same reality is sometimes viewed differently, the means of its verbal expression is frequently similar or even identical. Equivalence therefore appears to be theoretically defensible, although its lexicographical
manifestation is a complex and sometimes obscure approximation. One of the main problems of bilingual dictionaries is that the convergence of two lexical systems is interpreted by the unsophisticated as meaning that if the dictionary states that $a = b$, then this is an invariable truth. Less naive readers, on the other hand, may provide examples where $a = b$ is not true and conclude that the dictionary is unreliable or that its goal of equivalence is ineffectual.

An experienced lexicographer is aware that, although there are significant numbers of word pairs which demonstrate one-to-one correspondence in their lexical meaning, in practical terms many others are only marginally equivalent or can be equated only in very explicit circumstances. For this reason, the translations provided in a dictionary are only potentially equivalent because they are abstractions from real contexts and must be reconstituted in the environment which is appropriate to their semic structure. The general conclusion drawn from the theoretical analysis of Part I, in fact, is that the bilingual dictionary should be construed as a model of "lexical competence," a representation of "langue" rather than "parole." Using this as a working hypothesis, it becomes possible to justify the somewhat artificial co-extension in a dictionary of the lexical inventories of two languages. Moreover, we are provided with further evidence
of the requirement that polysemous words be clearly discriminated and of the desirability of including large quantities of explanatory material on the grounds that it furnishes instances of "parole" and serves to further specify the semantic extension of the words under consideration.

The linguistic methodology described in Part II is thus based on a model of "lexical competence" and "potential equivalence." Those who consult a bilingual dictionary are, in the main, learners of the second language whose particular needs must be catered to. Every aspect of practical lexical equivalence must therefore be carefully studied and appropriate information provided—whether it be in the form of details on the selectional restrictions which govern the use of a word in certain contexts or a warning to the effect that equivalence is only superficially or structurally applicable (as in the case of deceptive cognates). The means of indicating deviations from the ideal of perfect correspondence between lexical items may vary from one dictionary to another, but the motivation for doing so is the inviolable rule that everything possible must be done to avert native language interference and to remove the ambiguity that careless description imposes on the equivalence relation.

Bilingual lexicography thus lays claim to be an independent linguistic activity, susceptible to a set of
distinctive theoretical issues and possessing its own particular methodology. Like all branches of linguistics, many aspects of theory and practice are shared with other areas in the field, a feature which only serves further to assign bilingual lexicography to the domain of general linguistics. The application of technology to the compilation of bilingual dictionaries is the final step in the coming of age of bilingual lexicography which can now with justification take its place beside monolingual lexicography as a scientific linguistic endeavour.
APPENDIX I

CHRONOLOGICAL BIBLIOGRAPHY OF BILINGUAL FRENCH AND ENGLISH DICTIONARIES


1571 [Anon.]. *Dictionarie frech and english.* (Probably a reprint of Harryson).


Miège, Guy. *Nouvelle méthode . . . avec nomenclature francais-anglais.* Leyde:


<table>
<thead>
<tr>
<th>Year</th>
<th>Author</th>
<th>Title</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>1802</td>
<td>Garner, J.</td>
<td><em>Nouveau dictionnaire universel français-anglais et anglais-français.</em></td>
<td>Rouen</td>
</tr>
<tr>
<td>1804</td>
<td>Perrin, J.</td>
<td><em>Eléments de conversation français et anglais.</em></td>
<td>Paris</td>
</tr>
<tr>
<td>1810</td>
<td>Chambaud, Louis</td>
<td><em>Treasure of the French and English Languages.</em></td>
<td>London</td>
</tr>
<tr>
<td></td>
<td>Dufief, N. G.</td>
<td><em>Dictionnaire nouveau et universel des langues française et anglaise.</em></td>
<td>Philadelphia</td>
</tr>
<tr>
<td>1812</td>
<td>Poppleton, G.</td>
<td><em>Phraséologie française et anglaise.</em></td>
<td>Paris</td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>Nouveau éléments de conversation anglais-français.</em></td>
<td>Paris</td>
</tr>
<tr>
<td>1816</td>
<td>Harmonière, G.</td>
<td><em>Nouveau dictionnaire de poche français-anglais et anglais-français.</em></td>
<td>Paris</td>
</tr>
<tr>
<td>1819</td>
<td>Descarrières, A.</td>
<td><em>Dictionnaire français-anglais.</em></td>
<td>London</td>
</tr>
<tr>
<td>1821</td>
<td>Richardson, J.</td>
<td><em>Vocabulaire anglais et français ou dialogues familiers.</em></td>
<td>Paris</td>
</tr>
<tr>
<td>1822</td>
<td>Boniface, A.</td>
<td><em>Dictionnaire français-anglais et anglais-français.</em></td>
<td>Paris</td>
</tr>
<tr>
<td></td>
<td>Wanostrpocht, N.</td>
<td><em>A Classical Vocabulary.</em></td>
<td>London</td>
</tr>
<tr>
<td>1823</td>
<td>Stone, S.</td>
<td><em>Dictionnaire de poche français-anglais, anglais-français.</em></td>
<td>Paris</td>
</tr>
<tr>
<td>1826</td>
<td>O'Sullivan, D.</td>
<td><em>Nouveau dictionnaire de poche français-anglais et anglais-français.</em></td>
<td>Paris</td>
</tr>
<tr>
<td>1830</td>
<td>Sadler, P.</td>
<td><em>Manuel de phrases françaises et anglaises.</em></td>
<td>Paris</td>
</tr>
<tr>
<td>1834</td>
<td>Devereux, M.</td>
<td><em>Dictionnaire français-anglais.</em></td>
<td>Paris</td>
</tr>
<tr>
<td></td>
<td>Tibbins, J.</td>
<td><em>Dictionnaire français-anglais.</em></td>
<td>Paris</td>
</tr>
<tr>
<td>1835</td>
<td>Bossut, M.</td>
<td><em>French and English Word-Book.</em></td>
<td>Paris</td>
</tr>
<tr>
<td></td>
<td>[Anon.].</td>
<td><em>Vocabulaire du voyageur, anglais-français, français-anglais.</em></td>
<td></td>
</tr>
</tbody>
</table>
1836 O'Sullivan, D. **New English Reader (Dictionnaire anglais-français)**. Paris.


1844 Sadler, P. **Nouveau dictionnaire portatif anglais-français et français-anglais**. Paris.


1848 de Levizac, J. **A Dictionary of the French and English Languages**. London.


<table>
<thead>
<tr>
<th>Year</th>
<th>Author</th>
<th>Title</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>1855</td>
<td>Barrett, Th.</td>
<td><em>Dictionnaire anglais-français et français-anglais</em></td>
<td>Lyon</td>
</tr>
<tr>
<td></td>
<td>Nugent, T.</td>
<td><em>Dictionnaire de poche français-anglais et anglais-français</em></td>
<td>Paris</td>
</tr>
<tr>
<td></td>
<td>Sadler, P.</td>
<td><em>Petit dictionnaire classique anglais-français et français-anglais</em></td>
<td>Paris</td>
</tr>
<tr>
<td>1856</td>
<td>Tibbins, J.</td>
<td><em>Dictionary of the French and English Languages</em></td>
<td>Paris</td>
</tr>
<tr>
<td>1860</td>
<td>Elwall, A.</td>
<td><em>Nouveau dictionnaire français-anglais</em></td>
<td>Paris</td>
</tr>
<tr>
<td></td>
<td>Smith, L.</td>
<td><em>International English and French Dictionary</em></td>
<td></td>
</tr>
<tr>
<td>1861</td>
<td>Glashin, D.</td>
<td><em>Nouveau dictionnaire de poche français-anglais et anglais-français</em></td>
<td>Paris</td>
</tr>
<tr>
<td>1863</td>
<td>Clifton, M.</td>
<td><em>Manual of Conversation (English-French Vocabulary)</em></td>
<td>Paris</td>
</tr>
<tr>
<td></td>
<td>Contanseau, L.</td>
<td><em>Pocket Dictionary of the French and English Languages</em></td>
<td>London</td>
</tr>
<tr>
<td>1867</td>
<td>Contanseau, L.</td>
<td><em>A Practical Dictionary of the French and English Languages</em></td>
<td>London</td>
</tr>
<tr>
<td>1868</td>
<td>Boyer, A.</td>
<td><em>Nouveau dictionnaire français-anglais et anglais-français</em></td>
<td>Paris</td>
</tr>
<tr>
<td>1882</td>
<td>Nugent, T.</td>
<td><em>Nouveau dictionnaire de poche français-anglais</em></td>
<td>Paris</td>
</tr>
<tr>
<td>1883</td>
<td>Van Muyden, G.</td>
<td><em>Petit vocabulaire français</em></td>
<td>Paris</td>
</tr>
</tbody>
</table>


1920 Samson, D. N. English into French: 5,000 English Locutions. London.


<table>
<thead>
<tr>
<th>Year</th>
<th>Author</th>
<th>Title</th>
<th>Publisher/Details</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Kettridge, J. O.</td>
<td>French for English Idioms and Figurative Phrases, with many Quotations from French Authors.</td>
<td>London: Routledge.</td>
</tr>
</tbody>
</table>


SUMMARY OF FRENCH EXPRESSIONS EMPLOYING THE WORD "COUP" WITH ENGLISH EQUIVALENTS

1. [COUP + AGENT] (human or animal limb, etc.)

- coup d'aile: flap, wing-stroke
- " de bec: peck
- " du bout du doigt: dig, tap, rap
- " de corne: thrust, butt
- " de coude: nudge, elbow
- " de dent: bite, snap
- " de doigt: dig, tap
- " d'épaule: shove, shoulder
- " de genou: knee
- " de gosier: gulp, shout
- " de griffe: scratch
- " de gueule: roar, bellow
- " de langue: lick
- " de main: slap, (helping) hand
- " d'œil: glance
- " d'ongle: scratch
- " de patte: kick
- " de pied: kick, boot
- " du plat de la main: slap, cuff
- " de poing: punch, cuff, box
- " de pouce: jab (with a thumb)
- " de queue: flick (of the tail)
- " de reins: heave
- " de sabot: kick
- " de sang: stroke, fit
- " de tête: butt
donner des coups de corne to gore
" " " de coude to nudge
" " " de griffé to scratch
" " " d'ongle to scratch
" " " de pied to kick
" " " de poing to punch

2. [COUP + INSTRUMENT]
coup d'archet upstroke, downstroke, bowing
d'aviron (oar) stroke
de baguette wave, touch (of the conductor's baton)
de balai sweep
de barre (donner) steer (a ship)
de bâton hit, whack
de bistouri (surgeon's) incision
de brosse sweep, brush (-up)
de canne hit, strike
de chapeau (saluer) to raise one's hat
de chiffon dust, dusting
de ciseau snip
de cognée club
de couteau stab. slash
de crayon stroke (of the pen)
de épée stab, thrust
de épingle prick
de éponge sponge (off)
de éventail tap, rap (of fan)
de fer pressing, ironing
de fil phone call
de filet (fishing) cast, haul
" de fleuret
" de fouet
" de frein (donner)
" de hache
" de lance
" de marteau
" massue
" peigne (se donner)
" de pinceau
" de poignard
" de raquette
" de sabre
" de tampon
" de téléphone
" de torchon
" de volant
thrust
lash, flick
to brake
chop
thrust
hammer, knock
blow
comb, combing
(brush) stroke
stab
(tennis) stroke
thrust, slash
stamp
phone call
dust, dusting
swerve

traduire une version à
coups de dictionnaire
avoir un bon coup de
fourchette
donner des coups de bâton
enfoncer à coups de
marteau
to look every word up in
the dictionary
to be a big eater
to club, cudgel
to hammer (in)

3. [COUP + INSTRUMENT = SOUND]
coup de canon
" de carabine
" de cloche
" de feu
" de foudre
" de fouet
" de fusil
" de gong
bang, report, explosion
report, shot
ring, peal
shot, report, bang
thunderbolt
rack, slash (of whip)
shot, report
strike of the gong
coup de grisou: fire-damp explosion
" de piston piston stroke
" de midi the stroke of twelve
" de mine blast
" de roulis roll
" de sifflet whistle (blast)
" de sirène siren (blast)
" de sonnette ring
" de tambour beat; roll
" de tam-tam beat
" de timbre ring, peal
" de tonnerre clap, peal

sonner trois coups to strike three
donner un coup de trompette to sound the trumpet
donner un coup discret to tap
donner un coup violent to rap, thump

4. [COUP + WEATHER]
coup d'air rush (of air)
" de chaleur heat wave, sun stroke
" de foudre thunderbolt
" de froid cold snap, chill
" de mer heavy sea
" de soleil sunburn, sunstroke
" de tangage pitch, roll
" de tonnerre clap of thunder
" de vent blast, gust
5. **[COUP + effect on someone]**

<table>
<thead>
<tr>
<th>French</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>coup d'audace</td>
<td>piece of daring, bold stroke</td>
</tr>
<tr>
<td>&quot; de bonheur</td>
<td>stroke of luck</td>
</tr>
<tr>
<td>&quot; du ciel</td>
<td>godsend</td>
</tr>
<tr>
<td>&quot; dur</td>
<td>hard blow</td>
</tr>
<tr>
<td>&quot; mortel</td>
<td>mortal blow, fatal blow</td>
</tr>
<tr>
<td>&quot; de dé</td>
<td>throw (of the dice)</td>
</tr>
<tr>
<td>&quot; de désespoir</td>
<td>fit of despair</td>
</tr>
<tr>
<td>&quot; de destin</td>
<td>blow of fate</td>
</tr>
<tr>
<td>d'éclat</td>
<td>distinguished action, bold feat</td>
</tr>
<tr>
<td>&quot; d'état</td>
<td>rebellion, coup d'etat</td>
</tr>
<tr>
<td>&quot; de folie</td>
<td>fit, foolish act</td>
</tr>
<tr>
<td>&quot; de grace</td>
<td>finishing stroke</td>
</tr>
<tr>
<td>&quot; du hasard</td>
<td>mere chance</td>
</tr>
<tr>
<td>&quot; de maître</td>
<td>master stroke</td>
</tr>
<tr>
<td>&quot; de malheur</td>
<td>cruel blow</td>
</tr>
<tr>
<td>&quot; de la Providence</td>
<td>piece of luck</td>
</tr>
<tr>
<td>&quot; de mort</td>
<td>death blow</td>
</tr>
<tr>
<td>&quot; du sort</td>
<td>chance</td>
</tr>
<tr>
<td>&quot; de tête</td>
<td>impulse</td>
</tr>
</tbody>
</table>

6. **[COUP in sports]**

<table>
<thead>
<tr>
<th>French</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>coup adroit</td>
<td>skillful shot</td>
</tr>
<tr>
<td>&quot; au but</td>
<td>goal shot</td>
</tr>
<tr>
<td>&quot; bien joué</td>
<td>nice shot</td>
</tr>
<tr>
<td>&quot; d'approche</td>
<td>approach shot (golf)</td>
</tr>
<tr>
<td>&quot; droit</td>
<td>straight shot</td>
</tr>
<tr>
<td>&quot; heureux</td>
<td>successful shot</td>
</tr>
<tr>
<td>&quot; malheureux</td>
<td>unsuccessful shot</td>
</tr>
<tr>
<td>&quot; manqué</td>
<td>miss</td>
</tr>
<tr>
<td>&quot; roulé</td>
<td>run-up (shot)</td>
</tr>
<tr>
<td>&quot; de dés</td>
<td>throw (of the dice)</td>
</tr>
</tbody>
</table>
coup d'envoi  kickoff
beau coup  fine shot
jouer à coup sur  to bet on a sure thing
parier à coup sur (golf) stroke play
concours par coups

coup de billiard  billiard stroke
  " de boxe  punch
  " de dames  move
  " de fer  shot, putt, drive, hit
  " de tennis  shot, stroke
  " de crosse  (hockey) shot
  " franc  free kick
  " place  place kick
  " droit  home thrust (boxing)

7. [COUP = quantity]

boire un coup  to take a drink
boire qqch d'un coup  to drink at a draught,
boire qqch à petits coups  to gulp down

8. [COUP in adverbial expressions]

encore un coup  once more
au premier coup  from the start
du premier coup  to begin with
d'un coup  all at once
d'un seul coup  with a single blow
tout d'un coup  suddenly
coup sur coup  in rapid succession
<table>
<thead>
<tr>
<th>French Expression</th>
<th>English Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>pour le coup</td>
<td>this time</td>
</tr>
<tr>
<td>à tout coup</td>
<td>every time</td>
</tr>
<tr>
<td>à tous les coups</td>
<td>with absolute certainty</td>
</tr>
<tr>
<td>à coup sur</td>
<td>by ..., with the help of ...</td>
</tr>
<tr>
<td>à coups de ...</td>
<td>after it's all over</td>
</tr>
<tr>
<td>après coup</td>
<td></td>
</tr>
</tbody>
</table>

8. **[COUP in locutions]**

<table>
<thead>
<tr>
<th>French Expression</th>
<th>English Translation</th>
</tr>
</thead>
<tbody>
<tr>
<td>tué sur le coup</td>
<td>killed outright</td>
</tr>
<tr>
<td>jouer un sale coup à</td>
<td>to do the dirty on</td>
</tr>
<tr>
<td>un coup dur</td>
<td>a staggering blow</td>
</tr>
<tr>
<td>faire les 400 coups</td>
<td>to sow one's wild oats</td>
</tr>
<tr>
<td>être aux 100 coups</td>
<td>mad with grief, beside oneself</td>
</tr>
<tr>
<td>faire à qqn le coup du Père François</td>
<td>tp throttle someone</td>
</tr>
<tr>
<td>en mettre un coup</td>
<td>to put some sting into it</td>
</tr>
<tr>
<td>faire d'une pierre 2 coups</td>
<td>to kill 2 birds with one stone</td>
</tr>
<tr>
<td>frapper un grand coup</td>
<td>to strike hard</td>
</tr>
<tr>
<td>manquer son coup</td>
<td>to miss the mark</td>
</tr>
<tr>
<td>porter un coup à</td>
<td>to deal a blow to</td>
</tr>
<tr>
<td>réussir son coup</td>
<td>to bring it off</td>
</tr>
<tr>
<td>risquer le coup</td>
<td>to take a chance</td>
</tr>
<tr>
<td>sous le coup de</td>
<td>under the influence of</td>
</tr>
<tr>
<td>tenir le coup</td>
<td>to keep a stiff upper lip</td>
</tr>
<tr>
<td>faire un coup de tête</td>
<td>to act rashly</td>
</tr>
<tr>
<td>Ça m'a fait un coup</td>
<td>That came as a shock</td>
</tr>
<tr>
<td>tomber sous le coup d'une loi</td>
<td>to come under (the provisions)</td>
</tr>
<tr>
<td>un coup de rouge</td>
<td>a shot of red wine</td>
</tr>
<tr>
<td>entrer en coup de vent</td>
<td>to burst in</td>
</tr>
<tr>
<td>meutri de coups</td>
<td>bruised</td>
</tr>
<tr>
<td>Field Label</td>
<td>Abbreviation</td>
</tr>
<tr>
<td>-------------</td>
<td>--------------</td>
</tr>
<tr>
<td>Acct</td>
<td>Accounting</td>
</tr>
<tr>
<td>Acoust</td>
<td>Acoustics</td>
</tr>
<tr>
<td>Admin</td>
<td>Administration</td>
</tr>
<tr>
<td>Aeron</td>
<td>Aeronautics</td>
</tr>
<tr>
<td>Agric</td>
<td>Agriculture</td>
</tr>
<tr>
<td>Alcohol</td>
<td>Alcohol</td>
</tr>
<tr>
<td>Amerind</td>
<td>Amerindianism</td>
</tr>
<tr>
<td>Amus</td>
<td>Amusement</td>
</tr>
<tr>
<td>Anat</td>
<td>Anatomy</td>
</tr>
<tr>
<td>Anthrop</td>
<td>Anthropology</td>
</tr>
<tr>
<td>Antig</td>
<td>Antique</td>
</tr>
<tr>
<td>Archaeol</td>
<td>Archaeology</td>
</tr>
<tr>
<td>Archit</td>
<td>Architecture</td>
</tr>
<tr>
<td>Art</td>
<td>Art, Arts</td>
</tr>
<tr>
<td>Astrol</td>
<td>Astrology</td>
</tr>
<tr>
<td>Astronaut</td>
<td>Astronautics</td>
</tr>
<tr>
<td>Astronom</td>
<td>Astronomy</td>
</tr>
<tr>
<td>Auto</td>
<td>Automobile</td>
</tr>
<tr>
<td>Biol</td>
<td>Biology</td>
</tr>
<tr>
<td>Bot</td>
<td>Botany</td>
</tr>
<tr>
<td>Chem</td>
<td>Chemistry</td>
</tr>
<tr>
<td>Cine</td>
<td>Cinema</td>
</tr>
<tr>
<td>Col</td>
<td>Colours</td>
</tr>
<tr>
<td>Com</td>
<td>Commerce</td>
</tr>
<tr>
<td>Comput</td>
<td>Computers</td>
</tr>
<tr>
<td>Constr</td>
<td>Construction</td>
</tr>
<tr>
<td>Cost</td>
<td>Costumes</td>
</tr>
<tr>
<td>Cout.</td>
<td>Couture</td>
</tr>
<tr>
<td>Crim</td>
<td>Crime</td>
</tr>
<tr>
<td>Crustac</td>
<td>Crustacean</td>
</tr>
<tr>
<td>Culin</td>
<td>Culinary Arts</td>
</tr>
<tr>
<td>Dent</td>
<td>Dentistry</td>
</tr>
<tr>
<td>Drug</td>
<td>Drugs</td>
</tr>
<tr>
<td>Ecol</td>
<td>Ecology</td>
</tr>
<tr>
<td>Econ</td>
<td>Economics</td>
</tr>
<tr>
<td>Educ</td>
<td>Education</td>
</tr>
<tr>
<td>Elect</td>
<td>Elections</td>
</tr>
<tr>
<td>Electric</td>
<td>Electric(al)</td>
</tr>
<tr>
<td>Electron</td>
<td>Electronics</td>
</tr>
<tr>
<td>Engin</td>
<td>Engineering</td>
</tr>
<tr>
<td>Entom</td>
<td>Entomology</td>
</tr>
<tr>
<td>Equit</td>
<td>Equitation</td>
</tr>
<tr>
<td>Ethn</td>
<td>Ethnic</td>
</tr>
<tr>
<td>Ethnol</td>
<td>Ethnology</td>
</tr>
<tr>
<td>Pharm</td>
<td>Pharmacy</td>
</tr>
<tr>
<td>--------</td>
<td>----------</td>
</tr>
<tr>
<td>Philat</td>
<td>Philately</td>
</tr>
<tr>
<td>Philos</td>
<td>Philosophy</td>
</tr>
<tr>
<td>Photog</td>
<td>Photography</td>
</tr>
<tr>
<td>Phys</td>
<td>Physics</td>
</tr>
<tr>
<td>Police</td>
<td>Police</td>
</tr>
<tr>
<td>Post</td>
<td>Postal Service</td>
</tr>
<tr>
<td>Print</td>
<td>Printing</td>
</tr>
<tr>
<td>Psych</td>
<td>Psychiatry</td>
</tr>
<tr>
<td>Pub</td>
<td>Publicity</td>
</tr>
<tr>
<td>Radio</td>
<td>Radio</td>
</tr>
<tr>
<td>Realty</td>
<td>Real Estate</td>
</tr>
<tr>
<td>Rail</td>
<td>Railways</td>
</tr>
<tr>
<td>Recr</td>
<td>Recreation</td>
</tr>
<tr>
<td>Rel</td>
<td>Religion</td>
</tr>
<tr>
<td>Restaur</td>
<td>Restaurant</td>
</tr>
<tr>
<td>Rdgs</td>
<td>Roads</td>
</tr>
<tr>
<td>Sci</td>
<td>Society</td>
</tr>
<tr>
<td>Soc</td>
<td>Sociology</td>
</tr>
<tr>
<td>Sport</td>
<td>Sports</td>
</tr>
<tr>
<td>Staionery</td>
<td>Stationary</td>
</tr>
<tr>
<td>Statist</td>
<td>Statistics</td>
</tr>
<tr>
<td>Techn</td>
<td>Technical</td>
</tr>
<tr>
<td>Tel</td>
<td>Telephone</td>
</tr>
<tr>
<td>Telecom</td>
<td>Telecommunications</td>
</tr>
<tr>
<td>Text</td>
<td>Textiles</td>
</tr>
<tr>
<td>Theat</td>
<td>Théâtre</td>
</tr>
<tr>
<td>Title</td>
<td>Titre</td>
</tr>
<tr>
<td>Topon</td>
<td>Toponymie</td>
</tr>
<tr>
<td>Tour</td>
<td>Tourisme</td>
</tr>
<tr>
<td>Transp</td>
<td>Transport</td>
</tr>
<tr>
<td>TV</td>
<td>Télévision</td>
</tr>
<tr>
<td>Typo</td>
<td>Typographie</td>
</tr>
<tr>
<td>Union</td>
<td>Syndicat</td>
</tr>
<tr>
<td>Univ</td>
<td>Université</td>
</tr>
<tr>
<td>Urban</td>
<td>Urbanisme</td>
</tr>
<tr>
<td>Vet</td>
<td>Vétérinaire</td>
</tr>
<tr>
<td>Weap.</td>
<td>Arme</td>
</tr>
<tr>
<td>Wine</td>
<td>Vin</td>
</tr>
<tr>
<td>Zoo</td>
<td>Vins</td>
</tr>
<tr>
<td>Zool</td>
<td>Zoologie</td>
</tr>
</tbody>
</table>

**REGISTER**

<table>
<thead>
<tr>
<th>Dimin</th>
<th>Diminutive</th>
<th>Dimin</th>
<th>Diminutif</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fam</td>
<td>Familier</td>
<td>Fam</td>
<td>Familier</td>
</tr>
<tr>
<td>Fig</td>
<td>Figuré</td>
<td>Fig</td>
<td>Figuré</td>
</tr>
<tr>
<td>Loc</td>
<td>Location</td>
<td>Loc</td>
<td>Location</td>
</tr>
<tr>
<td>Mel</td>
<td>Melioratif</td>
<td>Mél</td>
<td>Mélioratif</td>
</tr>
<tr>
<td>Lit</td>
<td>Literary</td>
<td>Litt</td>
<td>Littéraire</td>
</tr>
<tr>
<td>------</td>
<td>----------------</td>
<td>------</td>
<td>--------------</td>
</tr>
<tr>
<td>Pej</td>
<td>Pejorative</td>
<td>Péj</td>
<td>Péjoratif</td>
</tr>
<tr>
<td>Pop</td>
<td>Popular</td>
<td>Pop</td>
<td>Populaire</td>
</tr>
<tr>
<td>Prov</td>
<td>Proverb</td>
<td>Prov</td>
<td>Proverbe</td>
</tr>
<tr>
<td>Slang</td>
<td>Slang</td>
<td>Argot</td>
<td>Argot</td>
</tr>
<tr>
<td>Vulg</td>
<td>Vulgar</td>
<td>Vulg</td>
<td>Vulgaire</td>
</tr>
<tr>
<td>[X]</td>
<td>Deceptive</td>
<td>[X]</td>
<td>Faux Ami</td>
</tr>
<tr>
<td></td>
<td>Cognate</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX IV

SAMPLE LIST OF DOCUMENTATION USED IN CD2

Note: The abbreviations correspond to codes used to identify the source of information and are recorded on file cards.

ABC ABC of Plain Words (Gower)
AC Air Canada (miscellaneous pamphlets)
ACELF CELF Bulletin
ACT Actualité terminologique
ADC Annuaire du Canada
ADS American Dialect Society
ADT Arrêtés de terminologie (Enrichissement de la langue française)
AED American Everyday Dictionary
AERO Lexicon of Aeronautical Terms (ICAO)
AFU Affaires universitaires
AGD Annuaire du gouvernement (Federal government directory)
AIM The Arts in Montreal
AMF Anglais modern/Modern French (Macquinghen)
AMI L'argot de milieu (Lacassaîgne et Devaux)
AMO L'argot moderne (Sandry et Carrère)
AMS American Speech
AMT American Translators' Association
ATA The ATA Magazine
AVF Améliorez votre français
BAT Banque de terminologie (Montréal)
BBN Bulletin de la bibliothèque nationale
BCC B.C. Calendar of Events
BDE Barnhardt Dictionary of New English Since 1963
BEL Dictionnaire général de la langue française au Canada (Belisle)
<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>BFD</td>
<td>Bellow's French and English Dictionary</td>
</tr>
<tr>
<td>BID</td>
<td>Bibliographie de dictionnaires</td>
</tr>
<tr>
<td>BIK</td>
<td>Birk's Jewellers bilingual catalogue</td>
</tr>
<tr>
<td>BIN</td>
<td>Bulletin d'information du laboratoire d'analyse lexicologique (Besançon)</td>
</tr>
<tr>
<td>BIS</td>
<td>Bibliography of Interlingual Scientific and Technical Dictionaries</td>
</tr>
<tr>
<td>BM</td>
<td>Banque de mots (Montréal)</td>
</tr>
<tr>
<td>BNF</td>
<td>Brèves nouvelles de France</td>
</tr>
<tr>
<td>BOIS</td>
<td>Le bois--vocabulaires normalisés (Montréal)</td>
</tr>
<tr>
<td>BOR</td>
<td>Dictionnaire BORDAS du français vivant</td>
</tr>
<tr>
<td>BPL</td>
<td>Bulletin pédagogique Larousse</td>
</tr>
<tr>
<td>BT</td>
<td>Bulletin de terminologie (Ottawa)</td>
</tr>
<tr>
<td>BUS</td>
<td>Bulletin signalétique (CNRS)</td>
</tr>
<tr>
<td>CAB</td>
<td>Cahiers de la Publicité</td>
</tr>
<tr>
<td>CAD</td>
<td>C'est-à-dire</td>
</tr>
<tr>
<td>CASS</td>
<td>Cassell's French and English Dictionary</td>
</tr>
<tr>
<td>CAUT</td>
<td>Canadian Association of University Teachers (Newsletter)</td>
</tr>
<tr>
<td>CBC</td>
<td>CBC Handbook for Announcers</td>
</tr>
<tr>
<td>CBW</td>
<td>Glossary of Canadian and British Words</td>
</tr>
<tr>
<td>CCV</td>
<td>Che Cos'è Vocabulario (Migliorini)</td>
</tr>
<tr>
<td>CDR</td>
<td>Codoroute (France)</td>
</tr>
<tr>
<td>CEN</td>
<td>Canadian English (Orkin)</td>
</tr>
<tr>
<td>CET</td>
<td>Comité d'étude des termes techniques français</td>
</tr>
<tr>
<td>CGG</td>
<td>Concise Glossary of Geographical Terms</td>
</tr>
<tr>
<td>CGQ</td>
<td>Chasse et gibier du Québec</td>
</tr>
<tr>
<td>CIB</td>
<td>Canadian Imperial Bank of Commerce (Newsletter)</td>
</tr>
<tr>
<td>CIV</td>
<td>Glossary of Civil Aviation Definitions</td>
</tr>
<tr>
<td>CNL</td>
<td>Canadian National (Linguistic services)</td>
</tr>
<tr>
<td>CNT</td>
<td>Canadian National (Terminology)</td>
</tr>
<tr>
<td>COF</td>
<td>Concise Oxford French Dictionary</td>
</tr>
<tr>
<td>COE</td>
<td>Concise Oxford English Dictionary</td>
</tr>
<tr>
<td>COL</td>
<td>Cahiers de l'Office de la langue française au Canada</td>
</tr>
<tr>
<td>Code</td>
<td>Title</td>
</tr>
<tr>
<td>------</td>
<td>----------------------------------------------------------------------</td>
</tr>
<tr>
<td>COW</td>
<td>Choice of Words (Collins)</td>
</tr>
<tr>
<td>CPR</td>
<td>Canadian Pacific Railway (miscellaneous documents)</td>
</tr>
<tr>
<td>CTR</td>
<td>Communication et traduction (Darbelnet)</td>
</tr>
<tr>
<td>CUC</td>
<td>Canadian Universités and Colleges</td>
</tr>
<tr>
<td>CUV</td>
<td>Culture vivante</td>
</tr>
<tr>
<td>CYB</td>
<td>Canada Year Book</td>
</tr>
<tr>
<td>DAL</td>
<td>Dictionnaire des affaires (Larousse)</td>
</tr>
<tr>
<td>DAN</td>
<td>Dictionnaire analogique (Maquet)</td>
</tr>
<tr>
<td>DAP</td>
<td>Dictionary of American Politics</td>
</tr>
<tr>
<td>DAR</td>
<td>Dictionnaire d'argot (larue)</td>
</tr>
<tr>
<td>DATA</td>
<td>Datamation</td>
</tr>
<tr>
<td>DBA</td>
<td>Dictionaries British and American (Hulbert)</td>
</tr>
<tr>
<td>DBC</td>
<td>Dictionnaire Beauchemin canadien</td>
</tr>
<tr>
<td>DBI</td>
<td>Dictionary of Biology (Penguin)</td>
</tr>
<tr>
<td>DCE</td>
<td>Dictionary of Canadian English (Gage)</td>
</tr>
<tr>
<td>DCFS</td>
<td>Dictionary of Colourful French Slanguage</td>
</tr>
<tr>
<td>DCHP</td>
<td>Dictionary of Canadianisms on Historical Principles</td>
</tr>
<tr>
<td>DCP</td>
<td>Dictionary of Computers (Penguin)</td>
</tr>
<tr>
<td>DDC</td>
<td>Dictionnaire des difficultés de la langue française au Canada</td>
</tr>
<tr>
<td>DDF</td>
<td>Dictionnaire des difficultés de la langue française (Thomas)</td>
</tr>
<tr>
<td>DDL</td>
<td>Dictionnaire de la loi</td>
</tr>
<tr>
<td>DDM</td>
<td>Dictionnaire des média</td>
</tr>
<tr>
<td>DDS</td>
<td>Dictionnaire des synonymes</td>
</tr>
<tr>
<td>DEF</td>
<td>Défense de la langue française</td>
</tr>
<tr>
<td>DEL</td>
<td>Dictionnaire de l'électrotechnique et l'électronique</td>
</tr>
<tr>
<td>DFC</td>
<td>Dictionnaire du français contemporain</td>
</tr>
<tr>
<td>DFO</td>
<td>Dictionnaire fondamental (Gougenheim)</td>
</tr>
<tr>
<td>DFP</td>
<td>Dictionnaire de la francophonie</td>
</tr>
<tr>
<td>DFW</td>
<td>Dictionary of Foreign Words (Newmark)</td>
</tr>
<tr>
<td>DGP</td>
<td>Dictionary of Geography (Penguin)</td>
</tr>
<tr>
<td>DGL</td>
<td>Dictionary of Geology (Penguin)</td>
</tr>
<tr>
<td>DIF</td>
<td>Dictionnaire inverse de la langue française (Juillard)</td>
</tr>
<tr>
<td>DIJ</td>
<td>Dissertation juridique</td>
</tr>
<tr>
<td>Code</td>
<td>Title</td>
</tr>
<tr>
<td>------</td>
<td>-------</td>
</tr>
<tr>
<td>DIM</td>
<td>Dictionnaire des idées suggérées par les mots (Rouaix)</td>
</tr>
<tr>
<td>DIT</td>
<td>Dictionnaire international du tourisme</td>
</tr>
<tr>
<td>DJD</td>
<td>Daniel Jones: English Pronunciation Dictionary</td>
</tr>
<tr>
<td>DMB</td>
<td>Description and Measurement of Bilingualism (Kelly)</td>
</tr>
<tr>
<td>DML</td>
<td>Dictionnaire moderne Larousse</td>
</tr>
<tr>
<td>DMN</td>
<td>Dictionnaire des mots nouveaux</td>
</tr>
<tr>
<td>DMS</td>
<td>Dictionnaire des mots sauvages</td>
</tr>
<tr>
<td>DNE</td>
<td>Dictionary of New Words in English (Berg)</td>
</tr>
<tr>
<td>DNW</td>
<td>Dictionary of New Words (Reifer)</td>
</tr>
<tr>
<td>DOR</td>
<td>Dictionnaire orthographique (Garnier)</td>
</tr>
<tr>
<td>DPF</td>
<td>Dictionnaire de la prononciation française (Warnant)</td>
</tr>
<tr>
<td>DPL</td>
<td>Dictionnaire pratique Larousse</td>
</tr>
<tr>
<td>DPP</td>
<td>Dictionary of Terms Used in the Paper and Printing Industry</td>
</tr>
<tr>
<td>DRE</td>
<td>Dictionnaire de rédaction (Godaert)</td>
</tr>
<tr>
<td>DSI</td>
<td>Dictionary of Science (Penguin)</td>
</tr>
<tr>
<td>DSY</td>
<td>Dictionnaire des synonymes (Bénac)</td>
</tr>
<tr>
<td>DTAF</td>
<td>Dictionnaire technique--anglais-français</td>
</tr>
<tr>
<td>DTG</td>
<td>Dictionnaire technique général</td>
</tr>
<tr>
<td>DTM</td>
<td>Dictionnaire français et anglais de terminologie mathématique</td>
</tr>
<tr>
<td>DUP</td>
<td>Dupont of Canada (various documents)</td>
</tr>
<tr>
<td>EAT</td>
<td>Eaton's Catalogue</td>
</tr>
<tr>
<td>ÉCE</td>
<td>Economie de l'énergie</td>
</tr>
<tr>
<td>ENO</td>
<td>Entre Nous</td>
</tr>
<tr>
<td>EQM</td>
<td>Equivalent Measures</td>
</tr>
<tr>
<td>ERT</td>
<td>Etude de Radio/Télévision</td>
</tr>
<tr>
<td>ESA</td>
<td>English Synonyms, Antonyms and Prepositions</td>
</tr>
<tr>
<td>EVV</td>
<td>Enrichissez votre vocabulaire (Reader's Digest)</td>
</tr>
<tr>
<td>EXP</td>
<td>L&quot;Express</td>
</tr>
<tr>
<td>FAM</td>
<td>Les faux amis (Derocquigny et Koessler)</td>
</tr>
<tr>
<td>FCB</td>
<td>Folk Names for Canadian Birds (McAtee)</td>
</tr>
<tr>
<td>FHQ</td>
<td>Forces Hydro-Québec</td>
</tr>
</tbody>
</table>
FLA Le français, langue des affaires (Clas et Horguelin)
FLE Foreign Language-English Dictionaries
FOC Formulaire commerciale
FPB French Phrase Book (Collins)
FRA France
FVR French Vocabularies for Repetition
GASC Gasc's French and English Dictionary
GBG Guide to Butchart's Gardens (multilingual)
GDA Guide de l'automobiliste
GPFC Glossaire du parler français au Canada
GLT Glossary of Linguistic Terminology
GNC English Geographical Names in Canada
GPI Glossaire du papetier et de l'imprimeur
GTI Glossaire des termes de l'informatique
GUC Guide des carrières
GUT Gestion de vos titres
HAP1 Harrap's Concise French and English Dictionary
HAP2 Harrap's Shorter French and English Dictionary
HAP3 Harrap's Standard French and English Dictionary
HNS Harrap's New Standard French-English Dictionary
HSC Harrap's French-English Dictionary of Slang and Colloquialisms
HSD The Howard Smith Dictionary
HT New York Herald Tribune
ICA International Civil Aviation (terminology)
INC The Incorporated Linguist
INCO INCOterms, 1953 (International Chambers of Commerce)
ITD International Travel Dictionary
ITM Informations touristiques (Monde)
JOLI Joliette—annuaire de téléphone
JOU Petit dictionnaire du joual en français
KCF Ketteridge's Commercial and Financial Dictionary
KED King's English Dictionary
KEV  Key to English Vocabulary
KTD  Ketteridge's Technical Dictionary
LAL  Lexique de l'aluminium
LAP  Larousse de poche (French and English)
LCA  Lexicon of Civil Aviation
LCE  Le canard enchaîné
LEM  Lexique de menuiserie
LFA  Les faux amis (Koessler)
LGA  Larousse gastronomique
LFE  Le français elementaire
LFI  Le français international (Vinay et al)
LIN  Lingot (journal of Alcan company)
LLI  Lexique americain-français de la langue idiomatique
LMC  Langue maritime commerciale
LMD  Le monde diplomatique
LMJ  Le mot juste
LMM  Larousse moderne (Marguerite Dubois)
LTL  Lexique de terminologie linguistique (Marouzeau)
MAC  Ministère des affaires culturelles (rapport annuel)
MAP  Mots anglais perfides (Deroquigny)
MARS  Marshall's Fabrics of Montreal (pamphlet)
MCL  MacLean's Magazine
MDE  Musée des erreurs
META  Meta (Journal des traducteurs/Translators' Journal)
MFA  Mots amis et faux amis (Kirk-Greene)
MFC  Manuel du français commercial à l'usage des étrangers
MIA  Musical Instruments Through the Ages
MIF  Made in France (Encyclopédie de la qualité française)
MIL  Military Dictionary/Dictionnaire militaire
MANU  Manufrance
MDV  Les mots dans le vent (Larousse)
MOA  Les mots anglais
MOE  Montréal et environs
MOF  Les mots français
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>MRC</td>
<td>Monnaie royale du Canada (pamphlet)</td>
</tr>
<tr>
<td>MRN</td>
<td>Ministère du revenu national, impôt (tax guide book)</td>
</tr>
<tr>
<td>MWD</td>
<td>Merriam Webster Pocket Dictionary</td>
</tr>
<tr>
<td>MXD</td>
<td>Mieux dire</td>
</tr>
<tr>
<td>NAN</td>
<td>Nation nouvelle (Revue du Canada français)</td>
</tr>
<tr>
<td>NBF</td>
<td>New Briefs from France</td>
</tr>
<tr>
<td>NCC</td>
<td>Nautical and Commercial Conversation</td>
</tr>
<tr>
<td>NLB</td>
<td>Nouveau Larousse bilingue (Mergault)</td>
</tr>
<tr>
<td>NLU</td>
<td>Nouveau Larousse universel</td>
</tr>
<tr>
<td>NMD</td>
<td>New Mathematics Dictionary</td>
</tr>
<tr>
<td>NPL</td>
<td>Nouveau petit Larousse</td>
</tr>
<tr>
<td>NRM</td>
<td>Noms des rues de la ville de Montréal</td>
</tr>
<tr>
<td>NUG</td>
<td>Nugent's Up-to-date Dictionary</td>
</tr>
<tr>
<td>OCDE</td>
<td>Catalogue des publications de l'économie</td>
</tr>
<tr>
<td>OED</td>
<td>Oxford English Dictionary</td>
</tr>
<tr>
<td>OLF</td>
<td>Office de la langue française au Québec</td>
</tr>
<tr>
<td>ORL</td>
<td>Orthographie et lexicographie (Catach et al)</td>
</tr>
<tr>
<td>ORT</td>
<td>Vocabulaire orthographique (Ters et al)</td>
</tr>
<tr>
<td>PDI</td>
<td>Petit dictionnaire en images</td>
</tr>
<tr>
<td>PDQ</td>
<td>Poissons du Québec</td>
</tr>
<tr>
<td>PED</td>
<td>Penguin English Dictionary</td>
</tr>
<tr>
<td>PEX</td>
<td>Par exemple</td>
</tr>
<tr>
<td>PFR</td>
<td>Proverbes français</td>
</tr>
<tr>
<td>PHC</td>
<td>Practical Handbook of Canadian French</td>
</tr>
<tr>
<td>PLF</td>
<td>Petit dictionnaire des locutions françaises</td>
</tr>
<tr>
<td>POD</td>
<td>Poissons d'eau douce</td>
</tr>
<tr>
<td>PLW</td>
<td>Présence francophone</td>
</tr>
<tr>
<td>PROB</td>
<td>Petit Robert</td>
</tr>
<tr>
<td>PVF</td>
<td>Parlez-vous franglais?</td>
</tr>
<tr>
<td>PDD</td>
<td>Petit dictionnaire raisonné des difficultés et exceptions de la langue française</td>
</tr>
<tr>
<td>QA</td>
<td>Québec-Amérique</td>
</tr>
<tr>
<td>QEB</td>
<td>Québec en bref</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Title</td>
</tr>
<tr>
<td>--------------</td>
<td>--------------------------------------------</td>
</tr>
<tr>
<td>RC</td>
<td>Radio Canada (services linguistiques)</td>
</tr>
<tr>
<td>RBC</td>
<td>Royal Bank of Canada (newsletter)</td>
</tr>
<tr>
<td>RCL</td>
<td>Revue canadienne de linguistique</td>
</tr>
<tr>
<td>REF</td>
<td>Réflexions pour 1985</td>
</tr>
<tr>
<td>RIM</td>
<td>Revue Imperial Oil</td>
</tr>
<tr>
<td>RIO</td>
<td>Revue internationale d'onomastique</td>
</tr>
<tr>
<td>RLQ</td>
<td>Renseignements généraux (Université Laval)</td>
</tr>
<tr>
<td>RLS</td>
<td>Regional Language Studies</td>
</tr>
<tr>
<td>ROB</td>
<td>Le dictionnaire Robert (Vols. 1-6)</td>
</tr>
<tr>
<td>ROB7</td>
<td>Robert Supplement</td>
</tr>
<tr>
<td>ROG</td>
<td>Roget's Thesaurus</td>
</tr>
<tr>
<td>RRO</td>
<td>Robert's Rules of Order</td>
</tr>
<tr>
<td>SCD</td>
<td>Science dimension</td>
</tr>
<tr>
<td>SCF</td>
<td>Speaking Canadian French (Orkin)</td>
</tr>
<tr>
<td>SCT</td>
<td>Sous-comité des termes-techniques</td>
</tr>
<tr>
<td>SFA</td>
<td>Stylistique comparée du français et de l'anglais</td>
</tr>
<tr>
<td>SFL</td>
<td>Semantic Frequency List (Eaton)</td>
</tr>
<tr>
<td>SFQ</td>
<td>Service de la faune du Québec</td>
</tr>
<tr>
<td>SGW</td>
<td>Sir George Williams University (newsletter)</td>
</tr>
<tr>
<td>SHELL</td>
<td>Shell News</td>
</tr>
<tr>
<td>SLE</td>
<td>General Service List of English Words and Phrases</td>
</tr>
<tr>
<td>SOL</td>
<td>Le Soleil</td>
</tr>
<tr>
<td>SSB</td>
<td>Sophisticated Synonym Book</td>
</tr>
<tr>
<td>SSB</td>
<td>Sophisticated Synonym Book</td>
</tr>
<tr>
<td>SSC</td>
<td>Simpsons-Sears Catalogue (French and English versions)</td>
</tr>
<tr>
<td>STT</td>
<td>Scientific and Technical Translation</td>
</tr>
<tr>
<td>SVY</td>
<td>The Substitute for &quot;Very&quot;</td>
</tr>
<tr>
<td>SYB</td>
<td>The Statesman's Yearbook</td>
</tr>
<tr>
<td>TB</td>
<td>Treasury Board of Canada (bilingual documents)</td>
</tr>
<tr>
<td>TBD</td>
<td>Thorndike-Barnhardt Handy Dictionary</td>
</tr>
<tr>
<td>TERM</td>
<td>Terminologie (Bulletin du Bureau de terminologie. Luxembourg)</td>
</tr>
<tr>
<td>TIG</td>
<td>Travaux de l'institut de géographie (Laval)</td>
</tr>
<tr>
<td>TIME</td>
<td>Time Magazine</td>
</tr>
<tr>
<td>TIN</td>
<td>Translation Inquirer</td>
</tr>
<tr>
<td>TNQ</td>
<td>Toponymie de côtes de Nouveau Québec</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>TRAD</td>
<td>Traduire</td>
</tr>
<tr>
<td>TTF</td>
<td>Termes techniques francais</td>
</tr>
<tr>
<td>TUP</td>
<td>Tu parles</td>
</tr>
<tr>
<td>UAB</td>
<td>Usage and Abusage</td>
</tr>
<tr>
<td>UMC</td>
<td>Université de Montréal Calendar</td>
</tr>
<tr>
<td>VAP</td>
<td>Vous avez la parole</td>
</tr>
<tr>
<td>VCP</td>
<td>Vocabulaire--cours supérieur</td>
</tr>
<tr>
<td>VM</td>
<td>Vocabulaire de médecine</td>
</tr>
<tr>
<td>VOM</td>
<td>Vocabulaire moderne français-anglais (Cornu)</td>
</tr>
<tr>
<td>VOP</td>
<td>Vocabulaire pratique de l'anglais au français</td>
</tr>
<tr>
<td>VOR</td>
<td>Vocabulaire orthographique de base</td>
</tr>
<tr>
<td>VOY</td>
<td>Le voyageur canadien</td>
</tr>
<tr>
<td>VRC</td>
<td>Vocabulary of Railway Cybernetics</td>
</tr>
<tr>
<td>VTC</td>
<td>Vocabulaire technique industriel et commercial</td>
</tr>
<tr>
<td>WAW</td>
<td>What a Word!</td>
</tr>
<tr>
<td>WCD</td>
<td>Webster's New Collegiate Dictionary</td>
</tr>
<tr>
<td>WIII</td>
<td>Webster's Third International Dictionary</td>
</tr>
<tr>
<td>WEM</td>
<td>Weekend Magazine</td>
</tr>
<tr>
<td>WIN</td>
<td>Winston's Dictionary of Canadian English</td>
</tr>
<tr>
<td>WNW</td>
<td>Webster's New World Dictionary of the American Language</td>
</tr>
<tr>
<td>WOR</td>
<td>Words ... Their Spelling, Pronunciation, Definition</td>
</tr>
</tbody>
</table>
### APPENDIX V

**SAMPLES OF PHONETIC DESCRIPTION USED IN CD2**

<table>
<thead>
<tr>
<th>ENGLISH VOWELS</th>
<th>ENGLISH CONSONANTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>[i:]</td>
<td>[p]</td>
</tr>
<tr>
<td>[e]</td>
<td>[f]</td>
</tr>
<tr>
<td>[æ]</td>
<td>[t]</td>
</tr>
<tr>
<td>[ɔ]</td>
<td>[d]</td>
</tr>
<tr>
<td>[ɔ]</td>
<td>[k]</td>
</tr>
<tr>
<td>[u]</td>
<td>[g]</td>
</tr>
<tr>
<td>[ʌ]</td>
<td>[s]</td>
</tr>
<tr>
<td>[ə]</td>
<td>[z]</td>
</tr>
<tr>
<td>[o]</td>
<td>[f]</td>
</tr>
<tr>
<td>[i:]</td>
<td>[v]</td>
</tr>
<tr>
<td>[u:]</td>
<td>[l]</td>
</tr>
<tr>
<td>[ɛə]</td>
<td>[r]</td>
</tr>
<tr>
<td>[æ]</td>
<td>[m]</td>
</tr>
<tr>
<td>[ʌ]</td>
<td>[n]</td>
</tr>
<tr>
<td>[ʌ]</td>
<td>[ŋ]</td>
</tr>
</tbody>
</table>

**ENGLISH STRESS MARKS**

- **Primary stress** [']
  - [əˈɡɛn] again
  - [ˈfut, bɔːl, -bɔl] football

- **Secondary stress** [,]
  - [ˈvərəˌleɪt] violate
  - [ɪmˌpɛˈnɪstɪk] impressionistic
APPENDIX VI

INVENTORY OF CANADIANISMS RELATING TO CLIMATE

A French-English

<table>
<thead>
<tr>
<th>French</th>
<th>English</th>
</tr>
</thead>
<tbody>
<tr>
<td>autoneige nf</td>
<td>snowmobile</td>
</tr>
<tr>
<td>balise nf</td>
<td>winter road marker</td>
</tr>
<tr>
<td>baliser vtr</td>
<td>to mark a winter road with posts</td>
</tr>
<tr>
<td>banc de neige nm</td>
<td>snowbank, snowdrift; sastrugi</td>
</tr>
<tr>
<td>bannot, banneau nm</td>
<td>sled for transporting snow</td>
</tr>
<tr>
<td>banquise nf</td>
<td>ice bank, ice field, ice pack, pack ice</td>
</tr>
<tr>
<td>barlot, berlot nm</td>
<td>snow sled (on skis)</td>
</tr>
<tr>
<td>berline, barligne, barline, borline nf</td>
<td>winter carriage</td>
</tr>
<tr>
<td>blizzard noir nm</td>
<td>black blizzard (dust storm)</td>
</tr>
<tr>
<td>bordages nmpl</td>
<td>bank ice, ice foot</td>
</tr>
<tr>
<td>bordée (de neige) nf</td>
<td>snowfall, blizzard</td>
</tr>
<tr>
<td>botte de neige nf</td>
<td>snowboot</td>
</tr>
<tr>
<td>boueige nf</td>
<td>slushy snow, slush</td>
</tr>
<tr>
<td>bourguignon, bourdignon nm</td>
<td>ice-shove</td>
</tr>
<tr>
<td>bourrasse nf</td>
<td>snow squall, flurry</td>
</tr>
<tr>
<td>bousculeil nm</td>
<td>ice scattered on lake or river</td>
</tr>
<tr>
<td>bousiller vtr &amp; intr</td>
<td>by wind, tide and current</td>
</tr>
<tr>
<td>canot nm</td>
<td>to fill in holes in a winter road</td>
</tr>
<tr>
<td>canot à glace nm</td>
<td>snow ridge, cahot</td>
</tr>
<tr>
<td>carnaval d'hiver nm</td>
<td>ice boat, ice canoe</td>
</tr>
<tr>
<td>carriole nf</td>
<td>winter carnival</td>
</tr>
<tr>
<td>charrue à neige nf</td>
<td>snow sleigh (on skis), carryall</td>
</tr>
<tr>
<td>cloture à neige nf</td>
<td>snow plough</td>
</tr>
<tr>
<td>charnière nf</td>
<td>snow fence</td>
</tr>
<tr>
<td></td>
<td>winter morgue (for storage of corpses till ground can be dug)</td>
</tr>
</tbody>
</table>
chasse de neige *nm*  
chasse d'hiver  
chemin d'hiver *nm*  
chemin boulant  
chemin croche  
chemin houleux  
chemin moulineux  

chino"ok, chinouque *nm*  
claireur *nm*  
compagnon d'hivarnement *nm*  
cor"ee des chemins *nf*  
couteau à neige *nm*  
craque *nf*  [Acadie]  
crâmi *nm*  
cro"ûte *nf*  
croûté *adj*  
cuter *nm*  
dé"barris *nm*  
se décalotter *vpron*  [Acadie]  
décles *nm*  
dérive *nf*  [Acadie]  

dériver  *vintr*  [Acadie]  
doux-temps *nm*  
écaille de neige *nf*  [Iles de la Madeleine]  
em"âcle *nm*  
s'embourber *vpron*  
s'encabaneur *vpron*  

drifting snow  
winter hunt  
winter road (made of packed snow)  
winter snow track  
winter snow road (for vehicles equipped with skis)  
rough surfaced snow road  
winter road which is difficult to negotiate because of rough surface  

chinook (wind)  

[person] snow clearer (of roads)  
wintering partner  
snow clearing bee (roads)  
snow knife  

ice crevasse  

soft snow  
hard crust formed on snow layer  
snow covered with crust  
cutter (sleigh)  

shore ice  

weather to clear  

ice-free pond  
snow embankment on either side of winter road  
to slide from one side to the other on a winter road  
thaw  

fall of fine, dry snow; flurry  

ice jam  
to become caught in the snow  
to move from summer to winter quarters
fleurette de neige nf
foudrille nm [Acadie]
foudrillement nm [Acadie]
foudroyer vinit [Acadie]
foulange nm [Acadie]

frasil, frazil nm

froid noir nm
galot nm
gazon de glace nm
giboulée de mare nf [Acadie]
glace bleue nf
glace boueuse
glace de fond
glace de rivage
glace en chandelles
glace fondante
glace /mouvante, cassée/ [Acadie]
glace pourrie
glace vive
glaces flottantes
gelure nf
glissière nf
grand-sleigh nm/f
gratte nf

grignon, grignot nm
haleine de vent nf [Acadie]
hivarnante, hivernante nf
hivernant nm

fresh snow, first snowfall
violent snowstorm; blizzard
[snow] to swirl, drift
sea foam on shore; mixed
snow and salt water
frazil ice; needle ice; floating
ice fragments or crystals; brash
cold snap
frozen chunk of earth on road
floating ice on rivers
mixed snow and rain (large flakes)
blue ice
slush hole
bottom ice, anchor ice
shore ice; ice shelf
candle(d) ice
slob ice
loose ice
blue ice
glare ice
drift-ice
icy mist on windows
ice runway
forestry sleigh
[Acadie] beam dragged sideways
by a horse to clear snow road;
[Quebec] tool for removing snow
from between railway lines
chunk of frozen mud or earth
squall
snow trapped beneath the soil
wintering partner
hivarnement nm

hivarner vintr
houle de neige nf [Acadie]
hummock nm
iglou, igloo nm
ile de glace nf [Acadie]
jarces nfpl [Acadie].
laid (il fait) [Acadie]

lisse nf

motoluge nf
" faire de la
motoneige nf
motoneigiste nm
motté de neige nf
motter vintr

neige collante nf [Acadie]
neige mouvante [Acadie]
neige poudreuse
neigailler vintr
neigasser vintr
ouvrir les chemins vtr
paumelle nf [Acadie]
peau de carriole nf
pêche sous la glace nf
pêcheur sous la glace nm
peloter vintr
pergelisol nm
piste d'hiver nf
pont (de glace) nm

wintering (of animals in barns, of ships in dry dock); winterization
to winter over large snow bank
hummock (of ice)
igloo, snowhouse
iceberg
ice crevasses
the weather is bad; cf il fait beau
metal slide on sleigh; metal lining on sleds
snowmobile
to snowmobile, go snowmobiling
snowmobile
snowmobiler
snowball
[snow] to form into chunks or balls of snow
soft snow
soft, melting snow
powdered snow
to snow lightly
to snow lightly
to clear winter roads
sleigh ski
fur blanket for sleigh use
ice fishing, ice angling
ice fisherman
[snow] to form into balls, chunks
permafrost
winter lumbering road
ice bridge
poste d'hivernement nm 
wintering quarters
poudraille nf [Acadie]
light and brief snow squall
poudrailler, poudriller,
poudroyer vintr [Acadie] [snow] to swirl, whirl, drift
powdered snow
poudre nf 
snowstorm of moderate violence
poudrement nm [Acadie] (between "poudraille" and
"poudrerie")
poudrer vintr [snow, water] to swirl, fly
about
drifted snow; snowstorm, blizzard, snowdrift
poudrerie nf white-out
poudre noire nf stormy
poudreux adj light snowstorm
poudrin nm [Acadie] boat ice-bound; horse
pris (dans la glace) adj bogged down (in snow)

quartiers d'hiver nmpl wintering grounds
rampart nm [Acadie] ice shelf left along shore after
breakup
raquette nf raquette, snowshoe
to snowshoe
raquettes (aller, voyager en)
raquetteur nm
raquetter vintr
renchausser vtr
robe de carriole nf
fur blanket for sleigh
roue de neige nf [Acadie]
small heap of snow
roulis de neige nm [Acadie] piled up snow; dunes
route d'hiver nf winter road
saignée nf [Acadie] gaping crack in ice
skidoo nm
skidoo, snowmobile
sleigh nm/f
sleigh to travel on snowshoes
sleigh à billots to winterize (a house)
forestry sleigh
sloche nm
sorcière nf
souffleuse nf
suisse nm
suisse plat
tempête de neige nf
tombée nf
torchon de neige nf
tracée des chemins nf
traîne sauvage nf
traîne /plate, à sommier,
à bâtons/
tuque nf
ventre-de bœuf nm

sluch, sluchy snow
twister
snow-blower
agricultural sleigh
working sleigh with broad skids
snowfall, snowstorm
fall, Autumn
piles of snowflakes
the opening of winter roads
sled, toboggan
sled, toboggan
tuque, toque
frost-boil

B English-French

anchor ice n
banana belt n
black blizzard n

black ice n
black snow n
blizzard n
blue ice n
bordage n
bottom ice n
brash n
break-up, breaking-up n
cahot n
candle(d) ice n
chinook (wind) [West]
clinkerbell n [Nfld]
cold snap n
glace de fond
le sud, les régions tempérées
poudrerie noire, blizzard noir,
orage de poussière
verglas
cf black blizzard
poudrerie, blizzard
glace /pourrie, bleue/
bordages; cf ice shelf
glace de fond
frasil
débâcle (des glaces)
cahot
glace en chandelles
chinouque, chinook
glaçon
vague de froid, froid noir
crust n
debacle n
drift n
drift storm n
drifting snow n
drift-ice n
dwye n Nfld
fall n
frazil, frazzle n
frost-boil n
glare ice n
glidder, glitter n [Nfld]
ground drift n
heap of snow n
hummock n
ice-bank n
ice bear n
ice blink n
ice boat n
ice bridge n
ice canoe n
ice creeper n
iced up adj
ice field n
ice fisherman n
ice fishing, angling n
ice-fishing hut n
ice foot n
ice hockey n
ice house n
ice hunter n [Nfld]
ice jam n
ice-pack n
croûte de neige
bourdigueau, bourdignon
banc de neige
poudrerie, tempête de neige
chasse de neige, poudrerie
glaces flottantes
tempête de neige
tombée, automne
frazil, frazil
ventre-de bœuf
glace vive
gelée blanche
chasse de neige
banc de neige
hummock (de glace)
banquise
ours polaire
clarté des glaces
canoë à glace
tout de glace
canot à glace
givre; pris dans la glace
banquise
pêcheur sous la glace
pêche sous la glace
cabane de pêche
bordage
hockey sur glace
iglou, igloo
[seals] chasseur de phoques
embâcle
embâcle (sur un fleuve, une rivière
ice road n
dice run n
ice runway n
ice-shelf n
ice-shove n
ice storm n
jam n
jam vintr
mush snow' n
needle ice n
nipped in adj
open water n
pack (ice) n
pancake ice n
permafrost n
powdered snow n
powdering n
raquette n
rubber ice n
sastrugi n
shore-ice n
silver thaw n  B.C.
slob (ice) n  Nfld
slush n
slushy adj
slush hole n
snowbank n
snow blindness n
snow blower n
snow boot n
snow bunting n
snow-crust n
pont de glace
glissière
pergelisol, permafrost, permagel
poudre, neige poudreuse
poudrerie
raquette
glace mince (des lacs et rivières)
(gros) banc de neige
glace de rivage
glace fondante
sloche, boue
trempé de neige
glace boueuse
banc de neige
cécité des neiges, mal de neige
souffleuse (de neige)
botte de neige
bruant des neiges; oiseau /blanc,
de neige/
croûte
snowed in adj
snow fence n

snow drift

snowfall n
snowglasses npl
snow goggles npl
snow goose n
snow house n
snow knife n
snowmobile n

snowmobile vintr
snowmobiler n
snowplough n

snowshoe n
snowshoe vintr
snowshoer n
snow storm n
snowy owl n
sundown n
twister n
whirling snow n
white-out n
winter vintr
winter carnival n
winterer n
winter hunt n
wintering grounds npl
wintering partner n

wintering post n

bloqué, enseveli/ dans la neige
paraneige, taillis en paraneige
clôture à neige
banc de neige, poudrerie,
congère
bordée de neige, chute de neige
lunettes de neige
lunettes de neige
oise blanche
iglou, igloo
couteau à neige
autoneige, skidoo, motoneige,
motoluge
faire de la motoluge
motoneigiste
charrue à neige, chasse-neige;
gratte
raquette
aller, voyager/ en raquettes
raquetteur
tempête de neige, poudrerie
harfang des neiges, hibou blanc
coucher du soleil, brunant
sorcière
neige poudrante, poudrerie
poudrerie
hiverner
carnaval d'hiver
hivernant, homme du nord
chasse d'hiver
quartiers d'hiver
compagnon d'hivernement, hiver-
nant
poste d'hivernement
winterkill  vintr  mourir de, froid
winter road  n  route, chemin/ d'hiver; [bois]
piste d'hiver.
The usual internationally recognized abbreviations are used in the Bibliography for learned journals and linguistic organizations. In addition, the following shortened forms have been used for frequently recurring publications:


Lexicologie et lexicographie, 1961.


Probleme der Semantik, 1968.


Problems in Lexicography, 1967.

1. BIBLIOGRAPHICAL SOURCES.


2. GENERAL LEXICOGRAPHY.


Barnhart, Clarence L. "Contributions od Fr. Thorndike to Lexicography." Teachers' College Record, 51 (1949), 35-42.


Bulletin d'information du Laboratoire d'analyse lexicographique. Besançon: Univ. de Besançon, 1959-.

Bulletin du Centre de recherche pour un Trésor de la langue française. Nancy, Centre national recherche scientifique.


Drysdale, Patrick D. "Dictionaries for Schools." *Education* 4, (1962), 105-119.


---------. "What is a Dictionary?" *Conference on Lexicography,* 1970.


Guiraud, Pierre. "La place de l'argot et de la langue populaire dans le Thesaurus de la langue française". In Lexicologie et lexicographie, 1961, 177-184.


Hatzfeld, H. "Lexicography and Stylistics". In French and Provencal Lexicography, 1964, 13-29.


Imbs, Paul. "La place du vocabulaire ancien dans un Thesaurus de la langue française." In Lexicologie et lexicographie, 1961, 133-139.


--------. "Problèmes de lexicographie." Revue belge de philologie et d'histoire, 43 (1965), 847-848.


Leksikograficeskiy sbornik. Moscow: Institut jazykoznaniya, Akademija nauk SSSR, 1957-. [Series of 6 volumes containing papers on lexicography].


McMillan, James B. "Five College Dictionaries." College English, 10 (1949), 214-221.


---------. "Dictionaries and the English Language." English Journal, 52 (1963), 336-345.


--------. "Review of La lexicographie." Language, 49 (1973), 961-967.


--------. "Le dictionnaire historique de la langue italienne." In Lexicologie et lexicographie, 1961, p. 41.


Mower, Morris Leon and Barney De Roy. "Which are the Most Important Dictionary Skills?" *Elementary English*, 45 (1968), 468-471.


"Dictionnaire de la langue française d'Emile Littré, abrégé par A. Beaujean, révision et mise à jour de G. Venzac." Compte-rendu in Zeitschrift für Romanische Philologie (1967), Band 83, Heft 1/2, 143-163.

"Datations nouvelles recueillies pour le dictionnaire Le Petit Robert." Le français moderne, 36, No. 3 (1968), 226-231.

"A propos des dictionnaires français." Le français dans le monde, 58 (1968), 13-16.

"Les bases théoriques de la description lexicographique du français; tendances actuelles." Travaux de linguistique et de littérature, 6, No. 1 (1968), 55-72.


"Le dictionnaire comme discours sur la chose et discours sur le signe." Semiotica, 1; No. 2 (1969), 185-195.


Sabrsula, J. "A propos de quelques problèmes linguistiques de la lexicographie. (L'aspect verbal, l'ordre du procès et les dictionnaires)." Romanistica Pragensia, 3; No. 3 (1963), 87-103.


Togeby, K. "Qu'est-ce qu'un mot?" Travaux du cercle linguistique de Copenhague, 5 (1959), 99-111.

---------. "La documentation lexicographique et ses besoins." Cahiers de lexicologie, 3 (1962), 101-118.


Vidos, B. E. "La place du vocabulaire technique dans le Thesaurus de la langue française." In Lexicologie et lexicographie, 1961, 185-194.


Wagner, R-L. "Le dictionnaire de Huguet; l'inventaire de la langue française." In Lexicologie et lexicographie, 1961, 29-32.


3. **BILINGUAL LEXICOGRAPHY.**


Lutaud, O. "Translation, traduction, tradition." Cahiers de lexicologie, 13, No. 2 (1968), 49-59.


META. (Journal des traducteurs/Translators' Journal). Montréal: Presses de l'Université de Montréal, 1956-.


Swanson, Donald C. "Recommendations on the Selection of Entries for a Bilingual Dictionary." In Problems in Lexicography, 1967, 63-78.


------------. "La traduction littéraire, est-elle un genre à part?" META, 14, No. 1 (1969), 5-21.


------------. "Regards sur l'évolution des théories de la traduction depuis vingt ans." META, 22, No. 1 (1975), 7-27.


------------. "Analysis of the Problems of Meaning Discrimination in Spanish and English Bilingual Lexicography." Babel, 6, No. 3 (1960), 121-125.


4. LEXICOLOGY AND LEXICOGRAPHY.


----------. "Distribution, ensemble et marque dans le lexique." *Cahiers de lexicologie*, 4 (1964), 5-16.


----------. "Le problème de la migration des mots d'un champ conceptuel à l'autre." *Lingua*, 10 (1961), 57-78.

----------. "L'homonymie et la polysémie." *Vox Romana*, 21

----------. "La structure du lexique et quelques problèmes sémantico-lexicaux." Revue romane de linguistique, 10 (1965), 559-569.


----------. "La signification et sa manifestation dans le discours." Cahiers de lexicologie, 5 (1964), 17-27.


"Les structures élémentaires de la signification." B.S.L.P., 60 (1965), 97-114.


Kohler, E. Reflexions sur le problème sémantique dans les langues romanes." In Lexicologie et lexicographie, 1961, 81-84.


Lounsbury, F. G. "The Varieties of Meaning." Georgetown University Monograph Series, 8 (1955), 158-164.


----------. "Les dimensions dans le vocabulaire d'aujour- 
d'hui." Journal de psychologie normale et pathologique, 

----------. "Le vocabulaire contemporain et l'espace."

Meschonnic, H. "Essai sur le champ lexical du mot idée." 
Cahiers de lexicologie, 5 (1964), 57-68.

Mitterand, H. "Deux dictionnaires français: Le Petit 
Robert et le Dictionnaire du français contemporain." 
Le français dans le monde, 59 (1968), 24-29.

Mounin, Georges. "Les analyses sémantiques." Cahiers de 
l'Informatique. Sciences économiques appliquées, Série 
M, 3 (1962), 105-124.

----------. "Essai sur la structuration du lexique de 
l'habitation." Cahiers de lexicologie, 6 (1965), 
9-24.

----------. "Un champ sémantique: la dénomination des animaux 
domestiques." La linguistique, 1 (1965), 31-54.

Muller, Ch. "Polysémie et homonymie dans l'élaboration du 
lexique contemporain." Etudes de linguistique 
n appliquée, 1 (1962), 49-54.

Nagy, G. O. "The Function of Lexemes and Synonymy. The 
Base of a Linguistically Oriented Dictionary of 

Neuhaus, H. J. "Towards a Diachronic Analysis of Vocabulary." 
Cahiers de lexicologie, 18, No. 1 (1971), 29-42.

Nida, Eugene A. "Semantic Components." Babel, 8 (1962), 
175-181.

Pietrzyk, A. "Notes on Lexical Problems in Linguistic 

Phal, H. "Les groupes de mots." Cahiers de lexicologie, 
4 (1964), 45-60.

Pottier, Bernard. Recherches sur l'analyse sémantique en 
linguistique et en traduction automatique. Publications 
linguistiques de la Faculté des Lettres et des Sciences 
humaines de Nancy, Série A: Linguistique appliquée 
et traduction automatique, 2, 1963.
"Du très grand au trop particulier en analysé linguistique." Travaux de linguistique et de littérature, 1 (1963), 9-16.

"Vers une sémantique moderne." Travaux de linguistique et de littérature, 3, No. 1 (1964), 107-137.


Prator, C. N. "Adjectives of Temperature." English Language Teaching, 17 (1963), 158-164.


"Les relations entre le signe et la chose dans le discours métalinguistique." Travaux de linguistique et de littérature, 7, No. 1 (1969), 113-129.


Saareste, A. "Quelques remarques sur le rapport entre la


Sokolova, G. "Remarques sur la conversion en français moderne." Cahiers de lexicologie, 7, No. 2 (1965), 51-64.


Togeby, K. "Grammaire, lexicologie et sémantique." Cahiers de lexicologie, 6 (1965), 3-7.

----------. "Qu'est-ce qu'un mot?" Travaux du Cercle linguistique de Copenhague, 5 (1959), 97-110.


"Le point de vue structurale en sémantique historique." Lexicologie et lexicographie, 1961, 85-86.


Vidos, B. E. "La place du vocabulaire technique dans le Thesaurus de la langue française." In Lexicologie et lexicographie, 1961, 185-194.


Watermann, J. T. "B. L. Worth and Linguistic Field Theory." Southwest Journal of Anthropology, 13, No. 3 (1957), 201-211.


5. **GRAMMAR IN LEXICOGRAPHY.**


Gates, J. Edward. "The Treatment of Lexemes Larger than the Word." *Studies in Lexicography as a Science and as an*
Art. Papers made available in mimeograph form by the University of Delaware Department of Modern Languages and Literature, Newark, 1977, 37-54.


------------. "'Gerund/Noun' and 'Participle/Adjective'." American Speech, 40 (1965), 40-46.


Mantchev, K. "Hiérarchie sémantique des verbes français contemporains." Cahiers de lexicologie, 10 (1967), 31-46.


Prator, C. N. "Adjectives of Temperature." English Language Teaching, 17 (1963), 158-164.


6. PRONUNCIATION IN LEXICOGRAPHY.


Straka, G. "La transcription phonétique de la prononciation dans les dictionnaires." In Lexicologie et lexicographie, 1961, 123-129.
7. REGISTER IN LEXICOGRAPHY.


Monson, Samuel C. "Restrictive Labels--Descriptive or Prescriptive?" In *Lexicography in English*, 1973, 208-212.


8. REGIONAL VARIATION.


Baldinger, K. "L'importance du vocabulaire dialectal dans un Thesaurus de la langue française." In Lexicologie et lexicographie, 1961, 149-173.


Deak, E. "Y a-t-il une langue américaine?" Babel, 6, No. 2 (1960), 68-71.


Duckert, Audrey R. "Regional and Social Dialects." In Lexicography in English, 1973, 51-54.


Harris, Barbara P. "Selected Political, Cultural and Socio-economic Areas of Canadian English as Contributors to the Vocabulary of Canadian English." Unpublished Ph. D. dissertation, Victoria, Univ. of Victoria, B. C., 1975. [See in particular Bibliography No. 2: "Canadian English"].


---------. "Canadianisms from Western Canada, with Special Reference to British Columbia." Transactions of the Royal Society of Canada, Series 4, No. 6 (1968), 181-185.


Vinay, Jean-Paul. "Le français au Canada I. Répartitions des groupes francophones." La classe de français, 5, No. 7 (1955), 312-317.

---------. "Le français au Canada II. Phonétique et vocabulaire." La classe de français, 6, No. 6 (1956), 353-361.

---------. "Le français en Amérique du Nord: problèmes et
réalisations." In Current Trends in Linguistics, 10 (1973), 323-406. [Note in particular the important bibliography following the article, pages 441-463].


9. HISTORICAL LEXICOGRAPHY.

Bousser, M. "Les différentes conceptions lexicographiques de Littré à nos jours, d'après les préfaces de dictionnaires français." Paris, 1971 (mimeographed).

Barnhart, Clarence L. "Contributions of Dr. Thorndike to Lexicography." Teachers' College Record, 51 (1949), 32-42.


10. COMPUTATIONAL LEXICOGRAPHY.


Grimes, Joseph E. "Computing in Lexicography." The Linguistic Reporter, 12, Nos. 5-6 (1970), 1-10.


"The Canadian Dictionary--An Experiment in Automated Processing." In Vingt-cinq ans de linguistique au Canada: Hommage à Jean-Paul Vinay. (forthcoming)