

Romain Gary/Émile Ajar:

Étude stylistique d'une fraude littéraire

par

Vina Tirvengadam

Thèse

**déposée à la faculté d'études supérieures
remplissant en partie les exigences
du diplôme de**

Doctorat

Département de français, d'espagnol et d'italien

Université du Manitoba

Winnipeg, Manitoba

Octobre 1999



National Library
of Canada

Acquisitions and
Bibliographic Services

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale
du Canada

Acquisitions et
services bibliographiques

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

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

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

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

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

0-612-51673-3

**THE UNIVERSITY OF MANITOBA
FACULTY OF GRADUATE STUDIES

COPYRIGHT PERMISSION PAGE**

**Romain Gary/Émile Ajar:
Étude stylistique d'une fraude littéraire**

BY

Vina Tirvengadum

**A Thesis/Practicum submitted to the Faculty of Graduate Studies of The University
of Manitoba in partial fulfillment of the requirements of the degree
of
Doctor of Philosophy**

VINA TIRVENGADUM ©1999

Permission has been granted to the Library of The University of Manitoba to lend or sell copies of this thesis/practicum, to the National Library of Canada to microfilm this thesis and to lend or sell copies of the film, and to Dissertations Abstracts International to publish an abstract of this thesis/practicum.

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

TABLE DES MATIÈRES

	Page
ABSTRACT	xv
REMERCIEMENTS	xvii
LISTE DES GRAPHIQUES	xx
LISTE DES TABLES	xxi
ABRÉVIATIONS	xxvi
GLOSSAIRE	xxix
CHAPITRE 1: L'ILLUSION D'ÊTRE SOI	1
1.1 Méthodologie	5
1.2 Étude de style	6
1.3 Organisation du travail	8
CHAPITRE 2: IL ÉTAIT UNE FOIS ROMAIN GARY	10
2.1 <i>La Promesse de l'aube</i>	12
2.2 Les Années formatrices à Wilno et à Varsovie	13
2.3 Adolescence: La France	18
2.4 Carrière militaire	20

2.5	<i>La Forêt de la colère</i>	23
2.6	Débuts diplomatiques	26
2.7	<i>Les Racines du ciel</i>	31
2.8	Expériences à Los Angeles	36
2.9	Le Consul sexuel	40
2.10	Vie commune	44
2.11	Le Solitaire à travers le monde	50
2.12	Déclin et mort	55
CHAPITRE 3: LES ROMANS ANALYSÉS		61
3.1	<i>Au-delà de cette limite votre ticket n'est plus valable</i>	61
3.2	<i>Clair de femme</i>	65
3.3	<i>Gros-Câlin</i>	67
3.4	<i>La Vie devant soi</i>	69
3.5	Le Choix des textes	73
CHAPITRE 4: L'AFFAIRE ÉMILE AJAR OU "JE EST UN AUTRE"		75
4.1	La Raison profonde du faux	76
4.2	Pierre Michaud et <i>La Solitude du python</i>	81
4.3	<i>La Tendresse des pierres</i>	85

4.4	Paul Pavlowich	93
4.5	Un Extraordinaire ballet de masques: <i>Pseudo</i>	95
4.6	Revendication de la vieillesse et des racines juives	101
4.7	Retentissement dans le milieu littéraire	104
CHAPITRE 5: ROMAIN GARY ET ÉMILE AJAR: RESSEMBLANCES		107
5.1	Les Ressemblances linguistiques	108
5.2	Thèmes et leitmotive communs aux oeuvres de Gary et d'AJar	110
5.3	Dépassement de soi et dédoublement	113
5.4	Le Dibbouk	116
5.5	Personnages communs aux oeuvres de Gary et d'AJar	123
5.6	Les Marginaux	126
5.7	Nina Kacew et Madame Rosa	128
5.8	Motifs communs aux oeuvres de Gary et d'AJar	131
5.9	Le Motif du parapluie et du petit caniche	134
CHAPITRE 6: LA MYSTIFICATION LITTÉRAIRE EN EUROPE ET EN FRANCE		137
6.1	Historique de la mystification	138

6.2	Le Pseudonyme	139
6.3	Les Auteurs apparents et les auteurs anonymes	142
6.4	L'Apocryphe	144
6.5	La Supposition d'auteur	145
6.6	L'emprunt et les auteurs imaginaires	147
6.7	Les Auteurs hétéronymes, cryptonymes et polynomes	150
6.8	Histoire de la mystification littéraire	151
6.9	Le Dix-neuvième Siècle: anoblissement des roturiers	155
6.10	Une Amusante Supercherie: <i>Clara Gazul</i>	157
6.11	Les Mystificatrices	160
6.12	La Mystification Ajar	162
6.13	La Raison profonde de la mystification	167
	CHAPITRE 7: L'ANALYSE STATISTIQUE DU STYLE	170
7.1	Les Précurseurs	172
7.2	Index, dictionnaires des fréquences, concordances etc.	178
7.3	<i>L'Élaboration du français fondamental et Le Français fondamental</i>	182
7.4	<i>Le Trésor de la langue française</i>	186

7.5	Gunnell Engwall: <i>Fréquence et distribution du vocabulaire dans un choix de romans français</i>	189
-----	---	-----

CHAPITRE 8: ATTRIBUTION D'AUTEURS AUX OEUVRES ANONYMES ET PSEUDONYMES

		194
8.1	La Méthode traditionnelle	194
8.2	La Longueur des vocables	196
8.3	Le Nombre de syllabes par vocable	202
8.4	La Longueur des phrases	203
8.5	La Répartition des éléments grammaticaux	207
8.6	La Fréquence des lettres	210
8.7	L'Analyse des "vocables distinctifs"	214
8.8	La Fréquence des vocables dans un corpus	216
8.9	John Burrows et l'Analyse des vocables les plus fréquents	218
8.10	La Fréquence des synonymes	223
8.11	Conclusions	224

CHAPITRE 9: QUELQUES NOTIONS FONDAMENTALES SUR LA STATISTIQUE

9.1	Analyse statistique de l'oeuvre littéraire	228
-----	--	-----

9.2	<i>Mot, vocable, lemme, vocabulaire, lexique, type, occurrence</i>	229
9.3	La Statistique	231
9.4	La Statistique dans le domaine littéraire et linguistique	233
9.5	La Nature statistique du langage	235
9.6	La Quantification du langage	237
9.7	La Moyenne	238
9.8	La Variance (v) et l'écart-type (σ)	241
9.9	La Distribution gaussienne	244
9.10	Le Test- t de Student	247
9.11	Le Test du khi-deux (χ^2)	249
9.12	Le Coefficient de corrélation de Pearson	255
9.13	Le r^2 (Le Coefficient de détermination) et les limites de confiance de "r"	262
9.14	Conclusions en matière de corrélation	263
	CHAPITRE 10: ANALYSES PRÉLIMINAIRES	264
10.1	Méthodologie	265
10.2	Traitement des données	266
10.3	Constitution du corpus	269

10.4	Les Soixante types les plus fréquents dans les huit romans et dans le corpus d'Engwall	270
10.5	Constitution d'une liste unique des soixante types les plus fréquents dans les neuf textes	273
10.6	Les test- <i>t</i> de Student/Fisher	279
10.7	Les Limites de confiance	281
CHAPITRE 11: LE TEST DU χ^2		284
11.1	La Valeur théorique représentée par le corpus d'Engwall	285
11.2	La Valeur théorique représentée par <i>Clair</i> et <i>Ticket</i>	288
11.3	La Table de contingence	291
11.4	Comparaison entre les huit romans (abstraction faite du corpus d'Engwall)	294
11.5	Les Quatre Romans de Gary/Ajar	295
11.6	Conclusions	297
CHAPITRE 12: LE COEFFICIENT DE CORRÉLATION DE PEARSON		299
12.1	Corrélation entre les oeuvres	300
12.2	Les Textes de Gide	304
12.3	Les Textes de Gary/Ajar	305

12.4	Ce que révèlent les corrélations et les limites de confiance de r	310
------	--	-----

CHAPITRE 13:	LA FRÉQUENCE DES SYNONYMES COMME INDICE DE STYLE	312
--------------	---	-----

13.1	La Répartition des synonymes	313
13.2	Le Test du χ^2 pour les neuf textes	314
13.3	Le Test du χ^2 pour les huit romans	317
13.4	Le Test du χ^2 pour <i>L'Immoraliste</i> et <i>La Porte étroite</i>	318
13.5	Le Test du χ^2 pour <i>Clair</i> et <i>Ticket</i> , <i>Câlin</i> et <i>La Vie</i>	319
13.6	Conclusions	322

CHAPITRE 14:	CONCLUSION	324
--------------	------------	-----

14.1	Ce que révèlent les tests	325
14.2	Explications	327

APPENDICES AU CHAPITRE 10	330
---------------------------	-----

10.1.1	<i>Clair de femme</i>	330
10.1.2	<i>Au-delà de cette limite votre ticket n'est plus valable</i>	331
10.1.3	<i>Gros-Câlin</i>	332

10.1.4	<i>La Vie devant soi</i>	333
10.1.5	<i>L'Étranger</i>	334
10.1.6	<i>Le Noeud de vipères</i>	335
10.1.7	<i>L'Immoraliste</i>	336
10.1.8	<i>La Porte étroite</i>	337
10.1.9	Le Corpus d'Engwall	338
10.2	Les Types les plus fréquents dans les huit romans et dans le corpus d'Engwall	339
10.3	Les Neuf Textes: limites de confiance à 95%	341
10.4	Les Neuf Textes: limites de confiance à 99%	343
APPENDICES AU CHAPITRE 11		345
11.1	Les Neuf Textes: fréquences observées	346
11.2	Les Huit Romans: valeurs hypothétiques	347
11.3	Les Huit Romans: χ^2 (Engwall fournit la valeur calculée)	349
11.4	Les Huit Romans: χ^2 (abstraction faite des marqueurs de personne)	350
11.5	Les Six Romans: valeurs calculées (<i>Clair</i> et <i>Ticket</i> représentent la valeur hypothétique)	351

11.6	Les Six Romans: χ^2 (<i>Clair</i> et <i>Ticket</i> représentent la valeur théorique)	352
11.7	Les Neuf Textes: Valeurs calculées	354
11.8	Les Neuf textes: χ^2	356
11.9	Les Neuf Textes: Valeurs calculées (abstraction faite des marqueurs de personne)	357
11.10	Les Neuf Textes: χ^2 (abstraction faite des marqueurs de personne)	358
11.11	Les Huit Romans: Valeurs calculées	359
11.12	Les Huit Romans: χ^2	361
11.13	Les Huit Romans: Valeurs calculées (abstraction faite des marqueurs de personne)	362
11.14	Les Huit Romans: χ^2 (abstraction faite des marqueurs de personne)	363
11.15	Les Quatre Romans de Gary/Ajar: Valeurs calculées	364
11.16	Les Quatre Romans de Gary/Ajar: χ^2	366
11.17	Les Quatre Romans de Gary/Ajar: Valeurs calculées (abstraction faite des marqueurs de personne)	367
11.18	Les Quatre Romans de Gary/Ajar: χ^2 (abstraction faite des marqueurs de personne)	368

APPENDICES AU CHAPITRE 12	370
12.1 Limites de confiance de r pour la corrélation entre <i>Clair de femme</i> et les autres oeuvres (lorsque $n = 60$)	370
12.2 Limites de confiance de r pour la corrélation entre <i>Clair de femme</i> et les autres oeuvres (lorsque $n = 31$)	370
12.3 Limites de confiance de r pour la corrélation entre <i>Au-delà de cette limite votre ticket n'est plus valable</i> et les autres oeuvres (lorsque $n = 60$)	371
12.4 Limites de confiance de r pour la corrélation entre <i>Au-delà de cette limite votre ticket n'est plus valable</i> et les autres oeuvres (lorsque $n = 31$)	371
12.5 Limites de confiance de r pour la corrélation entre <i>Gros-Câlin</i> et les autres oeuvres (lorsque $n = 60$)	372
12.6 Limites de confiance de r pour la corrélation entre <i>Gros-Câlin</i> et les autres oeuvres (lorsque $n = 31$)	372
12.7 Limites de confiance de r pour la corrélation entre <i>La Vie devant soi</i> et les autres oeuvres (lorsque $n = 60$)	373
12.8 Limites de confiance de r pour la corrélation entre <i>La Vie devant soi</i> et les autres oeuvres (lorsque $n = 31$)	373

12.9	Limites de confiance de r pour la corrélation entre <i>L'Étranger</i> et les autres oeuvres (lorsque $n = 60$)	374
12.10	Limites de confiance de r pour la corrélation entre <i>L'Étranger</i> et les autres oeuvres (lorsque $n = 31$)	374
12.11	Limites de confiance de r pour la corrélation entre <i>Le Noeud de vipères</i> et les autres oeuvres (lorsque $n = 60$)	375
12.12	Limites de confiance de r pour la corrélation entre <i>Le Noeud de vipères</i> et les autres oeuvres (lorsque $n = 31$)	375
12.13	Limites de confiance de r pour la corrélation entre <i>L'Immoraliste</i> et les autres oeuvres (lorsque $n = 60$)	376
12.14	Limites de confiance de r pour la corrélation entre <i>L'Immoraliste</i> et les autres oeuvres (lorsque $n = 31$)	376
12.15	Limites de confiance de r pour la corrélation entre <i>La Porte étroite</i> et les autres oeuvres (lorsque $n = 60$)	377
12.16	Limites de confiance de r pour la corrélation entre <i>La Porte étroite</i> et les autres oeuvres (lorsque $n = 31$)	377
12.17	Limites de confiance de r pour la corrélation entre Engwall et les autres oeuvres (lorsque $n = 60$)	378
12.18	Limites de confiance de r pour la corrélation entre Engwall et les autres oeuvres (lorsque $n = 31$)	378

APPENDICES AU CHAPITRE 13	379
13.1.1 <i>Clair de femme</i> (33.755 vocables)	379
13.1.2 <i>Au-delà de cette limite votre ticket n'est plus valable</i> (57.915 vocables)	379
13.1.3 <i>Gros-Câlin</i> (53.264 vocables)	380
13.1.4 <i>La Vie devant soi</i> (65.095 vocables)	380
13.1.5 <i>L'Étranger</i> (34.623 vocables)	381
13.1.6 <i>Le Noeud de vipères</i> (59.426 vocables)	381
13.1.7 <i>L'Immoraliste</i> (41.260 vocables)	382
13.1.8 <i>La Porte étroite</i> (39.953 vocables)	382
13.2.1 Les Neuf Textes: Valeurs observées et table de contingence	383
13.2.2 Les Neuf Textes: Valeurs calculées (aucune valeur inférieure à 5)	383
13.2.3 Les Neuf Textes: Valeurs du χ^2	384
13.3.1 Les Huit Romans: Valeurs observées et table de contingence	385
13.3.2 Les Huit romans: Valeurs calculées	385
13.3.3 Les Huit Romans: Valeurs du χ^2	386
13.4.1 <i>Porte</i> et <i>Immoraliste</i> : Valeurs observées et table de contingence	387

13.4.2	<i>Porte et Immoraliste</i> : Valeurs calculées	388
13.4.3	<i>Porte et Immoraliste</i> : Valeurs du χ^2	388
13.5.1	<i>Clair et Ticket</i> : Valeurs observées et table de contingence	388
13.5.2	<i>Clair et Ticket</i> : Valeurs calculées	389
13.5.3	<i>Clair et Ticket</i> : Valeurs du χ^2	389
13.6.1	<i>Câlin et Vie</i> : Valeurs observées et table de contingence	390
13.6.2	<i>Câlin et Vie</i> : Valeurs calculées	390
13.6.3	<i>Câlin et Vie</i> : Valeurs du χ^2	391

BIBLIOGRAPHIE		392
I	Oeuvres de Romain Gary	392
II	Sous le pseudonyme d'Émile Ajar	393
III	Filmographie	394
IV	Romans de Gary/Ajar adaptés à l'écran	395
V	Ouvrages cités	406
VI	Ouvrages consultés	372

ABSTRACT

Theoreticians of style, from Buffon to Barthes, postulate that style is dictated by the author's personality and is unique to each writer. This thesis tests this hypothesis in light of the pseudonymous works of Romain Gary. It applies standard authorship attribution methods and evaluates their reliability.

In 1974 Gary, a recipient of the Prix Goncourt, published *Gros-Câlin* under the pseudonym of Émile Ajar and a year later, published *La Vie devant soi* under the Ajar name. Meanwhile, he also published two other novels *Au-delà de cette limite votre ticket n'est plus valable* (1975) and *Clair de femme* (1977) under his own name. When he received the Prix Goncourt for *La Vie devant soi*, he became the first author to achieve such a feat. After his suicide in 1980, Gary admitted in a posthumous confession that Ajar was his pseudonym.

Romain Gary thus provides an excellent example for authorship attribution study which is the analysis of stylistic idiosyncracies of an author as an index of authenticity.

Chapter 1 sets out the problem and details methodology used in the thesis. Chapters 2 to 4 give an account of Gary's life and describe the plot of each novel. Chapter 5 shows similarities between the Gary and Ajar novels and chapter 6 looks

at the history of literary fraud in France. Chapter 7 details the types of statistical work done on vocabulary distribution and chapter 8 describes the various methodologies used in the past in the field of authorship attribution. Chapter 9 defines statistical terms and approaches used in the thesis.

Chapters 10 to 13 deal with the statistical tests. They all show that Gary uses a similar style in novels written under his own name and in *Gros-Câlin*, but uses a different one in *La Vie devant soi*. The tests indicate that certain elements of style are not always unique to a given individual and that a writer can indeed disguise his style at will.

REMERCIEMENTS

Cette étude n'aurait pu être menée à bon terme sans l'appui continu de plusieurs personnes envers lesquelles je suis entièrement reconnaissante. Dans mes moments d'incertitude, ce travail prenait parfois l'allure d'une tâche sisyphienne, mais elle s'est enfin accomplie grâce à leur encouragement et leur bienveillance.

Je suis infiniment reconnaissante aux membres de mon comité, surtout aux professeurs Donna Norell et Isabelle Strong du département de français, qui ont bien voulu lire cette thèse, m'ont toujours donné de bons conseils et m'ont été un soutien précieux.

Je remercie également le professeur Kevin Keen du département de statistique pour sa grande expertise dans le domaine de la statistique, pour l'enthousiasme qu'il a exprimé pour ce projet, et pour ses excellentes suggestions. Je suis sincèrement reconnaissante au professeur William Winder du département de français à l'université de La Colombie Britannique, d'avoir accepté de faire partie de mon comité aussi bien que pour ses commentaires très utiles. Je dois aussi remercier le professeur Tony Szturm du département de "Medical Rehabilitation" de L'Université du Manitoba qui a bien voulu me permettre de traiter mes textes à l'aide d'un lecteur optique.

Au professeur Paul Fortier, mon directeur de thèse et mon maître à penser, je dois la plus sincère reconnaissance. Je le remercie d'avoir voulu diriger et approuver ce travail gigantesque. Dès mes premiers tâtonnements dans le domaine d'initiation aux méthodes de la statistique linguistique et stylistique, le professeur Fortier ne s'est jamais lassé de me guider de sa grande expertise et de m'offrir son aide précieuse. Il m'a toujours encouragée dans la voie vers laquelle je me suis lancée et continue toujours de le faire. Sous ses conseils, j'ai participé à plusieurs conférences nationales et internationales ayant trait à mon sujet. Je remercie également son épouse Penny Gilbert pour ses suggestions concernant plusieurs aspects de rédaction dans cette thèse.

Je dois un grand merci à plusieurs membres de ma famille: à ma mère qui m'a orientée dans le domaine des lettres et à (feu) mon père qui a inculqué en moi l'importance de l'investigation et du raisonnement rationnels. Je remercie également mes soeurs Mira et Amrita et mon frère Harrish qui ont toujours cru en moi, et mon ami Salem Fathi pour son encouragement formulé à distance.

À mon petit Jean-Luc (mon neveu) je dois la joie et le plaisir de sa présence dans ma vie pendant une grande partie de la rédaction de cette thèse. Durant mes moments de doute, il m'encourageait en me tenant des propos sur la statistique.

Mais c'est surtout envers mon fils Daren et mon mari Vijay que je suis le plus redevable. Sans l'aide précieuse de Daren, je n'aurais pu entreprendre cette

étude à dimensions interdisciplinaires. Il m'a initiée dans le domaine de l'informatique et m'a aidée à accomplir toutes les tâches nécessaires pour l'aboutissement de cette thèse. À Vijay, je dis un grand merci d'avoir toujours cru en moi, de m'avoir soutenue durant les moments difficiles et d'avoir pu tolérer mes nombreuses absences lorsque j'étais "attelée" à mon ordinateur.

LISTE DES GRAPHIQUES

	Page
10.1 Les Huit Romans et le corpus d'Engwall	275
10.2 <i>Clair, Ticket, Câlin et La Vie</i>	276
10.3 <i>L'Immoraliste et La Porte étroite</i>	277
10.4 <i>Clair et Ticket</i>	278
10.5 <i>Gros-Câlin et La Vie devant soi</i>	279

LISTE DES TABLES

	Page
CHAPITRE 7	
7.1 <i>L'Ulysses</i> de James Joyce	176
7.2 Constitution du corpus de Henmon	179
7.3 Constitution du corpus de Vander Beke	181
7.4 Constitution du corpus de Cheydleur	182
7.5 Liste des lemmes par fréquence décroissante	184
7.6 Textes et fréquences dans la banque de données du <i>Trésor de la langue française</i>	188
7.7 Les Dix premiers types dans le corpus d'Engwall	192
CHAPITRE 8	
8.1 La Longueur des phrases chez Hamilton et Madison	206
8.2 La Classification d'après Proudfoot et Horton	212
8.3 La Classification d'après Ledger et Merriam	213
8.4 Les Personnages de Jane Austen dont le dialogue comprend au moins 2.000 vocables	219
8.5 Les Trente Vocables les plus fréquents dans l'idiolecte des personnages de Jane Austen	220

8.6	Fréquence relative des vocables par tranche de mille vocables	224
-----	---	-----

CHAPITRE 9

9.1	Les Sept Étapes nécessaires pour l'analyse statistique du style	228
9.2	Calcul de la variance pour les trois auteurs	243
9.3	La Courbe gaussienne	245
9.4	Valeurs observées et table de contingence	253
9.5	Valeurs calculées	254
9.6	Valeurs du χ^2	254
9.7	Rapport entre le poids et la taille (a)	255
9.8	Corrélation entre le poids et la taille (a)	256
9.9	Rapport entre le poids et la taille (b)	257
9.10	Corrélation entre le poids et la taille (b)	257
9.11	Rapport entre le poids et la flexibilité	258
9.12	Corrélation entre le poids et la flexibilité	259
9.13	Rapport entre la taille et l'aptitude linguistique	260
9.14	Corrélation entre la taille et l'aptitude linguistique	260

CHAPITRE 10

10.1	Fréquences et Écarts réduits	269
10.2	Les Soixante Premiers Types dans chaque roman et dans le corpus d'Engwall	271
10.3	Répartition des 60 premiers types dans chaque roman et dans le corpus d'Engwall	273
10.4	Le Nombre d'observations se situant au-delà de deux écarts-types de la moyenne	280
10.5	Limites de confiance à 95% et à 99%	282

CHAPITRE 11

11.1	Valeurs du χ^2 pour les huit romans	286
11.2	Valeurs du χ^2 (sans les marqueurs de personne grammaticale)	288
11.3	Valeurs du χ^2 (<i>Clair</i> et <i>Ticket</i> représentent la valeur théorique)	289
11.4	Valeurs du χ^2 (sans marqueurs de personne): <i>Clair</i> et <i>Ticket</i> représentent la valeur théorique	290
11.5	Valeurs du χ^2 pour la table de contingence	292
11.6	Valeurs du χ^2 abstraction faite des vocables indiquant la perspective narrative	293

11.7	Valeurs du χ^2 pour les huit romans	294
11.8	Valeurs du χ^2 pour les huit romans (sans les marqueurs de personne)	295
11.9	Valeurs du χ^2 pour les quatre romans de Gary/Ajar	296
11.10	Valeurs du χ^2 pour les quatre romans de Gary/Ajar (abstraction faite des marqueurs de personne)	297

CHAPITRE 12

12.1	Coefficient de corrélation entre les huit romans et le corpus d'Engwall lorsque le nombre de types est de soixante	301
12.2	Coefficient de corrélation entre les huit romans et le corpus d'Engwall lorsque le nombre de types est de trente et un	302
12.3	Le r^2 entre <i>L'Immoraliste</i> et <i>La Porte étroite</i>	305
12.4	Le r^2 entre <i>Clair de femme</i> et quelques autres textes	306
12.5	Le r^2 entre <i>Au-delà de cette limite votre ticket n'est plus valable</i> et d'autres romans	307
12.6	Le r^2 entre <i>Gros-Câlin</i> et d'autres oeuvres	308
12.7	Le r^2 entre <i>La Vie devant soi</i> et quelques autres oeuvres	309

CHAPITRE 13

13.1	Engwall (500.000 vocables)	313
13.2	Les Neuf Textes: valeurs observées et valeurs calculées	315
13.3	Valeurs du χ^2 : les neuf textes	316
13.4	Valeurs du χ^2 : les huit romans	317
13.5	Valeurs du χ^2 : <i>Immoraliste</i> et <i>Porte</i>	319
13.6	Valeurs du χ^2 : <i>Clair</i> et <i>Ticket</i>	320
13.7	Valeurs du χ^2 : <i>Câlin</i> et <i>La Vie</i>	321

ABRÉVIATIONS

Toute citation qui figure dans ce travail est suivie (entre parenthèses) soit du nom de l'auteur et du numéro de la page, soit du titre de l'oeuvre et du numéro de la page. Les sigles suivants sont utilisés pour les oeuvres:

OEUVRES DE ROMAIN GARY

Cerfs-Volants: Romain Gary. Les Cerfs-Volants. Paris: Gallimard, 1980.

Chien: Romain Gary. Chien Blanc. Paris: Gallimard, 1970.

Clair: Romain Gary. Clair de femme. Paris: Gallimard, 1977.

Danse: Romain Gary. La Danse de Gengis Cohn. Paris: Gallimard, 1967.

Éducation: Romain Gary. Éducation européenne. Paris: Gallimard, 1945.

Gloire: Romain Gary. Gloire à nos illustres pionniers. Paris: Gallimard, 1962.

Nuit: Romain Gary. La Nuit sera calme. Paris: Gallimard, 1974.

Promesse: Romain Gary. La Promesse de l'aube. Paris: Gallimard, 1960.

Racines: Romain Gary. Les Racines du ciel. Paris: Gallimard, 1956.

Sganarelle: Romain Gary. Frère Océan I: Pour Sganarelle. Paris: Gallimard, 1965.

Stéphanie: Romain Gary. Les Têtes de Stéphanie. Paris: Gallimard, 1974.

Tête coupable: Romain Gary. La Tête coupable. Paris: Gallimard, 1968.

Ticket: Romain Gary. Au-delà de cette limite votre ticket n'est plus valable. Paris: Gallimard, 1975.

Vie et Mort: Romain Gary. Vie et Mort d'Emile Ajar. Paris: Gallimard, 1981.

OEUVRES PUBLIÉES SOUS LE PSEUDONYME D'ÉMILE AJAR

Câlin: Gros-Câlin. Paris: Mercure de France, 1974.

La Vie: La Vie devant soi. Paris: Mercure de France, 1975.

Pseudo: Pseudo. Paris: Mercure de France, 1976.

Salomon: L'Angoisse du Roi Salomon. Paris: Mercure de France, 1979.

LES AUTRES OEUVRES

Bruissement: Barthes, Roland. Le Bruissement de la langue. Paris: Seuil, 1993.

C.L.G.: Saussure (de) Ferdinand. Cours de linguistique générale. Publiée par Charles Bally et al. Paris: Payot, 1964-65.

Computation: Burrows, John. Computation into Criticism: A Study of Jane Austen's Novels and an Experiment in Method. Oxford: Clarendon Press, 1987.

Étranger: Camus, Albert. L'Étranger. Paris: Gallimard, 1942.

Fréquence: Engwall, Gunnell. Fréquence et distribution du vocabulaire dans un choix de romans français. Stockholm: Scriptor, 1974.

Immoraliste: Gide, André. L'Immoraliste. Paris: Mercure de France, 1902.

Inference: Mosteller, Frederick and David Wallace. Inference and Disputed

Authorship: The Federalist. Reading, Mass: Addison-Wesley, 1964.

Initiation: Muller, Charles. Initiation aux méthodes de la statistique linguistique.

Paris: Hachette, 1973.

Journal: Gide, André. Journal 1939-1949 Souvenirs. Paris, Gallimard, 1954.

Noeud: Mauriac, François. Le Noeud de vipères. Paris: B. Grasset, 1932.

Porte: Gide, André. La Porte étroite. Paris: Mercure de France, 1909.

Les abréviations suivantes sont utilisées dans les tables:

Clair: Clair de femme

Tick: Au-delà de cette limite votre ticket n'est plus valable

Câl: Gros-Câlin

Vie: La Vie devant soi

Étran: L'Étranger

Vip: Le Noeud de vipères

Imm: L'Immoraliste

Porte: La Porte étroite

Eng: Le corpus d'Engwall

GLOSSAIRE

Classement analytiques des symboles

x	une variable
y	une autre variable
χ_1	une valeur de la variable χ
i	l'intervalle de classe
μ	la moyenne
v	la variance
σ	l'écart-type
H_0	l'hypothèse nulle de la différence entre les valeurs observées et les valeurs de référence. C'est celle que l'on veut rejeter.
H_a	l'hypothèse contraire. C'est l'hypothèse que l'on veut appuyer.
α	le seuil de probabilité, ou le premier risque de rejeter H_0 lorsqu'elle est vraie.
ddl	le nombre de degrés de liberté
t	le test de Student-Fisher
Z	La valeur résultant de la transformation de r (transformation de Fischer)

χ^2 Le symbole du Khi-deux

Σ La somme de tout ce qui suit

Chapitre 1

L'illusion d'être soi

L'habitude de n'être que soi-même finit par nous priver totalement du reste du monde, de tous les autres; "je", c'est la fin des possibilités. (Gary, *Les Trésors de la mer Rouge* 78)

Que les lecteurs des romans soient *institués, coopératifs, invoqués* ou *apostrophés* (Maingueneau 30-6), ils réconcilient l'oeuvre avec son auteur, conscients du fait qu'un être en chair et en os a créé les personnages qui figurent dans les romans. Ils peuvent ainsi rapprocher l'ouvrage de la personnalité historique de l'auteur. Mais lorsqu'une oeuvre littéraire est publiée sous un pseudonyme, elle tend à dérouter les critiques et les lecteurs étant donné qu'ils ne peuvent rattacher l'oeuvre en question à une image ou à un nom.

Une telle attitude a existé tout au long de la vie fictive d'Émile Ajar. La coqueluche des médias durant l'existence d'Émile Ajar, Romain Gary continue toujours de fasciner les critiques et le public, et cela plusieurs années après sa mort. En fait, Gary est le seul romancier qui a réussi à transgresser un interdit de l'Académie Goncourt en recevant deux fois son prix, le prix littéraire le plus prestigieux de France, quoique cet auteur ait nié jusqu'au bout être Ajar.

Six mois après le suicide de Romain Gary, la publication de *L'Homme que l'on croyait* de Paul Pavlowitch et celle du testament littéraire posthume de Gary *Vie et mort d'Émile Ajar* (une sorte de confession épistolaire) désillent enfin les yeux des critiques et établissent la parenté littéraire de l'oeuvre signée Ajar. Avec ce testament, Romain Gary rejoint une fois pour toutes une légion de mystificateurs.

Dans ce petit livre de quarante-trois pages, Romain Gary s'en prend au "parisianisme" des critiques; il condamne leur "terreur des lettres", leurs "cliques à claques" et leur "copinage" (25). Il explique aussi les raisons profondes qui l'avaient poussé à entreprendre sa mystification: "J'étais las de n'être que moi-même. J'étais las de l'image Romain Gary qu'on m'avait collé sur le dos" (28). Cette confession suscite un grand remous dans le monde littéraire, d'autant plus que Gary se réjouit qu'avec son pseudonyme Ajar, il a réussi si bien à tromper et à déjouer les critiques qu'il accuse surtout de mal lire. Ceux-ci savent à peine réconcilier l'aventurier Ajar avec la carrière

respectable de Romain Gary, gaulliste convaincu, auteur engagé et défenseur du droit, de la dignité, de la fraternité et de la liberté de l'être.

Quelques-uns d'entre eux acceptent difficilement la mystification et ne lui pardonnent pas cet affront. En 1987 (sept ans après la mort de Gary), Angelo Rinaldo affirme dans une interview accordée à *Paris-Match*: "Je crois avoir été l'un des premiers à avoir senti qu'Ajar était un [pseudonyme]. Il y a des livres qui ont le goût du chef-d'oeuvre, mais ce ne sont pas des chefs-d'oeuvre" (4). Par ces propos Rinaldo semble suggérer que la biographie de l'auteur est indispensable à la compréhension du texte. Par contre, Jean Royer loue l'oeuvre de Gary, et déclare qu'en adoptant un pseudonyme, Romain Gary "a brisé les chaînes par lesquelles voulait le menotter une société franchement médiocre" (23). Ces deux propos démontrent le genre de divergence d'opinion qui entoure la valeur littéraire de l'oeuvre de Gary et d'Ajar.

Les écrits critiques sur Gary ne sont pas nombreux. Quelques articles paraissent périodiquement dans les quotidiens et les revues littéraires, mais jusqu'à récemment une seule thèse de doctorat de Guy Gallagher de l'Université Laval publiée en 1978 lui a été consacrée, aussi bien qu'un mémoire de maîtrise de Jane Rebecca McKee du Trinity College Dublin en 1978. Depuis quelque temps, les universitaires redécouvrent l'oeuvre de Romain Gary.

Deux thèses de doctorat consacrées à son oeuvre ont paru durant ces dernières années, l'une de Jørn Boisen de l'Université d'Odense, publiée en 1993, et l'autre de

Anne-Charlotte Östman de l'Université de Stockholm publiée en 1994. Viennent aussi de paraître, deux mémoires de maîtrise présentés à l'Université McGill, l'un de Hellène Lafond et l'autre de Dominique Fortier. Un livre d'Adam Pawlowski *Séries Temporelles en linguistique avec application à l'attribution de textes: Romain Gary et Émile Ajar* vient aussi d'être publié. Par contre, seulement quatre livres ont été consacrés à la vie de Romain Gary (Bayard, Bona, Catonné, Pavlowitch).

La plupart de ces auteurs expliquent la mystification Ajar aussi bien que le suicide de Romain Gary. Jean-Marie Catonné maintient que Romain Gary n'a jamais assumé la paternité de l'oeuvre Ajar puisqu'il ne voulait pas qu'Ajar apparaisse comme son véritable visage: "Gary n'était pas Ajar. Il l'a vraisemblablement moins inventé que subi" (95). Par ailleurs, le testament littéraire de Romain Gary, tout en expliquant la démarche Ajar, n'apporte pas non plus de réponse absolument convaincante sur cette mystification.

Ce travail explore, lui aussi, la mystification Émile Ajar, et cela par le biais d'une analyse de quatre romans de Gary, dont deux sont signés Ajar: *Au-delà de cette limite votre ticket n'est plus valable* (Gary, 1975), *Clair de femme* (Gary, 1977), *Gros-Câlin* (Ajar, 1974) et *La Vie devant soi* (Ajar, 1975). Puisque Romain Gary veut recommencer à zéro lorsqu'il assume le nom de plume d'Ajar, ce travail examine dans quelle mesure Gary a pu adopter un style différent dans l'oeuvre signée Ajar, à quel point il a pu s'affranchir de son écriture Gary lorsqu'il s'est camouflé derrière le nom fictif et inconnu d'Ajar.

1.1. Méthodologie

En ce qui concerne l'approche littéraire adoptée dans ce travail, elle est avant tout stylistique: elle est fondée sur le précepte de Ferdinand de Saussure que l'intérêt d'un texte réside dans les mots contenus dans le texte lui-même, que la langue est forme avant tout, et non substance (134). D'après cette optique, l'auteur est un artisan qui travaille à son art, tout comme un sculpteur cisèle la pierre ou le bois; sa compétence linguistique lui permet d'employer les mots qui sont puisés de son stock de vocabulaire, et son originalité réside dans la façon dont il retravaille la langue. Quoique ce travail considère le contenu des romans, il traite avant tout du style de ces romans, en s'attardant sur le vocabulaire.

Une telle approche est loin d'être traditionnelle. Au 19^e siècle, par exemple, Sainte-Beuve préconise l'étude de l'auteur dans son développement moral, intellectuel et émotionnel pour expliquer l'oeuvre littéraire (112). Mais Proust reconnaît qu'un livre est le produit "d'un autre *moi* que celui que nous manifestons dans nos habitudes, dans la société, dans nos vies. Ce moi-là, c'est [...] en essayant de le recréer en nous, que nous pouvons y parvenir" (221-2). Cet autre *moi* de l'auteur dont parle Proust se reconstruit à partir de la lecture de l'oeuvre. Ce travail se concentre lui aussi sur les caractéristiques linguistiques de l'oeuvre, selon la perspective de Proust.

Dans le domaine de l'analyse littéraire cette sorte d'approche s'écarte de la norme. Dans *La Critique littéraire au XX^e siècle*, par exemple, Jean-Yves Tadié détaille

les méthodes employées dans le domaine de l'analyse littéraire au 20^e siècle, telle que la critique psychanalytique, la critique sociologique etc. La plupart de ces approches privilégient le contenu du texte aux dépens de sa forme, et recherchent la réalité de l'oeuvre derrière la langue et non dans la langue: elles demeurent souvent muettes sur la forme et le style de l'oeuvre.

Le travail se range aussi dans la catégorie de l'esthétique-science proposée par Souriau (27). Il s'appuie, autant que possible, sur l'observation des faits linguistiques, et examine les romans par le biais de la statistique. Ce genre d'analyse, qui depuis longtemps déjà rend de nombreux services aux littéraires pour résoudre les problèmes d'attribution d'auteur et de datation de textes, présente des chiffres et des tables qui permettent de démontrer empiriquement les coutumes stylistiques de l'auteur.

1.2. Étude de style

De nos jours, il existe deux visions courantes du *style*. L'une postule que le style provient de la personnalité de l'auteur, que c'est le lien entre le sujet écrivant et son écriture. Pour Buffon "le style, c'est l'homme même" (Galdi 7). Il constitue une manière singulière d'écrire propre à un auteur, ce qui suggère l'existence d'un caractère génétique dans la forme de l'écrit. Si on envisage le langage des romans français du 20^e siècle au narrateur homodiégétique comme un fait de style pris au sens général, le style d'un auteur tel que Gide constitue alors un écart à cette norme; il a une valeur esthétique

individuelle qui permet de distinguer le style de Gide de la norme ou des autres auteurs. D'après cette tendance, il est toujours possible d'identifier un auteur, même lorsqu'il écrit sous un pseudonyme.

La deuxième vision est celle proposée par Léo Spitzer, l'un des fondateurs de la stylistique moderne. Spitzer affirme que le style d'un auteur constitue une façon particulière d'utiliser le langage; autrement dit, c'est une application spécifique de la langue à un certain moment de la carrière littéraire de l'auteur (89). Le style a une valeur esthétique individuelle qui s'écarte de l'usage normal, et il est possible de le mesurer. Cet écart reflète l'esprit de l'auteur à ce moment précis de sa carrière. Si cet aspect domine, l'auteur peut changer fondamentalement les caractéristiques de son style quand il le veut.

La tâche principale de ce travail est d'explorer ces deux concepts en les appliquant aux oeuvres de Gary/Ajar déjà mentionnées, afin de déterminer lequel domine. Cette recherche qui est statistique avant tout, identifie les unités de style à étudier, observe leur répartition, leur fréquence et la proportion de ces fréquences dans le texte et explique ensuite en quoi elles sont significatives.

Pour permettre une telle analyse, l'ordinateur est un outil indispensable. Il facilite l'analyse stylistique et statistique des oeuvres dans leur intégralité et permet de dénombrer avec rapidité les mots du texte. La machine fournit des données, dévoile des caractéristiques que l'être humain serait incapable de décerner et permet au

chercheur de faire une analyse objective. Mais l'ordinateur garde toujours sa juste valeur; il contribue sûrement à l'analyse statistique du style, mais ne la remplace pas.

1.3. Organisation du travail

Les chapitres de ce travail se disposent *grosso modo* en quatre parties. La première partie (chapitres 2 à 4) est consacrée en grande partie à la biographie de Romain Gary et aux principaux événements qui ont précédé le subterfuge Ajar. Le *moi* de Gary est abordé dans les chapitres 2 et 4 dans la mesure où il peut éclairer l'énigme Ajar et cerner la personnalité de Romain Gary au moment de la création de son double, Émile Ajar. Ce *moi* n'est ni basé ni calqué sur l'oeuvre romanesque de Romain Gary.

La vie de Romain Gary a pu être aisément vérifiée par les interviews qu'il a données aux journalistes et aux médias, par son testament littéraire *Vie et mort d'Émile Ajar*, par *La Nuit sera calme* (une espèce de confession de Gary). Les deux études biographiques consacrées à Romain Gary et écrites par ses proches ont aussi été consultées. C'est cela qui a pu justifier l'approche biographique dans ces deux chapitres. Le troisième chapitre est consacré à une brève description de l'intrigue des quatre romans de Gary.

La deuxième partie du travail (chapitres 5 et 6) aborde les similarités entre l'oeuvre de Gary et celle d'Ajar. Le chapitre 5 montre les ressemblances thématiques entre les romans de Gary/Ajar et le chapitre 6 examine les différents types de

mystification entrepris jusqu'ici dans le monde littéraire, et explique en quoi la mystification Ajar est différente des autres. La troisième partie (chapitres 7 et 8) fait un bilan du genre de recherche entrepris jusqu'ici dans le domaine de l'analyse statistique du style. La quatrième partie de l'étude comprend les chapitres 9 à 13. Le chapitre 9 est consacré à quelques définitions sur la statistique et les chapitres 10 à 13 sont surtout consacrés à l'étude statistique des romans déjà mentionnés. Enfin, le dernier chapitre élabore les conclusions qui s'imposent.

Chapitre 2

Il était une fois Romain Gary

Le romancier utilise la collaboration involontaire du lecteur antérieurement à la lecture [...] pour mieux le tromper [...] pour endormir son soupçon et sa vigilance, pour commettre son abus de confiance avec toute l'apparence, avec toute la force convaincante de la bonne foi. (*Sganarelle* 421-2)

Tout semble avoir réussi dans la vie de Romain Gary, homme à carrières et à faces multiples. Romancier célèbre et diplomate de carrière, c'est aussi un héros de la France libre et de l'aviation durant la Deuxième Guerre mondiale (recevant blessures, citations, croix de guerre), un Compagnon de l'Ordre de la Libération, un Chevalier de la Légion d'honneur, un cinéaste et un collaborateur de scénario à Hollywood.

Polyglotte, Romain Gary connaît la gloire littéraire dans une langue d'adoption, tout comme Vladimir Nabokov et Joseph Conrad. Mais il s'insère surtout dans la tradition française d'auteurs hommes d'action tels que Saint-Exupéry, Apollinaire et Malraux son maître à penser, "l'auteur-acteur de sa propre tragédie [...] le plus poignant mime de l'affaire homme que ce siècle ait connu (*Promesse* 190). Il est sans dons scolaires particuliers, étant "plus près des cancre que des fortes têtes" (Bona 27) et comme Jean Mermoz, l'aviateur français mort en action, il semble être plutôt fait pour la lutte fraternelle que pour la littérature. Il est sans doute ni athée, ni croyant mais plutôt agnostique, car pour lui "croire en Dieu, c'est calomnier Dieu, c'est un blasphème" (*Nuit* 57).

Gary fait partie d'un grand nombre d'écrivains contestataires et humanistes de la période de l'après-guerre qui, révoltés contre les attentats commis envers l'humanité, recherchent de nouvelles valeurs pour la reconforter. Pour ces écrivains, comme pour lui, le romancier doit s'engager dans ses écrits et dénoncer les injustices: "J'ai devant moi une oeuvre qui proteste, manifeste, pétitionne, appelle, crie et hurle, et qui est la seule contribution valable que je puisse faire (*Nuit* 98).

À l'apogée de sa carrière, Romain Gary fait partie du *jet-set* international et porte "sa fine moustache à la Clark Gable" (Catonné 19), ce qui semble contredire son aspect grave et sérieux. Il incarne tous les signes de la réussite et connaît une vie si pleine qu'il est difficile d'en dresser un bilan. Toutefois, ce chapitre se propose

d'élaborer les étapes les plus marquantes de la vie personnelle et professionnelle de Romain Gary à commencer par sa naissance à Moscou jusqu'à son suicide à Paris en 1980. Il s'attardera souvent sur les points les plus saillants de sa vie.

2.1. *La Promesse de l'aube*

Malgré tout le succès que connaît Romain Gary, il n'en reste pas moins qu'à sa naissance, tous les signes extérieurs présagent un avenir incertain, car il vient au monde à une époque des plus turbulentes de l'histoire de l'Europe. Il naît à Moscou le 8 mai 1914, à peine quelques années avant la révolution bolchevique et reçoit le nom civil de Roman Kacew. Dans *La Promesse de l'aube*, qui est une oeuvre autobiographique étoffée de plusieurs demi-vérités, ayant une structure qui "l'apparente un peu à la *Recherche* de Proust" (Catonné 29), Gary raconte l'histoire de cette enfance en Russie, aussi bien que l'histoire de sa mère, Nina Owczinski.

Fille d'un horloger, Nina Owczinski est née en 1883, à quelques cinq cents kilomètres au sud de Moscou. Réclamant son indépendance à un jeune âge, elle rompt avec la tradition familiale, se consacre à seize ans au théâtre, une "carrière rebelle" (Bona 27) et se joint à des troupes itinérantes, allant même jusqu'à Moscou. Comme c'est l'époque où l'on compare souvent la profession des comédiennes à celle des courtisanes, elle prend le nom de scène de Nina Borisovskaïa.

Par la suite, elle épouse Lebja Kacew et à trente et un ans, donne naissance à son enfant unique, Roman Kacew. Gary prétendra plus tard qu'elle avait trente-six ans à sa naissance (*Nuit* 61). Lebja Kacew n'est pas le père de cet enfant et Nina le quitte quelque temps après la naissance de l'enfant. Ainsi, comme Apollinaire, Romain Gary est né de père inconnu. C'est peut-être l'une des raisons pour laquelle il est constamment à la recherche d'une identité qui l'authentifie.

Romain Gary n'a vraisemblablement jamais rencontré Lebja Kacew. Ceci ne l'empêche pas pour autant de prétendre autrement. Il raconte dans *La Promesse de l'aube* que ce dernier était:

d'un aspect doux (et) avait de grands yeux bons et des mains toujours très soignées, et [...] lorsqu'il me regardait ainsi, tristement, avec, me semblait-il, un peu de reproche, je baissais toujours le regard, et j'avais, je ne sais pas pourquoi, l'impression de lui avoir joué un mauvais tour.

(*Promesse* 104)

Dans ce même roman, Gary raconte que Lebja Kacew a fini ses jours dans les chambres à gaz d'Auschwitz.

2.2. Les Années formatrices à Wilno et à Varsovie

En 1917 (peu après la chute du Tsar Nicolas II), Nina Kacew et son fils Roman quittent Moscou, "sans doute dans un wagon à bestiaux" (Catonné 210) et

commencent leur long cheminement vers l'Ouest. Ils se réfugient d'abord à Wilno, la capitale de la Lituanie, occupée alors par les troupes allemandes du général van Eichhorn. Cette ville pauvre et délabrée abrite un grand nombre de Juifs. À peine quelques années auparavant, la Russie et la Pologne se disputaient les frontières de Pologne, et Wilno appartenait aux Russes.

En 1918, Joseph Pilsudski établit un gouvernement indépendant en Pologne, mais les Russes et les Polonais continuent de se disputer les frontières. En 1919 Wilno passe à la Lituanie. Dans un tel climat d'incertitude et de bouleversement, les plus forts règnent en maîtres absolus à Wilno; les riches contrôlent tous les biens tandis que les pauvres vivent dans la misère, et que la famine bat son plein dans la ville de Wilno. La Lituanie passe aussi par une époque antisémitisme extrême et les Juifs sont souvent la cible d'un racisme féroce.

Dans *La Promesse de l'aube*, Romain Gary raconte les aventures de son enfance à Wilno où sa mère, qu'on pourrait qualifier de "mère poule", le protège de la misère et du climat d'intolérance qui y règne. Nina Kacew travaille comme une couturière et une confectionneuse de chapeaux pour subvenir à leur besoins matériels. Elle s'installe avec son fils au n° 16 rue Wielka-Pohulanka, la Grande Rue dans un quartier chic de Wilno que la plupart des autres Juifs ne connaissent pas (Catonné 210), car elle aime fréquenter la société bourgeoise.

Nina Kacew n'envoie pas son fils à l'école, mais elle lui apprend à lire et à écrire en russe. Elle lui fait aussi apprendre le français, le langage "des aristocrates plutôt que le yiddish que parlent les pauvres ouvriers" (Bona 36). À sept ans, le jeune Kacew sait lire les poèmes de Victor Hugo, réciter par coeur des fables de La Fontaine et chanter *La Marseillaise*. Sa mère lui offre aussi des leçons de danse et lui apprend l'art du baisemain russe, la marque d'un homme cultivé et distingué.

Par contre, Nina Kacew et son fils doivent, comme tout réfugié politique ou économique, lutter à chaque moment pour survivre dans le climat hostile de Wilno. Romain Gary raconte comment les dames riches de cette ville ont toutes voulu se faire habiller par Nina Kacew, qui avait ouvert une maison de haute couture. Mais une fois que Nina Kacew connaît la ruine financière, celles-ci font semblant de ne pas la voir en la croisant dans les rues. Cet incident a dû beaucoup blesser l'amour-propre du jeune Roman Kacew car des années plus tard, il écrit avec rancœur et sarcasme:

Je ne me souviens plus des noms de ces nobles créatures, mais, j'espère fermement qu'elles sont toujours en vie, qu'elles n'ont pas eu le temps de mettre leur viande à l'abri et que le régime communiste est venu leur apprendre un peu d'humanité. Je ne suis pas rancunier et je ne vais pas plus loin. (*Promesse* 126)

C'est pendant son séjour à Wilno que le jeune Roman Kacew ne sachant pas qui est son vrai père, s'invente un père putatif, Ivan Mosjoukine. Ce dernier est né en

1889 à Penza dans la Volga centrale qui est l'une des régions d'origine des Tartares. Grec orthodoxe, il est issu d'une famille de propriétaires terriens d'ancienne noblesse. C'était aussi une grande vedette et une idole populaire du film muet en Russie; il avait connu et fréquenté Nina Kacew à Moscou.

Dans *La Promesse de l'aube*, Romain Gary raconte que durant l'exil de Mosjoukine en France, ce dernier continue à fréquenter Nina Kacew jusqu'à janvier 1939, date à laquelle il meurt de tuberculose dans une maison de santé à Neuilly. Il est légitime de postuler que pour le jeune Kacew, s'imaginer être le fils d'un Russe orthodoxe noble vaut mieux que de s'imaginer le fils de Lebja Kacew, un Juif anonyme.

Quelle que soit la vérité sur cette paternité supposée, toujours est-il que la ressemblance entre Romain Gary et Ivan Mosjoukine est extraordinaire. Dominique Bona affirme:

ils ont le même visage oriental, les mêmes yeux clairs, la même peau bistre, les mêmes sourcils, les mêmes pommettes saillantes et le front large, arrondi [...]. Ils évoquent l'un et l'autre par leur stature et leur faciès brûlé, ces guerriers tartares des armées de Genghis Khan ou de Tamerlan. (35)

Dans plusieurs oeuvres Gary déclare avec désinvolture qu'il existe en lui un curieux mélange de sang sémite-russo-polonais. Il se proclame "bâtard asiatique"

(*Sganarelle* 227) étant “un peu Cosaque et Tartare, mâtiné de Juif - une sorte de Genghis Cohn” (*Sganarelle* 33). Dans *Chien Blanc* il évoque son sang mongol, tartare et juif et fait mention de ses ancêtres en ces termes: “du côté paternel ce sont des anciens mongols [...] par dessus le marché, mes ancêtres tartares paternels étaient des progromeurs et mes ancêtres juifs maternels des progromés” (163). D’après Catonné, Gary reconnaît ainsi en lui une nature divisée de juif-slave et revendique une bâtardise raciale de mi-juif, mi-slave et un alliage des maraudeurs des steppes et de la souffrance des ghettos juifs d’Europe (44).

Entre 1921 et 1923 Nina Kacew et son fils quittent Wilno “sans regret” (*Promesse* 126) pour échapper aux créanciers mais aussi sans doute pour fuir l’antisémitisme. Ils s’installent à Varsovie et rejoignent la famille du frère de Nina Kacew. À Varsovie, Nina Kacew gagne difficilement sa vie; elle mène une vie d’expédients et lutte jour et nuit pour nourrir son fils. Elle exerce tour à tour le métier de courtière de bijoux, de gérante d’immeubles. Souvent à court d’argent, Nina Kacew et son fils sont chassés par les propriétaires des immeubles pour non-paiement de loyer et se trouvent contraints de déménager plusieurs fois. Elle ne peut exaucer l’un de ses plus grands désirs, inscrire son fils au lycée français car cela dépasse ses moyens financiers.

Nina Kacew est obligée d’envoyer son fils au *Gymnasium Kreczmar*, le lycée réservé aux prolétaires. Au *Gymnasium*, Roman Kacew livré à lui-même, loin de sa

mère, apprend à se défendre contre l'intolérance et l'antisémitisme qui y règnent. C'est aussi à cette époque que le jeune Kacew rêve souvent de se surpasser, sous la poussée d'une mère ambitieuse. C'est elle qui lui inculque la notion de l'ambition et lui fournit la définition du succès:

Les poètes maudits, Baudelaire en particulier et les musiciens tragiques étaient soigneusement bannis de la collection [...]. Le succès était quelque chose qui devait nous arriver de notre vivant. (*Promesse* 108)

2.3. Adolescence: La France

D'après Romain Gary, sa mère avait toujours aimé la France et rêvait souvent de s'y établir. Pour elle, la France représente "la terre de toutes les libertés et les douceurs de vie" (Bona 36). La raison de cet amour démesuré est peut-être celle que fournit Romain Gary lui-même: "ma mère aimait la France sans raison aucune, comme chaque fois que l'on aime vraiment" (*Promesse* 215). C'est sans doute pour cela, mais aussi à cause de l'intolérance qui règne à Varsovie que lorsque Roman Kacew a treize ans, sa mère et lui quittent définitivement la Pologne pour aller vivre en France.

Nina Kacew et son fils s'installent à Nice et habitent d'abord un petit appartement de deux pièces dans la rue Shakespeare. Pour subvenir à leurs besoins, Nina Kacew accepte les métiers les plus humbles: elle fait la toilette des chats et des

chiens et se fait engager comme femme de ménage. Comme elle se fait souvent passer pour une grande dame aristocrate russe ruinée, l'hôtel-Pension Mermonts qui accueille la grande clientèle de l'époque, l'engage comme gérante. L'hôtel est situé dans un quartier "qui a la nostalgie de Kiev et de Saint Pétersbourg" (Bona 13). Peu de temps après, Nina Kacew et son fils déménagent à l'hôtel-Pension Mermonts et la fortune semble leur sourire. Ils vivent en sécurité, à l'abri des révolutions et de la pauvreté.

Quoique atteinte du diabète, Nina Kacew gère toutes les affaires de l'hôtel avec aplomb et courage. Elle ne cesse, pour autant, de rappeler à son fils que la seule chose qui compte dans la vie, c'est de réussir. Elle rêve que son fils devienne le plus grand écrivain de la France. Roman Kacew, qui veut toujours accomplir des prouesses pour sa mère, tâche d'exaucer les vœux de celle-ci en s'exerçant au métier d'écrivain: "Les cahiers s'amorçaient sur ma table, couverts de pseudonymes de plus en plus éloquents, de plus en plus superbes, de plus en plus désespérés" (*Promesse* 154).

Nina Kacew et son fils peuvent enfin vivre à leur aise. Le fils poursuit ses études au lycée tandis que la mère travaille sans relâche afin d'offrir à son fils tout le confort qui lui a manqué avant son arrivée à Nice. Il n'oubliera jamais le courage de cette mère si dévouée, déclarant plusieurs années après la mort de celle-ci:

si tous mes livres sont pleins d'appels à la dignité, à la justice, si l'on parle tellement et si haut de l'honneur d'être un homme, c'est peut-être parce que j'ai vécu [...] du travail d'une vieille femme malade et surmenée. (*Promesse* 185)

Après ses études au lycée, Roman Kacew s'inscrit à la Faculté de Droit à Aix-en-Provence et quitte sa mère en octobre 1934. C'est un moment douloureux et déchirant tant pour la mère que pour le fils. Il s'efforce alors à se prouver digne d'elle pendant qu'elle vit encore.

Quelque temps plus tard, il s'inscrit à la Faculté de Droit à Paris et s'installe à Paris. Il est forcé d'exercer de petits métiers étant donné que la mensualité que lui envoie sa mère suffit à peine à ses besoins. D'après Jean-Marie Catonné (212), entre 1933 et 1934, il passe un diplôme d'études slaves à l'université de Varsovie dans des conditions mystérieuses. Le 14 juillet 1935 à l'âge de vingt et un ans, il est enfin naturalisé français. Il éprouve dorénavant le sentiment de l'être complètement: "Ma formation culturelle est française, et je me sens Français sans complexes" (*Nuit* 54).

2.4. Carrière militaire

En 1938, Roman Kacew voulant se consacrer à une carrière dans l'armée, fait sa Préparation Militaire Supérieure, étape nécessaire à franchir pour ceux qui veulent entrer dans l'armée. Ceci lui permettra d'obtenir le galon de sous-lieutenant. Parmi

les trois cents élèves environ qui se présentent à l'examen final, il est quatrième. Malgré son succès, il n'est pas reçu officier et demeure le seul dans sa classe à être simple soldat. Ceci le fait rougir de honte et lui donne conscience d'une infériorité.

La raison officielle que l'administration invoque pour expliquer au jeune Roman Kacew cette sanction est que Romain Kacew n'est naturalisé français que récemment, et que seuls les Français de souche ou ceux qui sont naturalisés depuis dix ans sont admis comme officiers. Mais jusqu'à la fin de ses jours, Romain Gary est convaincu que cette sanction prononcée contre lui découle d'un sentiment antisémite: "Dans le mauvais coup qui venait de m'être fait, je n'avais aucune peine à reconnaître la main de Totoche, le dieu de la bêtise, celui qui devait bientôt faire d'Hitler le maître de l'Europe" (*Promesse* 218).

La Deuxième Guerre mondiale est déclarée en 1939, et comme les autres Français, Roman Kacew est mobilisé. Il est nommé instructeur de tir à l'École de l'Air de Salon-de-Provence, puis envoyé à Bordeaux-Mérignac au sud de la France, où il est nommé sergent. En 1940, la France subissant de lourdes défaites capitule, et le Général Pétain signe l'armistice. Roman Kacew, étant de souche juive et craignant sans doute un gouvernement collaborateur, quitte la France et s'installe à Londres. Il devient un Français libre, rejoint d'autres gaullistes et se soumet à un entraînement intense et accéléré de la Royal Air Force. C'est à ce moment qu'il change le nom de Roman Kacew à celui de Romain Gary.

Romain Gary participe à de nombreuses expéditions aériennes; il a plusieurs aventures dangereuses en Afrique et frôle souvent la mort. En septembre 1941, il contracte la fièvre typhoïde en Égypte et passe deux semaines au pavillon des contagieux à l'hôpital de Damas, entre la vie et la mort. Lorsqu'il devient comateux, l'hôpital lui administre cinq transfusions de sang et il reçoit l'extrême-onction. Il se rétablit, peu à peu, quoiqu'il demeure convalescent pendant six mois à l'hôpital de Damas.

Malade, il ne peut participer à la campagne de Libye menée contre les troupes du Maréchal Rommel et à laquelle participe tous ses compagnons du groupe Lorraine. Au mois d'août 1942, quoiqu'il ne soit pas complètement rétabli, il participe aux opérations de guerre et effectue une soixantaine de vols aériens. Sa mère Nina Kacew meurt cette même année en France, alors qu'il combat du côté des forces alliées.

La bravoure fait partie intégrante du psychisme de Romain Gary. Lors d'un vol d'opération à bombardement où il vise les centrales électriques et les gares, Romain Gary est atteint au ventre par une projectile. La boucle de son parachute lui sert de bouclier et le sauve d'une mort certaine. Le pilote de l'avion est touché aux yeux et devient aveugle. Guidé par la voix de Gary, le pilote peut quand même lâcher une bombe sur sa cible et fait atterrir son avion et tout son équipage à bord sain et sauf. En reconnaissance de ce geste de bravoure, le Général de Gaulle octroie la Croix de la Libération à Romain Gary, en janvier 1944.

Durant la guerre, Romain Gary rédige le roman *Éducation européenne*. Il le fait publier en Angleterre dans sa version anglaise *The Forest of Anger*, “translated from French by Viola Garvin”. Le roman est très bien reçu par le public anglais et la critique anglaise. Comme les talents d’écrivain de Gary commencent à être reconnus dans le monde littéraire, il espère de plus en plus que la guerre se termine bientôt, ce qui lui permettra de se consacrer entièrement à sa carrière d’écrivain.

Paris est libéré au mois d’août 1944 et le Général de Gaulle arrive au pouvoir au cours de cette même année. Quelque temps après, Romain Gary épouse Lesley Blanch, une correspondante du *Daily Telegraph*, et une russophile anglaise convaincue. C’est une amie intime du comédien anglais Peter Ustinov. Elle admire Romain Gary sans doute parce que, d’après elle, il “ressemble à Gogol” (Bona 99). L’Allemagne capitule au mois de mai 1945 et, au mois de juin 1945 l’*Éducation européenne* est publié en France.

2.5. La Forêt de la colère

Le monde littéraire français découvre Romain Gary en 1945 lors de la parution de ce premier roman. Dans un souffle épique rappelant les oeuvres d’André Malraux, le roman raconte les aventures d’un petit groupe de partisans polonais durant la Deuxième Guerre mondiale. Ils doivent se terrer dans des trous qu’ils ont creusés afin de ne pas se faire attraper par l’armée allemande qui occupe leur pays.

Si dans un premier temps le titre du roman paraît optimiste, il est néanmoins trompeur: il dément son thème principal aussi bien que son titre original *La Forêt de la colère*. Loin d'évoquer un pèlerinage culturel à travers l'Europe "civilisée", ce petit roman met au contraire en lumière toute la barbarie violente dont les pays européens sont capables.

L'Éducation européenne n'est pas un roman d'aventures, car rien de vraiment spectaculaire ne se produit, sauf l'effort constant des partisans de survivre au froid, à la neige et à l'hiver, les quelques attaques qu'ils montent contre des convois allemands et les représailles qu'ils infligent aux traîtres.

Les partisans savent que leur lutte est vaine et désespérée dans cet univers absurde. Le seul remède qu'ils ont contre la dépression la plus noire, le suicide et la folie, est leur foi en Nadejda, le héros "insaisissable, invincible, protégé par tout un peuple [qu'aucune] puissance au monde, aucune force matérielle, ne peuvent empêcher de continuer et de triompher" (*Éducation* 105).

Les partisans ont aussi l'espoir en l'armée Rouge qui se bat à Stalingrade, et, ils attendent le moment où ils pourront lutter aux côtés des Soviétiques contre les Allemands. Le héros du roman, Janek Twardowski, est un jeune adolescent de quatorze ans qui semble aussi être le porte-parole de Romain Gary. Tout comme Gary, Janek refuse de se livrer au fanatisme; il est conscient du ridicule de sa tâche et

du ridicule de la condition humaine. Plusieurs fois dans le roman, le jeune Janek compare le monde des humains au monde sans coeur des insectes:

Le monde où souffrent et meurent les hommes est le même que celui où souffrent et meurent les fourmis: un monde cruel et incompréhensible, où la seule chose qui compte est de porter toujours plus loin, une brindille absurde, un fétu de paille, toujours plus loin, à la sueur de son front et à ses larmes de sang, toujours plus loin!

(*Éducation* 282)

Dominique Bona affirme que *l'Éducation européenne* met déjà en évidence chez Romain Gary le sens de l'absurde et révèle son pessimisme face à la vie qu'il appelle "la grande pagaille" (137). Jean-Marie Catonné y trouve des reflets de Gogol dans "cette humanité réduite à des conditions de vie larvaires" (134). Quant à Jean-Paul Sartre, il se demande si dans une centaine d'années les lecteurs pourront juger que "*l'Éducation européenne* était oui ou non *le* livre de la Résistance" (*Situations II*, 40).

Peu de temps après sa publication, ce petit roman de Romain Gary obtient le Prix Critique du livre. Avec son premier roman, Romain Gary commence à se faire une réputation littéraire respectable qui se concrétise et se cristallise au cours des trente prochaines années avec la publication de plusieurs autres romans. Le 14 juillet

1945 le Général de Gaulle décerne la Croix de la Légion d'honneur à Romain Gary; Gary est décoré sous l'Arc de Triomphe.

2.6. Débuts diplomatiques

Le 25 octobre 1945, Romain Gary commence sa carrière de fonctionnaire au Quai d'Orsay, et Georges Bidault, alors ministre des affaires étrangères, lui propose une carrière diplomatique. Quelque temps après, Romain Gary se brouille avec le chef de cabinet d'André Malraux, Raymond Aron, car ce dernier ne lui a pas accordé le poste de conseiller culturel à Londres que lui avait promis Malraux. C'est sans doute pour cela qu'en décembre 1945, Romain Gary, accompagné de sa femme, quitte Paris pour prendre ses fonctions à Sofia.

Il règne à cette époque en Bulgarie, une atmosphère tendue et les intrigues entre communistes et anticommunistes, qui forment un gouvernement de coalition, augmentent chaque jour. Comme Gary se lie d'amitié avec plusieurs intellectuels bulgares, les communistes éprouvent envers lui une grande méfiance. Or, tout au long de son existence, Romain Gary n'a été ni anticommuniste, ni communiste et, comme l'a si bien noté Dominique Bona: "alors qu'il [a] fui la Russie à cause de la Révolution bolchevique, et qu'il la [retrouve] à l'oeuvre à Sofia, vingt ans plus tard, Gary ne [fait] jamais du communisme ni de l'anticommunisme un thème de son oeuvre (119).

Par ailleurs, Gary refuse les catégorisations raciales, civiles, nationales ou politiques. Il se proclame "citoyen du monde" (Bona 39). Il est d'une telle indifférence politique qu'il est attaqué par les extrémistes de tout bord. En France il est dénoncé comme Juif par la presse antisémite et en Israël il est dénoncé comme non-Juif. Si Gary affirme ses origines juives par sa mère, ce n'est pas par chauvinisme, mais plutôt pour apprendre à Israël qu'il n'a pas le droit de décréter qui est juif et qui ne l'est pas: "Demi-juif, je ne sais pas ce que ça veut dire. Demi juif, c'est demi-parapluie. C'est une notion à l'usage des racistes maniaques d'Israël" (*Nuit* 235).

Pour rétablir son équilibre mental, Romain Gary s'éloigne des intrigues qui sévissent en Bulgarie et séjourne brièvement à Paris. Peu de temps après, il retourne à Sofia où il assiste à l'abolition de la monarchie bulgare et à la proclamation de la république, un geste qui entraîne des exécutions sommaires. C'est sans doute pour cette raison que Romain Gary commence à rêver d'une Europe unie, sans mur ni rideau de fer qui engloberait à la fois, les pays de l'Est et de l'Ouest. Comme il recherche une Europe libre, affranchie du nationalisme, il revendique son ascendance slave ou même Tzigane. Comme lui, ces Tziganes venaient sans doute de l'Est et étaient les premiers voyageurs "sans passeports ni frontières" (Bona 131).

En janvier 1946, le Général de Gaulle démissionne. En 1948 Romain Gary quitte définitivement la Bulgarie et retourne à Paris où il est nommé à l'administration

centrale sous les ordres de Camille Paris, le gendre de Paul Claudel. Il se consacre alors à sa carrière politique autant qu'à sa carrière littéraire. C'est aussi à partir de cette époque que sa pensée philosophique commence à se concrétiser.

Gary commence à diviser le monde en deux, à savoir les amis (ceux qui ont cru en une France libre) et les "salauds", c'est-à-dire tous ceux qui d'une manière ou d'une autre ont collaboré avec les Allemands. Pour élargir ses horizons, il entreprend de nombreux voyages. Se trouvant en Suisse en mission, il rencontre Henri Hoppenot, un humaniste gauchiste.

De 1947 à 1952, Romain Gary publie trois autres romans chez Gallimard: *Tulipe* en 1947 (un échec total), *Le Grand Vestiaire* en 1949 (que personne ne remarque en France) et *Les Couleurs du jour* en 1952 (qui passe inaperçu). En décembre 1949, il est nommé premier secrétaire à Berne. En octobre 1951, Roman Kacew devient officiellement Romain Gary.

Lorsqu'il est envoyé à l'O.N.U. en février 1952 comme premier secrétaire et attaché de presse, portant le titre officiel de "Public Relations Officer" ou porte-parole, une nouvelle étape de sa carrière diplomatique commence. L'année 1952 marque l'apogée de la guerre froide. En France c'est aussi l'époque de la guerre d'Indochine. Romain Gary qui, dans ses oeuvres se proclame "du côté des faibles et des opprimés et [...] révolté par tant d'attentats contre la dignité humaine" (Outzen 345), doit néanmoins, jour après jour, défendre des positions qu'il ne partage pas,

telles que le colonialisme et la présence de l'armée française en Indochine. Il doit aussi expliquer le refus de la France d'accorder l'indépendance à la Tunisie, au Maroc et aux états d'Afrique noire.

L'un des thèmes essentiels des *Racines du Ciel* est un appel à la liberté africaine. En expliquant, plus tard, cette contradiction, Romain Gary dit qu'étant ambassadeur de France à l'O.N.U., il ne faisait qu'exercer un métier, le métier d'avocat "avec toute la virtuosité technique dont j'étais capable" (*Nuit* 185). Par contre aux États-Unis, les maccarthystes livrent une véritable chasse aux sorcières et se déclarent ouvertement contre ceux qui ont des liens avec les communistes ou les gauchistes.

Un jour à l'O.N.U., Krishna Menon, le délégué de l'Inde et le futur ministre de la Défense indienne, intellectuel libéral de gauche et socialiste avoué dans le sillage de Nehru, s'apprête à parler. Un journaliste du *Chicago Tribune* lui pose la question suivante: "Monsieur, est-ce que c'est vrai que vous êtes un communiste notoire?" (*Nuit* 211). Krishna Menon quitte la tribune dans une grande colère. Le journaliste pose la même question à Romain Gary. Gary marque un silence, et répond à ce journaliste, père de famille, en direct sur les ondes: "Monsieur, est-ce que c'est vrai ce que tout le monde raconte dans les couloirs, que vous êtes un homosexuel notoire?" (*Nuit* 211).

Cette anecdote démontre que Romain Gary a souvent les répliques à point et sait défendre une cause. Elle révèle aussi son "talent de retourner l'agression contre l'agresseur [...] de tourner les reproches contre son interlocuteur" (Bona 153). Ceci montre aussi sa méfiance de l'ultra conservatisme et de tous les excès commis au nom de la politique. Romain Gary expose ses expériences à l'O.N.U. dans un roman intitulé *L'Homme à la colombe*, qu'il publie sous son premier pseudonyme Fosco Sinibaldi.

Plus tard il explique à son ami François Bondy sa réaction au maccarthysme de la façon suivante:

J'avais donc usé du maccarthysme, [...] parfois tu n'as pas le choix des armes [...]. Il y a des moments où tu ne peux pas te permettre de te retirer sur les hauteurs et de dominer ton adversaire de toute ta grandeur morale, quand c'est au couteau, c'est au couteau. (*Nuit* 212)

Néanmoins, du point de vue politique, Romain Gary ne préconise ni le capitalisme ni le communisme: "Le capitalisme n'est pas menacé par la corruption: il se prolonge par la corruption [...]. Sans corruption, il n'y aurait pas eu de surplus" (*Nuit* 33).

Gary déclare, par ailleurs, que le communisme ne l'a jamais marqué: "il faut laisser au communisme le temps de rater son coup pour qu'il réussisse" (*Nuit* 123-4). Il aspire au socialisme au "visage humain, celui qui a accumulé tous les échecs mais n'a cessé de montrer la seule direction de marche qui [...] paraît digne d'être suivi"

(Nuit 99), mais pas au communisme de Staline “pervertie jusqu’à la moelle épinière” (Nuit 33).

Paradoxalement, Gary se plaît beaucoup aux États-Unis, le pays capitaliste par excellence. Au siège de l’O.N.U., Romain Gary vit à son aise, se déplaçant souvent en tenue décontractée, à l’américaine. Entre-temps sa femme Lesley Blanch, romancière, elle aussi, publie son premier roman *Les Rives sauvages de l’amour* qui connaît un grand succès. La compagnie Métro-Goldwyn-Mayer achète les droits de *La Sultane blonde*, autre grand succès de sa femme.

En 1955 lorsque Romain Gary est secrétaire d’ambassade à Londres, il brouille avec l’ambassadeur Jean Chavel. En février 1956, il est nommé consul général de France à Los Angeles.

2.7. *Les Racines du ciel*

En automne 1956, onze ans après la parution d’*Éducation européenne*, Romain Gary publie *Les Racines du ciel*, “le premier roman écologique de la littérature française, [une] défense chaleureuse des pays africains contre l’exploitation coloniale” (Nøjgaard 31). Le roman est probablement né des souvenirs du Tchad avec ses grandes plaines grises où Romain Gary a séjourné brièvement durant la Deuxième Guerre mondiale.

Les Racines du ciel raconte l'histoire de Morel, un Don Quichotte moderne et un idéaliste dur qui est incapable de s'adapter aux manigances et à l'hypocrisie du monde. Durant la guerre, Morel a été interné dans le camp de concentration à Auschwitz. Dans ses pires moments de souffrance et de désespoir, Morel imaginait le grand éléphant d'Afrique dans toute sa force et sa splendeur, traversant les plaines, écrasant tout sur son passage sans que rien ne puisse l'arrêter. Morel doit donc la vie aux éléphants, car l'éléphant imaginaire lui a redonné l'espoir de survivre et de vaincre toutes les atrocités. L'image de l'éléphant tout puissant est ainsi devenue le symbole de la liberté et de la force de l'âme humaine.

Ayant survécu à Auschwitz, Morel s'établit au Tchad afin de sauver les éléphants africains des chasseurs blancs et noirs, qui comme les pires Nazis torturent et mutilent ces éléphants. Au Tchad, les marginaux et les idéalistes se rallient à Morel. Parmi ce groupe d'idéalistes se trouve Forsythe, un *G.I.* américain proscrit par son pays. Sous la douleur de la torture il a avoué aux Coréens qu'il a bombardé la population coréenne de mouches infectées de peste et de choléra, et cela sous les ordres de l'armée américaine.

Minna, une ancienne prostituée allemande sans illusions se trouve aussi au Tchad: elle suit Morel dans la brousse et s'allie à lui parce que "son combat, c'est quelque chose de propre" (*Racines* 178). Un missionnaire franciscain s'y trouve aussi. Mais lui s'intéresse plutôt aux problèmes humains qui existent au Tchad. Il

veut mettre fin aux rites sexuels des tribus. À la fin, même ceux qui s'opposent le plus à Morel, tels que le gouverneur du Tchad, le prêtre jésuite Tassin et le missionnaire franciscain se convertissent à sa cause.

Le plus grand opposant à la cause de Morel est Waitari, un nationaliste noir, député français éduqué à la Sorbonne, et sympathisant des Maus-Maus. Waitari, un moderniste, veut débarrasser le Tchad des troupeaux d'éléphants sauvages qui errent à travers les plaines. Il les considère comme le plus grand archaïsme de son pays.

Les Racines du ciel contient un message écologique qui intéresse le public littéraire dès sa parution. Son intérêt réside sans doute dans le fait que chaque personnage présente un point de vue auquel le lecteur peut s'identifier. Le lecteur peut prendre le parti de Morel, de Minna, de Forsythe, du franciscain, du prêtre jésuite, qui veulent sauver les éléphants, ou inversement du nationaliste Waitari qui veut les détruire au nom de la modernité. En effet, comme l'indique Romain Gary: "À chacun ses éléphants, et à chacun ses Juifs" (*Nuit* 48).

Les Racines du ciel connaît un succès éclatant. Le 3 décembre 1956, lorsque Romain Gary est à La Paz (en Bolivie) où il exerce les fonctions de chargé d'affaires par intérim, le Prix Goncourt lui est décerné pour ce roman.

C'est aussi l'époque où la scène politique ne s'avère guère prometteuse sur le plan international. Les ennuis entre les Arabes et les Israéliens au Moyen-Orient augmentent, tandis qu'à Budapest la révolte hongroise est réprimée par l'armée

soviétique. Romain Gary ne cesse de réfléchir sur ce monde qui semble être en perdition. C'est peut-être pour cela qu'il fait la déclaration suivante à la presse bolivienne après l'annonce du Prix Goncourt:

Je suis déchiré, [...] entre la joie de me voir décerner le Prix Goncourt et la tristesse de constater que l'idéal de liberté et de dignité humaines que je défends dans mon livre n'a jamais été plus menacé. L'appel que lancent en ce moment tous les écrivains du monde dignes de ce nom pour que l'on respecte les droits de l'homme, ne recueille que la réponse des armes atomiques. Ce que j'ai défendu depuis 1936 sans interruption, les armes à la main, je continuerai à les défendre, aussi bien dans ma vie que dans mon oeuvre. (Bona 172)

Même avec l'obtention du prix Goncourt, *Les Racines du ciel* ne cesse de diviser la critique. André Billy du *Figaro* affirme que *Les Racines du ciel* est "un livre extraordinaire" (Bona 192). René Lalou, des *Nouvelles littéraires* remarque qu'il s'agit d'un roman "symphonique, d'une puissante intensité, et dont maints accents éveillent maints échos dans la conscience profonde de ses lecteurs" (Bona 182). Pour André Maurois du *Carrefour*, *Les Racines du ciel* est "un livre étrange, symbolique et vaste. L'auteur a un talent puissant, un esprit généreux. Un de nos espoirs" (Bona 182).

Par contre, d'autres critiques s'en prennent à la longueur du livre aussi bien qu'au style touffu de Romain Gary et au fait que la maison d'édition Gallimard ait pu imprimer un livre criblé de tant de fautes grammaticales. Stephen Hecquet du *Bulletin de Paris* affirme que "quand on a décidé de réécrire *La Condition humaine*, on ne triche pas avec les conjugaisons" (Bona 183). Il ajoute en outre que Romain Gary lui semble "pencher du côté de M. Pierre Boule plutôt que du côté de M. Camus ou de M. Malraux". Kléber Haedens y verse son propre venin et affirme:

Pour venir à bout des *Racines du ciel*, il faut surmonter la fatigue que donne la répétition implacable des mêmes idées et des mêmes thèmes (pourtant bien simples), répétition qui permet de supposer que Romain Gary n'a pas l'espoir d'être lu par des gens très intelligents. (Bona 183)

Carmen Tessier dans *Paris-Presse* remarque que *Les Racines du ciel* contient tellement d'erreurs de syntaxes et de conjugaisons que l'auteur est sans aucun doute analphabète. Elle accuse en même temps Albert Camus, qui, affirme-t-elle, s'est très mal acquitté de sa fonction d'éditeur chez Gallimard (Bona 181). Dans un communiqué de presse Albert Camus s'empresse de réfuter immédiatement cette accusation en la déclarant "totalement dénuée de fondement et également blessante pour tous ceux qu'elle met en cause" (Bona 181). Romain Gary répond lui aussi à cette attaque en la qualifiant de commérages.

Avec le recul, l'on constate qu'il n'y a rien de surprenant que Romain Gary ait obtenu le prix Goncourt pour son roman. Il est avant tout "un auteur moraliste, humaniste et idéaliste, défenseur des bonnes causes" (Outzen 345). Et, dans ses écrits Romain Gary se déclare altruiste et humanitaire.

Malgré cela, la vanité fait partie intégrante de sa personnalité. Lorsqu'il arrive à Paris pour fêter l'obtention de son Goncourt, il songe avant tout de soigner son image et de cultiver sa différence. Il sait que les critiques en France l'accusent de mal manier la langue française. Pour rassurer ses critiques, il revendique tranquillement une mère française, ce qui prouverait que le français est sa langue maternelle. Peu de temps après, il se prépare à prendre ses fonctions de Consul de France à Los Angeles. En 1957 il quitte la France pour Hollywood.

2.8. Expériences à Los Angeles

En 1957, Romain Gary habite une belle villa à Hollywood et fréquente les vedettes du cinéma. Souvent chaussé de bottes de cow-boy et vêtu d'un poncho, il donne des interviews devant la caméra dans le jargon californien. Il fait aussi la connaissance de plusieurs artistes américains et français de l'époque tels que Marilyn Monroe, Charles Boyer, Édith Piaf et Maurice Chevalier. Il rencontre le grand metteur en scène américain Cecil B. De Mille.

Son allure plaît tellement aux Américains que Walter Wanger lui propose d'interpréter le rôle de Jules César dans le film *Cleopatra*. Ce rôle sera en fin de compte incarné par l'acteur anglais Rex Harrison. Pour Gary, Hollywood est un cirque où s'agitent les millionnaires et où les seules valeurs qui comptent sont celles du *box-office*, du succès et du *glamour* (Bona 191). C'est cet aspect artificiel de la vie américaine que dénonce Romain Gary. Plus tard, en parlant de Hollywood, à son ami François Bondy il dit: "C'était une pyramide où chaque étage était scrupuleusement fixé en termes de succès, de fric, de réussite" (*Nuit 243*).

Même à Hollywood Romain Gary ne perd pas le talent de bien émouvoir son auditoire. Il lit un jour, lors de la fête des mères, à Los Angeles, des lettres émouvantes que sa propre mère, alors décédée, lui avait écrites afin de lui redonner courage lorsqu'il combattait loin d'elle pour la France. Ces lettres font fondre en larmes toute l'auditoire.

Les studios d'Hollywood qui s'intéressent de plus en plus aux romans de Romain Gary tournent le film *Les Racines du ciel*: le scénario est basé sur le roman de Gary. La production est assurée par le prestigieux Daryl F. Zanuck et la réalisation par John Huston. Les acteurs les plus connus de l'époque jouent les rôles principaux. Trevor Howard interprète le rôle de Morel et Juliette Gréco incarne celui de Minna. S'y trouvent aussi rassemblés dans ce film, Errol Flynn et Orson Welles. La première a lieu à New-York en octobre 1958. Loin d'être une défense du droit des animaux

dans leur habitat naturel, le film glorifie la chasse aux animaux sauvages; les éléphants d'Afrique ne servent plus que de toile de fond dans ce film. Romain Gary exige que son nom soit complètement rayé du générique.

Bien qu'il se tourmente des déceptions qu'il subit à Hollywood, il apprend par contre beaucoup sur l'industrie cinématographique. Il collabore comme co-scénariste au célèbre film sur l'invasion des forces alliées en Normandie, *Le Jour le plus long*. Il a aussi à cette époque plusieurs aventures amoureuses sur lesquelles sa femme, Lesley Blanch, ferme les yeux.

En 1958 a lieu le putsch d'Alger et de Gaulle revient au pouvoir en France. Pendant la guerre d'Algérie, en bon gaulliste, Romain Gary est obligé de défendre la France colonialiste. En décembre 1958, il fait publier un long article dans *Life Magazine* où il fait un plaidoyer en faveur du Général de Gaulle (Catonné 223).

Gary a toujours été un gaulliste convaincu. Il admire le Général de Gaulle parce qu'il "était pour moi l'image du père héroïque que je n'ai jamais eu" (*Nuit 20*), mais aussi parce que de Gaulle représente l'être extraordinaire qui a refusé de capituler devant la force ennemie: "C'était pour [moi] la situation même de l'homme, la condition de l'homme, et ce refus de capituler, [c'était] à peu près la seule dignité à laquelle [les êtres humains] peuvent prétendre" (*Nuit 21*).

Il compare de Gaulle à Soljénitsyne et à Napoléon, ce dernier étant le symbole le plus puissant de la gloire française. Selon Jean-Marie Catonné, le gaullisme de

Gary ne contient aucune trace de fascisme. Il est plutôt “essentiellement moral; voire métaphysique: accord sur une certaine idée, toute mythique, de la France, fidélité à l'épopée et aux camarades” (48). Mais, le gaullisme de Gary est mal vu par quelques-personnes.

Lorsque Gary rentre chez lui un jour, deux jeunes Algériens l'attaquent à la porte de sa villa. Il est sauvé d'une mort certaine par un ami qui ayant observé la scène de l'intérieur de sa villa, lui ouvre la porte permettant ainsi à Gary de s'enfuir de ses agresseurs.

Pendant toute cette période de tribulations et d'adaptation à son nouveau milieu, en pleine période de la guerre froide, Romain Gary révèle son dualisme et garde une grande nostalgie pour son pays natal, la Russie. Lors de la visite aux États-Unis de Nikita Khrouchtchev, le souvenir de la Russie le rend mélancolique. Lorsque Khrouchtchev visite Twentieth Century Fox à Los Angeles au mois de septembre 1959, Romain Gary s'y trouve aussi. Il admire Khrouchtchev ce “cosaque de Kousrk” qui lui rappelle sa propre ascendance russe et incarne miraculeusement tous les traits caractéristiques de sa race (Bona 197).

Romain Gary admire aussi les États-Unis, le pays des pionniers et des cow-boys à la Gary Cooper, tout en rejetant le côté superficiel de ce pays. C'est ce drame américain qu'il met en oeuvre plus tard dans *Les Mangeurs d'Étoiles*, roman qui fait partie de sa *Comédie américaine*. En 1969, Romain Gary publie chez Gallimard un

autre roman intitulé *Adieu Gary Cooper*, qui parle de l'aliénation, de la séparation et du refus de tendresse dans l'Amérique des années soixante.

2.9. Le "Consul sexuel"

Entre-temps, la femme de Romain Gary, fascinée par le Proche Orient, court à travers le monde pour découvrir l'exotisme et écrire des livres sur la culture et le monde musulmans. Romain Gary se trouve souvent seul à Los Angeles. Durant l'absence de sa femme, comme il est souvent en compagnie de jeunes starlettes, les journaux à sensation de Los Angeles le surnomment "le consul sexuel". De plus, avec son ascendance russe et tartare qu'il revendique, la presse à sensation le transforme en un Atila. Il acquiert la réputation d'être un "macho" qui séduit les jeunes ingénues et les starlettes pour les délaisser peu après.

Dans ses écrits, Gary dénonce cette catégorisation car presque tous ses héros que ce soit Momo (*La Vie devant soi*), Jacques Rainier (*Au-delà de cette limite votre ticket n'est plus valable*), Ludo (*Les Cerfs-volants*), le narrateur de *La Promesse de l'aube* agissent sous le signe de la femme. Dans ses romans, il revendique toujours la féminité: "Un homme qui n'a pas en lui une part de féminité - ne serait-ce que comme un état de manque -, d'aspiration, c'est une demie portion" (*Nuit 19*).

En décembre 1959, lors d'une soirée au consulat de France, Romain Gary rencontre Jean Seberg, la célèbre actrice américaine des années cinquante qui a

incarné le rôle de Jeanne d'Arc d'Otto Preminger et celui de Cécile dans *Bonjour Tristesse* de Françoise Sagan. Elle est accompagnée ce soir-là de son mari, l'avocat François Moreuil. À cette époque Gary est âgé de quarante-cinq ans et Jean Seberg de vingt et un. C'est le coup de foudre des deux côtés.

Jean Seberg fascine Gary par son air de femme-enfant. Très discrets au début de leur liaison, Gary et Seberg affichent leur amour plus ouvertement par la suite. À l'époque, ils forment le couple idéal de l'homme mûr et de la jeune starlette ingénue qui rappelle un autre couple célèbre des années cinquante, le dramaturge américain Arthur Miller et la vedette Marilyn Monroe.

Leur liaison offense les mœurs rigoureuses du Quai d'Orsay de l'époque qui juge qu'une carrière diplomatique n'est guère compatible avec une vie d'adultère. Pour le Général de Gaulle et madame de Gaulle qui considèrent le mariage comme une union sacrée, cette union doublement adultère est scandaleuse. Par ailleurs, la cinquième République n'approuve nullement les fonctionnaires divorcés. Mais, comme le signale Dominique Bona, Jean Seberg ne peut que fasciner Gary:

Sa blondeur et son innocence lui évoquent irrésistiblement toutes les femmes qu'il s'est inventées dès son premier livre, Zosia, Josette, Anne, Minna, qui toutes se retrouvent ou se reconnaissent dans le charme de Jean. Elle correspond si bien à l'idéal féminin de ses

romans, qu'il a l'impression de tomber amoureux de l'une de ses créations, et de voir son rêve prendre corps. (209)

Par ailleurs, Jean Seberg a été élevée dans le culte des acteurs américains tels que James Dean et Marlon Brando. Venant du Mid-Ouest américain (l'Iowa) elle a eu une formation religieuse austère et une enfance protégée, qui la rendent séduisante aux yeux de Romain Gary. Elle a beaucoup de mal à s'accoutumer au milieu hollywoodien et recherche une âme consolatrice qu'elle retrouve en Romain Gary.

Lesley Blanch qui a toujours soutenu Romain Gary dans sa carrière diplomatique demeure sa meilleure amie et sa confidente. Elle a toujours su passer outre aux "indiscrétions" de son mari. Cette liaison, que plusieurs considèrent scandaleuse ne l'est pas pour autant pour Lesley Blanch. D'ailleurs elle conseille au couple de continuer cette union sans mariage, car le divorce serait le suicide professionnel pour Romain Gary (Bona 222).

Cédant sans doute à la pression, Jean Seberg décide de se séparer de Romain Gary pour aller tourner à Paris le film *À Bout de Souffle* avec Jean-Paul Belmondo. Ce film est loué par Jean Cocteau et Jean-Paul Sartre lorsqu'il débute à Paris. Jean Seberg est fêtée et adulée par la presse parisienne dès la première du film. Tout Paris répète la fin du film où Jean Seberg dit dans son accent américain: "Dis Michel, qu'est-ce que ça veut dire ... dégueulasse?"

Pendant ce temps, Romain Gary se trouve à Los Angeles, où pour se divertir il écrit *La Promesse de l'aube*. Il y raconte l'histoire de sa mère tant aimée dont il en fait l'héroïne. Il évoque aussi Roman Kacew dès son enfance en Pologne jusqu'à ses expériences en France. Cette oeuvre si émouvante connaît un succès fou en France. Même le Général de Gaulle prend partie pour ce livre et, le 3 juillet 1960, il envoie la communication suivante au consulat de France à Los Angeles:

Mon cher Romain Gary,

La Promesse de l'Aube est un livre excellent. Je l'ai lu avec plaisir et non sans émotion. Le talent y est éclatant [...] moi, pour toutes les raisons qui furent et qui sont "nôtres", je me félicite de voir se développer votre grande carrière littéraire. (Bona 269)

Romain Gary retourne à Paris pour la parution de son roman, mais aussi, sans doute, pour retrouver Jean Seberg. Ils entreprennent ensemble un long voyage en Extrême-Orient. Les deux amants se croient contraints de démentir les rumeurs de divorce qui circulent dans la presse, tout en jouant à cache-cache avec les journalistes.

Tout semble annoncer que Romain Gary ne sera jamais nommé ambassadeur de France. Un de ses collègues, Maurice Couve de Murville, déclare même que la conduite de Gary est malséante, et s'oppose catégoriquement à toute promotion que le corps diplomatique pourrait éventuellement accorder à Romain Gary. En 1986,

(plusieurs années après) de Murville exprime encore son mépris pour la conduite de Romain Gary.

2.10. Vie commune

Avec la parution de *La Promesse de l'aube*, Romain Gary est fêté et adulé. En septembre 1960 Jean Seberg obtient le divorce, mais Romain Gary hésite de divorcer pour ne pas blesser Lesley Blanch sans doute, mais aussi pour ne pas dire adieu à sa carrière. Au début de 1961 le couple vit en concubinage dans un grand appartement à Paris, rue du Bac. C'est une période paisible pour Gary et Seberg qui est néanmoins agitée par le fait que Lesley Blanch refuse de divorcer d'avec Romain Gary. Elle lui accorde le divorce lorsqu'elle apprend que le couple Gary/Seberg attend un enfant.

Ayant préféré l'amour à une carrière diplomatique, Romain Gary se consacre entièrement à sa vocation littéraire. Il explique sa décision plus tard en ces termes: "Je ne pouvais plus travailler auprès de Charles de Gaulle parce que je voulais garder ma liberté sexuelle" (*Nuit* 320). Il écrit aussi à cette époque un autre roman directement en anglais qui s'intitule *Lady L* (un cadeau qu'il offre à Lesley Blanch et qui l'a inspiré).

C'est l'histoire d'une vieille aristocrate anglaise qui raconte en rétrospective sa vie et ses amours. Métro-Goldwyn Mayer, après avoir acheté les droits du roman, en produit une version cinématographique. Sophia Loren joue le rôle de Lady L. Se

trouvent aussi réunis dans ce film, le célèbre acteur américain Paul Newman et la grande vedette du cinéma français, Philippe Noiret.

En été 1963, Romain Gary et Jean Seberg sont invités à dîner à La Maison Blanche par le Président Kennedy et son épouse. Le 16 octobre 1963, le couple se marie; quelque temps après naît leur fils Alexandre-Diégo. Dorénavant, Romain Gary suit sa femme partout où elle tourne des films que ce soit en Suisse, à Marrakech ou en Amérique où elle incarne le rôle de "Lilith." La famille Gary est toujours accompagnée de la nourrice de Diego, Eugénia de Muñoz, une Espagnole à qui Romain Gary dédiera plus tard son plus gros livre, *Pour Sganarelle*. Dans ce roman, il élabore sa pensée au sujet du personnage et de l'auteur et donne des définitions de l'art, de la culture et du roman:

Jouant les Lacan, les Barthes, par gageure, mais jouant cependant dans la clarté d'un langage simple, il ouvre son champ à l'illimité, refuse les lois et les frontières au nom de la liberté de l'artiste. (Bona 249)

La famille Gary s'installe à Puerto de Andraitz au sud-est de l'île de Majorque, dans une maison que Gary fait construire et qu'il surnomme *Cimarron*. C'est une maison spacieuse qui ressemble à un château maure: ils y mènent une vie paisible et heureuse. Mais comme la municipalité a construit un égout qui se déverse à côté de la plage de Romain Gary, la famille déménage et vend la maison. Pendant le reste de sa vie, Gary cherche une demeure au soleil qui lui rappelle *Cimarron*. Il voyage

constamment à travers le monde, de la Grèce à l'Île Maurice à la recherche de cet autre *Cimarron*. Finalement il décide de s'établir à Jouani dans le Lot.

En décembre 1966 le Général de Gaulle est réélu, et en mars les gaullistes gagnent de justesse aux élections législatives. Au mois de mai 1968, le Général de Gaulle est déstabilisé. En juin 1968, Gary écrit un article dans *Le Monde* où il se déclare inconditionnellement gaulliste.

La deuxième partie des années soixante est une période de révolution de mœurs, d'une prise de conscience et d'une remise en question des valeurs traditionnelles par les jeunes en Europe et aux États-Unis. Jean Seberg qui appartient à cette génération de jeunes subit les influences de ce courant envahissant. Romain Gary est lui aussi marqué, quoique indirectement, par ce remous social.

Lorsque le 4 avril 1968, Martin Luther King est assassiné au balcon d'un motel où il séjournait à Memphis (Tennessee), Jean Seberg est à Hollywood en train de tourner le film *La Kermesse de l'Ouest*. Redevenue vedette célèbre aux États-Unis, elle décide d'y demeurer d'autant plus qu'elle rencontre une foule de contestataires et de hippies s'opposant à la guerre du Vietnam, et, avec lesquels elle sympathise. Son engagement politique commence à se cristalliser à partir de ce moment. Comme Romain Gary refuse de s'établir à Los Angeles, il est souvent obligé de se déplacer entre Paris et Los Angeles pour voir sa famille.

Étant donné que de longues absences les séparent, Romain Gary commence à douter de la fidélité de Jean Seberg. Il soupçonne Clint Eastwood, le partenaire de sa femme dans *La Kermesse de l'ouest* d'en être l'amant et veut même provoquer Eastwood en duel (Catonné 232).

Bientôt des circonstances indépendantes de leur volonté séparent Romain Gary et sa femme. Jean Seberg est née en Iowa, aux États-Unis où a eu lieu dans les siècles précédents l'irrésistible poussée vers l'Ouest. Elle est sans doute au courant du génocide qui s'est produit dans l'état d'Iowa au nom du progrès de l'homme blanc. Peut-être, comme le soutient Catonné, appartient-elle à cette bande de jeunes aux prises avec une culpabilité indéfinie et sans solution, et qui ne peut trouver la paix dans le silence et l'indifférence. Peut-être souffre-t-elle de ce crime sans visage, dont ni elle ni ses contemporains ne sont responsables. Comme plusieurs jeunes Américains des années soixante, elle ressent une culpabilité "impersonnelle, illimitée et inexpiable d'un dévouement poussé jusqu'au sacrifice [...] comme victime d'un péché originel jamais désigné" (Catonné 65).

Jean Seberg s'engage dans la lutte antiraciste aux États-Unis et, comme plusieurs autres vedettes de Hollywood, se déclare ouvertement contre la guerre au Vietnam; elle se livre corps et âme à cette cause. Son foyer devient le lieu de rencontre de militants blancs aussi bien que noirs. Son mari n'approuve nullement ce qui se passe dans sa maison à Los Angeles, bien qu'il se soit déclaré du côté des

opprimés. Mais comme il l'exposera dans son roman, *Chien Blanc*, il voit dans cette lutte antiraciste par les Blancs, de l'exploitation par des parasites: à savoir, les producteurs et vedettes milliardaires qui, tout en se déclarant antiracistes, font de la publicité sur le dos des Noirs. D'ailleurs son opinion des Noirs n'est guère plus favorable. D'après lui, ceux-ci, tout en acceptant l'aide des Blancs, ne peuvent s'empêcher de les insulter une fois que ces derniers tournent le dos.

Selon l'optique de Romain Gary, tous ceux qui fréquentent sa maison, Noirs et Blancs, profitent de la générosité de sa femme. De plus, le remords de conscience des Américains blancs n'intéresse nullement Romain Gary, car lui s'intéresse plutôt à l'être humain quelle que soit sa couleur, quel que soit son sexe, quelle que soit sa race. Comme l'indique Dominique Bona, "Il refuse de partager la culpabilité des Blancs face aux Noirs comme il refuse de porter l'étoile jaune de la vengeance et de la haine" (283).

Entre-temps, Jean Seberg, milliardaire et jeune femme au bon coeur, impliquée dans la lutte des Noirs, ne peut refuser ni son argent ni son temps à la cause. De plus en plus frustré et exaspéré, Romain Gary affirme: "Je n'en peux plus. Dix-sept millions de Noirs américains à la maison, c'est trop, même pour un écrivain professionnel" (*Chien* 120).

Jean Seberg refuse de renoncer à ses convictions politiques et philosophiques et à sa lutte en faveur des Noirs. Elle s'associe même aux activistes noirs

d'Amérique, les "Panthères Noires" qui portant béret, blouson noir, verres fumés et fusils, épouvantent la majorité des Blancs. S'attachant ouvertement à ces groupes de militants, Jean Seberg reçoit plusieurs menaces de mort par téléphone; quelqu'un tire une balle par la fenêtre de sa cuisine.

C'est à cette époque que le couple Gary/Seberg décide de divorcer à l'amiable; Romain Gary, voulant demeurer le plus loin possible de ce grand bouleversement, s'établit encore une fois à Paris. Mais, en septembre 1968, lorsque Romain Gary et Jean Seberg se séparent, ils ne peuvent pas pour autant vivre loin l'un de l'autre. D'ailleurs Romain Gary l'immortalise dans ses romans:

Jean restait proche [...]. Sa silhouette pathétique va hanter plusieurs [de ses livres]. Jean est présente, reconnaissable et immortalisée dans *Les Mangeurs d'étoiles*, *La Danse de Gengis Cohn*, *La Tête coupable*, *Adieu Gary Cooper*, *Chien Blanc*, *Les Têtes de Stéphanie* et *Clair de femme*. (Pavlowitch 30)

Jean Seberg vient souvent avec son fils rejoindre Romain Gary à Paris, où ils reprennent la vie de couple. Lorsqu'elle se retrouve enceinte une deuxième fois, le couple est heureux, quoique leur divorce soit sur le point d'être prononcé. Un article dans la revue *Newsweek* met en doute la paternité de l'enfant. Seberg dénonce catégoriquement cette insinuation, mais plongée dans un état de profond désarroi, elle

fait une fausse couche le 23 août 1970. Gary et Seberg nomment le bébé Nina Hart Gary et l'enterrent au cimetière de Riverside à Marshalltown dans l'Iowa.

2.11. Le Solitaire à travers le monde

En novembre 1970, la France est en deuil car le Général de Gaulle vient de mourir. Trois cent cinquante Compagnons de la Libération participent aux obsèques officielles à Colombey-les-deux-églises. Romain Gary fait partie de la délégation. Il porte, pour cette occasion, sa vieille vareuse d'aviateur, ce qui le distingue des autres compagnons. Il porte aussi fièrement sa médaille de la Légion d'honneur, sa Croix de la Libération et sa Croix de Guerre. On l'attaque de tous les côtés: on lui reproche surtout son exhibitionnisme et son manque de pudeur.

Romain Gary a toujours admiré le Général de Gaulle. Le 20 novembre, il fait l'éloge du Général dans la revue *Life Magazine* dans l'article intitulé "*Ode to the man who was France*" (Ode à l'homme qui était la France):

Pour la première fois dans les temps modernes les Français ont un modèle. La plus grande oeuvre de de Gaulle sera peut-être posthume, sa mort marquera plus profondément le pays que sa politique même [...]. De Gaulle sera peut-être plus puissant mort que vivant. (Bona 266)

Romain Gary âgé alors de cinquante-sept ans est à nouveau divorcé. Pour retrouver son équilibre mental, il parcourt le monde: il voyage tour à tour à l'Île Maurice, à Mombasa, aux Seychelles, à Kuala Lumpur, à Singapour, à Bangkok, en Pologne ou à Mexico. Il fait publier des articles sur les lieux qu'il visite dans les journaux et les revues tels que *Life Magazine* et *Travel and Leisure*. Chaque voyage lui apporte aussi des expériences qui figurent, plus tard, dans ses romans.

Souvent, lorsque Gary retourne chez lui, il retrouve Jean Seberg vieillie et fatiguée malgré ses trente-deux ans, et perturbée psychologiquement. Pour lui redonner goût à la vie, il lui fait interpréter le rôle du personnage principal dans un de ses films qui s'intitule *Kill*. Ce film a aussi pour interprètes principaux James Mason et Stephen Boyd. C'est un film de suspense baigné d'humour noir, voire macabre. Jean Seberg y interprète le rôle douloureux d'Emily, une nymphomane qui, désespérée et troublée, porte au début du film une perruque de style africain. Lorsqu'elle est accusée par son mari de vivre en pleine négritude, elle tire au pistolet sur une cible suspendue au mur. D'après Bona c'est un geste de révolte et une revendication du passé gauchiste et pro-Panthères Noires de Jean Seberg (300). Ce film a eu très peu de succès en France.

En avril 1971, Romain Gary intente un procès à la revue *Newsweek*, en France, où il met en cause la méchanceté de la presse et la puissance de l'argent. Il accuse la revue *Newsweek* d'être responsable de la mort de sa fille. Mais, la cour rejette

l'accusation de Gary. Malgré toute cette contrariété, Jean Seberg se remarie avec un jeune cinéaste et continue à vivre très près de son ex-mari.

Romain Gary se consacre entièrement à la rédaction des romans. Comme il frise la soixantaine, il a des crises d'angoisse qui provoquent une hypocondrie grandissante. En 1972, il publie *Europa* et en 1973, il publie *Les Enchanteurs*.

C'est à ce moment que paraît à Paris, *Gros-Câlin*, d'un nouvel auteur qui se nomme Émile Ajar. Ce roman raconte les aventures de Michel Cousin et révèle l'angoissante solitude d'un être réduit à cohabiter avec un python. Le texte à la fois tendre et sérieux fait réagir le monde littéraire d'autant plus que son auteur semble vouloir garder l'anonymat le plus complet. La presse française soupçonne qu'Émile Ajar est probablement le pseudonyme d'un auteur connu. Elle nomme même une fois Romain Gary comme le véritable Ajar, mais celui-ci s'empresse de nier catégoriquement cette insinuation.

Étant donné que Gary veut se camoufler sous ce nom fictif d'Ajar, il invente l'une des plus grandes mises en scène du monde littéraire. Il charge son neveu Paul Pavlowitch de se faire passer pour Émile Ajar, le pseudonyme d'un nommé Hamil Raja. Cette histoire réussit si bien que, du vivant de Romain Gary, personne ne devine qu'Émile Ajar est le pseudonyme de Romain Gary. Tous croient qu'Émile Ajar est le nom de plume de Raja et plus tard de Pavlowitch.

L'année suivante, Émile Ajar réapparaît sur la scène parisienne avec la publication du roman *La Vie devant soi* qui s'impose comme le roman de l'année. Le livre émeut le public parisien tant par sa franchise au sujet de la vieillesse que par son langage. Le langage argotique et imagé qui reflète le langage de l'immigré et de l'enfant de la rue, séduit tout le monde. La critique parisienne est encore une fois divisée quoique plusieurs considèrent que c'est un roman extraordinaire.

Malgré son éclatant succès, Émile Ajar demeure introuvable, ce qui provoque désormais une chasse à l'auteur. Quelques temps après, le Prix Goncourt est attribué à Émile Ajar pour *La Vie devant soi*. Interrogé avec d'autres anciens lauréats sur le Goncourt attribué à Émile Ajar, Romain Gary offre la réponse suivante: "J'avais aimé *Gros-Câlin*, je n'ai pas encore lu *La Vie devant soi*. Je ne crois pas que son auteur continuera à préserver pour longtemps son anonymat" (Bona 348).

En 1975, Romain Gary publie chez Gallimard, *Au-delà de cette limite votre ticket n'est plus valable* sous son propre nom. Ce roman raconte les angoisses d'un homme qui, approchant la soixantaine, craint de vieillir et de perdre sa virilité. À sa parution, ce livre choque les lecteurs, compte tenu de sa franchise sur un sujet tabou (la défaillance sexuelle des sexagénaires) et de son langage cru, grossier et même vulgaire à certains endroits. Les lecteurs prennent ce roman pour une confession publique de Gary de sa propre impuissance.

À la même époque il écrit un autre roman, *Pseudo*, qu'il fait publier en novembre 1976 sous le pseudonyme d'Ajar. Ce roman, écrit à la première personne, prétend être les divagations de l'auteur Émile Ajar qui se fait soigner pour troubles mentaux dans une clinique en Suisse.

En 1977 Romain Gary publie *Clair de femme* sous son propre nom. Cette fois-ci certains critiques affirment qu'il plagie Émile Ajar. En effet, vers la fin des années soixante-dix, Émile Ajar est mieux connu que Romain Gary. C'est aussi à cette époque que le nihilisme et le non-conformisme de Gary sont de plus en plus évidents. En 1977, dans une interview accordée à *Paris-Normandie*, en faisant allusion à son roman *Clair de femme*, il déclare:

Il y a dans ce roman la dérision et le nihilisme qui guettent notre foi humaine et nos certitudes sous le regard amusé de la mort. Les dieux païens nous guettent installés sur l'Olympe de nos tripes. Notre vie n'est peut-être que le divertissement de quelqu'un. (23)

En 1978 Romain Gary fait publier *Charge d'Âme* sous le nom de Gary et, en 1979, il publie son quatrième roman signé Émile Ajar, *L'Angoisse du roi Salomon* qui raconte l'histoire d'un ancien roi du prêt-à-porter, Salomon Rubinstein. Comme *Au-delà de cette limite votre ticket n'est plus valable* et *La vie devant soi*, *L'Angoisse du roi Salomon* raconte une tendre histoire d'amour entre "le roi du pantalon" de quatre-vingt-quatre ans et une ancienne chanteuse de cabaret de soixante-cinq ans.

2.12. Déclin et Mort

Avec la parution du dernier livre signé Émile Ajar, la critique française est plutôt favorable envers Émile Ajar et au mois de juillet, ce roman intéresse même les producteurs de cinéma. Mais voici que Romain Gary subit le plus grand chagrin de sa vie, après celui causé par la mort de sa mère: le suicide de Jean Seberg. Elle s'était séparée de son troisième mari en janvier 1979 et vivait seule. Elle avait disparu depuis dix jours, mais le huit septembre 1979, son corps, recouvert d'une couverture, est retrouvé dans sa voiture garée dans la rue du général Appert, dans le XVI^e arrondissement. Elle s'est suicidée en prenant des barbituriques. Elle a trouvé la mort lors de sa huitième tentative de suicide. Après l'autopsie les médecins concluent publiquement qu'aucune réponse définitive ne peut être prononcée sur sa mort.

Deux jours après le décès de Jean Seberg, Romain Gary tient une conférence de presse dans son bureau chez Gallimard. Il dément catégoriquement la rumeur selon laquelle Jean Seberg s'est suicidée sous le coup d'une dépression nerveuse, après avoir vu la version cinématographique de *Clair de femme*. Par contre il accuse nettement le F.B.I. d'avoir contribué à sa mort:

Jean Seberg a été détruite par le F.B.I. [...] cet organisme a fait paraître dans un grand journal américain (que) Jean attendait un enfant d'un extrémiste des "Black Panthers." [...] cette calomnie a rendu Jean malade, et l'enfant, une petite fille, est morte à sa naissance. Dès lors,

Jean est devenue psychotique. Chaque année, à l'anniversaire de cette fausse-couche, elle tentait de se donner la mort. (Bona 379)

Le 14 septembre, Jean Seberg est enterrée au cimetière Montparnasse. Son fils Alexandre-Diégo, ses trois maris, plusieurs amis et des photographes sont à l'enterrement. Ils déposent chacun une rose sur son cercueil (Bona 381).

En 1980 Romain Gary vit avec Leïla Chellabi, une divorcée d'origine turque âgée de quarante ans. C'est à cette époque qu'il publie son dernier roman *Les Cerfs-Volants*. Ce roman raconte l'histoire du jeune orphelin Ludo, élevé en Normandie par son oncle Ambroise Fleury; il ne parle ni de mère excessive, ni de la peur de vieillir. Ce dernier roman de Romain Gary est un livre paisible: il ne contient ni haine ni rancune, car Gary ne divise plus le monde entre les bons et les méchants, entre les "compagnons" et les salauds; il existe au contraire des résistants ridicules et des collaborateurs sympathiques, des Allemands humanistes et des Français barbares.

Après la publication de ce livre et suite à la mort de Joseph Kessel, Maurice Shumann, sénateur et ancien député du Nord, Compagnon de la Libération et membre de l'Académie française depuis 1974, propose à Romain Gary un fauteuil à l'Académie française. Gary refuse de poser sa candidature. Il refuse aussi cette année-là, le Prix Paul Morand que décerne l'Académie pour la première fois à un auteur français reconnu pour "la qualité de sa pensée, de son style et pour son esprit d'indépendance et de liberté" (Bona 384).

Dominique Bona suggère que Paul Morand, avec sa personnalité "pétainiste" (384), heurte sans doute l'orgueil du vieux Résistant chez Romain Gary. Une autre raison à ce refus est, sans doute, qu'après toutes les gloires qu'il a vécues, les nouvelles gloires littéraires ne l'intéressent plus.

C'est à ce moment que les relations entre Romain Gary et Paul Pavlowitch se dégradent. Paul Pavlowitch réclame une augmentation de salaire pour sa participation à la grande comédie que Gary a orchestrée. De plus, Pavlowitch tient à incarner Émile Ajar à sa manière. Pour comble de malheur le contrôle fiscal commence à inquiéter Romain Gary.

Dans *La Promesse de l'aube*, *La Nuit sera calme*, *Clair de femme* et *Europa*, Romain Gary évoque le suicide. Dans *La Promesse de l'aube*, il raconte qu'il a failli se suicider au moins trois fois dans sa vie: une fois quand il était enfant en Pologne, une autre fois quand il était adolescent à Nice, et une troisième fois encore durant la guerre lorsqu'il était en Afrique. D'après un de ses amis, Jean Béliard, durant les années soixante alors que Romain Gary était diplomate à New-York, une fois dans la nuit, Gary l'a appelé au téléphone annonçant qu'il allait se tuer. Mais "Béliard arriva à temps à l'hôtel Plaza pour empêcher Gary de se jeter par la fenêtre" (Bona 234).

Vers la fin de novembre 1980, Gary est rempli d'angoisse suite aux contrecoups qu'il a subis, surtout le suicide de Jean Seberg. Il est pris dans son propre piège Ajar; il est atteint d'une hypocondrie grandissante et se trouve forcé d'affronter

les symptômes de la vieillesse. Le 2 décembre 1980, Gary réussit à sa dernière tentative de suicide. Allongé sur son lit par un après-midi pluvieux à Paris, il se tire une balle de revolver à la gorge. Lorsque sa compagne Leïla Chellabi découvre son corps, sa tête est recouverte d'un peignoir de bain rouge qu'il a placé là par pudeur sans doute pour ne pas causer d'effroi à sa compagne et à son fils Diégo. Malgré le fait qu'il s'est tiré une balle à la gorge, sa tête est restée intacte.

Gary laisse au pied de son lit une lettre dont il a envoyé une copie à Claude Gallimard pour que le contenu soit rendu public. C'est dans cette lettre (qui est par la suite publié sous le titre de *Vie et mort d'Émile Ajar*) qu'il dévoile pour la première fois que, comme l'avaient soupçonné certains, Émile Ajar et Romain Gary ne sont que deux aspects du même auteur. Il laisse aussi une autre lettre à Leïla Chellabi, sa compagne, dans laquelle il pose la question suivante:

Alors pourquoi? Peut-être faut-il chercher la réponse dans le titre de mon ouvrage autobiographique *La Nuit sera calme* et dans les derniers mots de mon roman: "car on ne saurait mieux dire" [...] Je me suis enfin exprimé entièrement. (Bona 398)

Quoiqu'il en soit, la nouvelle du suicide explose dans la France étonnée et incrédule. Son suicide ne semble pas être l'effet du cafard mais a l'air prémédité. Par ailleurs, quatre jours auparavant Gary avait envoyé à Raymond Aron, un message destiné à la presse où il soutient qu'il n'y a aucun rapport entre son suicide et celui de

Jean Seberg (Catonné 128). De plus, au bas de la lettre envoyé chez Gallimard, il tient encore une fois à avoir le dernier mot en déclarant: "Je me suis bien amusé. Au revoir et merci" (*Vie et mort* 43). Le 9 décembre dans l'après-midi, onze aviateurs portent le cercueil de Romain Gary aux Invalides et il est inhumé au Cimetière Père-Lachaise.

En se suicidant, Romain Gary tue à jamais la légende Émile Ajar, et comme le proverbial phénix, Romain Gary renaît de ses cendres. Le 3 janvier 1981, Paul Pavlowitch annonce à l'émission télévisé *Apostrophes* qu'il n'est pas Émile Ajar. Le testament littéraire *Vie et mort d'Émile Ajar* publié à titre posthume en juillet 1981 établit une fois pour toutes la paternité de l'oeuvre Ajar. Dans ce petit livre d'une cinquantaine de pages, Romain Gary revient sur l'unité de l'oeuvre entière et en signale les constantes stylistiques et structurales.

Gary souligne que dans ses romans signés Gary ou Ajar, il oppose souvent des couples antithétiques: le maître/le valet; le Juif/le Nazi; Madame Rosa, la Juive/Momo, l'orphelin arabe. Il ajoute:

aucun des critiques n'avait reconnu ma voix dans *Gros Câlin*. Pas un dans *La Vie devant soi*. C'était, pourtant exactement la même sensibilité que dans *Éducation européenne*, *Le Grand Vestiaire*, *La Promesse de l'aube*, et souvent, les mêmes phrases, les mêmes tournures, les mêmes humains. Il eut suffi de lire *La Danse de Gengis*

Cohn pour identifier immédiatement l'auteur de *La Vie devant soi*. Les jeunes gens amis du jeune héros de *L'Angoisse du Roi Salomon* sont tous sortis d'*Adieu Gary Cooper*: [...]. Tout Ajar est déjà dans *Tulipe*.
(*Vie et mort* 18)

Il déclare en outre qu'il n'a jamais redouté qu'une simple et facile analyse de texte ne vienne le tirer de son anonymat (*Vie et mort* 18). Les critiques remarquent alors les nombreuses similarités entre l'oeuvre signée Gary et celle signée Ajar, similarités qui seront examinées en détail au chapitre 5.

Après ces précisions biographiques sur la vie de Romain Gary, le prochain chapitre aborde l'intrigue des quatre romans étudiés, et les trois chapitres qui le suivent examinent en détail le subterfuge Ajar.

Chapitre 3

Les Romans analysés

Le présent chapitre se propose de donner un aperçu de l'intrigue des romans étudiés, en commençant par *Au-delà de cette limite votre ticket n'est plus valable* (1975) et en terminant par *La Vie devant soi* (1975). Les raisons pour lesquelles ces romans figurent dans la présente étude feront l'objet de la deuxième partie du chapitre.

3.1. *Au-delà de cette limite votre ticket n'est plus valable* / Gary (1975)

Jacques Rainier, un industriel français vivant à Paris après la défaite du gaullisme, a des ennuis financiers suite à la crise du pétrole en Europe et à la restructuration industrielle en France sous le giscardisme. Il est divorcé et a un fils Jean-Pierre, âgé de vingt-neuf ans. Au cours d'un voyage à Venise, Jacques Rainier rencontre Jim Dooley, une ancienne connaissance et millionnaire Texan qui a dépassé la soixantaine. Dans un moment de complicité masculine, Dooley avoue ses angoisses et ses défaillances sexuelles.

À la suite de ces confidences, Jacques Rainier, qui a cinquante-neuf ans, est atteint lui aussi de doutes obsédants sur sa propre virilité; la hantise de l'impuissance sexuelle l'envahit complètement. Son angoisse résulte surtout du fait qu'il est amoureux de Laura, une jeune Brésilienne brune de vingt-trois ans, même si tout s'est toujours bien passé entre Laura et lui jusque-là, malgré leur différence d'âge. Il veut cacher à Laura sa fragilité d'homme d'un certain âge, mais petit à petit il éprouve des problèmes sexuels. Il a alors une grande peur de la perdre.

Lorsque Rainier raconte ses inquiétudes au Frère Jacques, celui-ci lui dit qu'il exploite Laura, que cette liaison est sans avenir parce que lui Rainier est "en déclin [...] en baisse" (87). Le Frère Jacques lui propose alors d'aller voir le docteur Mingard, un sexologue gériatrique. Le médecin lui suggère de lire le manuel *Races et phantasmes* et explique que les Occidentaux font "souvent appel aux Noirs et aux Arabes dans leurs phantasmes sexuels" (173) et que lui Rainier devrait en faire de même lorsqu'il fait l'amour à Laura. D'après le médecin, ceci est dû au fait que les satisfactions les plus élémentaires qui sont à la portée de ces travailleurs africains pauvres échappent parfois à un homme cultivé comme Rainier.

Suivant les conseils du docteur Mingard, lorsqu'il couche avec Laura, Jacques Rainier essaie de faire appel aux pauvres immigrants africains, aux arabes ou mêmes aux bêtes dans ses phantasmes. Il y renonce toujours, de peur d'exploiter ces derniers et de les

réduire au niveau de la brutalité animale, aussi bien que par une certaine pudeur idéologique (171).

Rainier crée alors le personnage de Ruiz, un voyou d'origine espagnole, qui ressemble un peu à Rudolphe Nouréïev, mais qui est encore plus jeune et plus sauvage, avec les traits "d'un de nos conquérants" (220). Dans ce phantasme, Ruiz apparaît dans la chambre de Rainier la nuit pour le cambrioler. Il porte un maillot de corps blanc et un pantalon de cuir noir, l'image même de la virilité: "S'il m'était permis de 'redevenir' c'est cet article-là que j'aurais choisi. Jamais je n'avais vu un visage aussi proche à la fois de la sauvagerie primitive et de la finesse frémissante" (172). Jacques Rainier a alors recours aux services de Ruiz; sur son ordre, Ruiz possède Laura brutalement. Jacques Rainier ne peut pas se passer de Ruiz et le cherche parfois dans les endroits les plus malfamés de la Goutte d'Or.

Parfois dans sa hantise, Rainier croit qu'il devrait peut-être céder sa place à son fils Jean-Pierre qui est plus près de l'âge de Laura. Il craint que Laura et Jean-Pierre s'aiment mais qu'ils n'osent le lui avouer de peur de le blesser: "J'avais cru remarquer que Jean-Pierre la regardait avec beaucoup d'amitié, peut-être même de tendresse" (188). Il organise alors un déjeuner chez lui, afin d'observer entre Laura et Jean-Pierre "un échange de regards ou un moment de silence un peu gêné, lorsqu'on découvre l'autre pour la première fois" (188). Mais rien de tel ne se produit et lorsqu'il dit en plaisantant à

Laura qu'il voudrait arranger un mariage entre Jean-Pierre et elle, celle-ci lui répond stupéfaite: "Mon Dieu, quelle horreur!" (189).

À la fin du roman Rainier se rend compte du fait que Laura l'aime malgré ses défaillances sexuelles. Il comprend aussi que les hommes plus jeunes que lui n'intéressent guère Laura, qu'elle veut tout simplement "sentir à ses côtés la présence d'un homme mûr et réfléchi, se laisser guider par une main expérimentée" (189), comme celle de Jacques Rainier. Il accepte alors son bonheur avec Laura, même si de temps en temps il doit avoir recours aux services de Ruiz.

Romain Gary avait abordé le thème du dysfonctionnement sexuel avant *Au-delà de cette limite votre ticket n'est plus valable* dans le film *Les oiseaux vont mourir au Pérou* (1968) et le roman *La Danse de Gengis Cohn* (1967), mais dans ces deux cas, il s'était concentré sur la frigidité des femmes. Dans *Au-delà de cette limite votre ticket n'est plus valable*, Gary déculpabilise la femme et traite l'un des tabous les plus intimes: la vulnérabilité par laquelle le temps frappe l'homme en premier. Il le fait dans un langage cru, franc et même obscène. Mais, en traitant l'impuissance sexuelle chez les hommes, Gary traite en même temps le déclin de la jeunesse que glorifie à l'excès la société occidentale.

Puisque le roman traite un sujet presque tabou, il a soulevé un grand débat en France et a connu un grand retentissement. Même s'il aborde le sujet de la sexualité, il n'est cependant pas un roman érotique.

3.2. *Clair de femme* / Gary (1977)

Michel Folain âgé de quarante-cinq ans est pilote de ligne pour *Air France*. Une nuit, il erre dans les rues de Paris en attendant de prendre son vol pour Caracas le lendemain matin. Yannik, sa femme tant aimée, se donne la mort cette nuit même. Atteinte d'un cancer, elle a choisi de se suicider avant le déclin total.

Yannik a dit préalablement à Michel de ne pas vivre seul et lui a demandé de trouver une autre femme à aimer; car aimer une autre femme, c'est triompher sur la mort. "La plus cruelle façon de m'oublier, ce serait de ne plus aimer" (53) lui a-t-elle dit. Elle pourra ainsi survivre grâce à cette autre femme.

Par le plus simple des hasards Michel rencontre Lydia Towarski, une femme de quarante-trois ans aux yeux noirs, aux pommettes saillantes et aux longs cheveux blancs. Michel Folain va passer quelques heures chez Lydia; il lui parle de Yannik et de l'espoir de celle-ci qu'il continuerait à l'aimer à travers une autre femme: "Vous ne saurez jamais comment elle croyait en vous et vous faisait confiance" (141). Ce même soir, il a des rapports sexuels avec Lydia; celle-ci devient alors cette autre femme que Yannik avait recherchée pour Michel et aide ce dernier dans ce moment atroce.

Lydia, elle aussi, est meurtrie par le destin. Elle emmène Michel rencontrer son mari Alain qui ne vit plus avec elle, mais habite chez sa mère, Mme Sonia Towarski, dans un grand appartement, avenue Malesherbes. Le mari, blessé dans un accident de voiture qui a causé la mort de leur petite fille, est maintenant un aphasique ridicule et suicidaire

qui ne communique que par monosyllabes incompréhensibles. Il célèbre ce soir même son anniversaire. Lydia avoue alors à Michel qu'elle aime beaucoup cet homme, bien qu'elle le tienne responsable de la mort de leur fille et qu'elle "essaie donc de l'aimer encore plus fort" (69). Lydia fait comprendre à Michel que sa situation à lui n'est pas aussi unique qu'il ne le croyait.

Quelque temps après le suicide de Yannik, lorsque Michel arrive à son appartement accompagné de Lydia, tout le monde le prend pour un goujat. Mais Lydia et lui savent qu'il n'a fait que respecter les vœux de Yannik en recherchant la compagnie d'une autre femme. À la fin du roman Lydia se sépare de Michel pour lui donner le temps de s'adapter à sa nouvelle situation de veuf. Elle lui offre, néanmoins, une fraternité dans le malheur d'où naîtra peut-être quelque chose dont avait rêvé Yannik.

Clair de femme est un hymne à l'amour du couple: il traite un sujet qui de prime abord paraît mélodramatique, mais ne l'est pas. Dans ce roman l'unité du couple représente une valeur bien vivante. Au-delà de toute notion sexiste, séparatiste, le couple est une unité indissoluble représentant "une autre espèce, un autre sexe, un autre pays" (129); il doit toujours se refaire. Ainsi, s'il s'agit de la mort du couple Yannik/Michel dans *Clair de femme*, la permanence du couple est assurée par le couple Lydia/Michel. La rencontre entre Michel et Lydia leur permet de triompher de la mort de Yannik et de former un autre couple.

3.3. *Gros-Câlin* / Ajar (1974)

Gros-Câlin est un "traité sur les pythons" (10) raconté par Michel Cousin, un jeune employé du département des statistiques au Jardin des Plantes. Michel Cousin, qui est sans famille et sans amis, souffre de troubles de la personnalité; il passe ses journées "à compter par milliards" (21), et n'arrive pas à se situer dans le monde: "Je me (tiens) là discrètement, avec mon petit chapeau, mon noeud papillon jaune à pois bleus, mon cache-nez et mon pardessus, très correctement habillé, veston, pantalon et tout, à cause des apparences et de la clandestinité" (21).

Comme Cousin cherche quelqu'un à aimer, il tombe amoureux de la belle Mlle Dreyfus, une de ses collègues, et a l'intention de l'épouser. Cette dernière, une Noire de la Guyane française, s'habille de façon assez provocante portant minijupes et bottes à mi-cuisses. Se sentant repoussé dans ses avances amoureuses, Cousin nourrit un amour platonique pour Mlle Dreyfus et fréquente celles qu'il appelle "les bonnes putes" (34).

Michel Cousin se console alors avec un python sans venin, Gros-Câlin, qu'il a ramené clandestinement à Paris, suite à un voyage organisé qu'il a fait à Abidjan. Ce python aime "s'enrouler affectueusement autour de vous, des pieds à la tête" (14). Cousin qui est un véritable original se promène sur les Champs-Élysées, son python enroulé autour du cou. Lorsque son chef de bureau lui demande pourquoi il a adopté un python,

il répond le plus naturellement du monde: "Les pythons sont très attachants. Ils sont liants par nature" (15).

Le dilemme de Cousin est qu'il sait que la plupart des femmes ne voudraient pas vivre dans un appartement avec un python car "cela demande beaucoup de tendresse et de compréhension" (61). Mlle Dreyfus lui raconte qu'elle l'a croisé un jour sur les Champs-Élysées: elle fait allusion au fait qu'il avait Gros-Câlin autour du cou, ce jour-là, ce qui explique peut-être pourquoi elle ne s'intéresse pas à lui.

Gros-Câlin est surtout un récit humoristique. Lorsque Michel Cousin adopte le python, il a du mal à le nourrir. Il se rend vite compte du fait qu'un python comme Gros-Câlin se nourrit exclusivement de proies vivantes telles que "des souris, des cochons d'Inde, ou même un petit lapin de temps en temps" (12). Ayant acheté la souris Blondine pour l'offrir en pâture à Gros-Câlin, Cousin s'attache à elle, comme il s'était attaché à Gros-Câlin. Il est alors difficile de nourrir Gros-Câlin.

Une autre fois, la femme de ménage portugaise s'évanouit en voyant Gros-Câlin pour la première fois. Courant au commissariat de police pour porter plainte contre Cousin, dans un français approximatif, elle l'accuse d'être un "monsieur sadista, monsieur exhibitionnista" (37), ce qui provoque une scène d'hilarité auprès des policiers, qui croient que Cousin a commis un outrage à la pudeur.

À la fin du roman, Michel Cousin qui se sent victime de la vie moderne, commence à se reconnaître en son python, tandis que Mlle Dreyfus finit prostituée et rejoint le groupe de “bonnes putes” que fréquente Cousin.

3.4. *La Vie devant soi* / Ajar (1975)

Ce roman raconte l'histoire d'un petit garçon algérien, Mohammed, dit Momo. Il a été recueilli dès son plus jeune âge par madame Rosa, une vieille juive polonaise dans un foyer qu'elle tient clandestinement à rue Bisson à Belleville,. C'est le jeune Momo qui fait le récit de sa vie, aussi bien que de la vie de ceux qui l'entourent.

Dans ce foyer, il y a des enfants qui arrivent le matin et repartent le soir, tels que les deux enfants de Monsieur Moussa l'éboueur, déposés à six heures du matin. Madame Rosa héberge aussi chez elle plusieurs autres enfants, surtout des enfants de prostituées. Il y a Banania qui est “un Malien ou un Sénégalais ou un Guinéen ou autre chose” (21), Moïse, l'enfant juif illégitime qui fait honte à sa mère, Michel le petit orphelin vietnamien, Salima que sa mère a réussi à sauver de l'assistance publique en la cachant dans une poubelle lorsque les voisins ont dénoncé cette mère “comme pute” (67), Antoine âgé de deux ans, le seul français de souche, aussi bien que les enfants d'autres prostituées. Ces enfants, nés de pères inconnus, ont trouvé le jour par hasard, leur mère n'ayant pas pu se faire avorter à temps. Ces mères envoient des chèques alimentaires de façon irrégulière à Madame Rosa, et viennent leur rendre visite quand elles le peuvent. Momo

est le seul dans ce foyer à ne pas recevoir la visite de sa mère. D'ailleurs il ne l'a jamais vue.

Au début du roman, Momo ignore tout sur sa lignée généalogique. Il ne sait pas qui sont ses vrais parents, mais une chose toutefois est sûre, il doit avoir des parents quelque part car là-dessus "la nature est intraitable" (83). Il a appris à l'école qu'il est arabe, ce qui signifie qu'il doit être musulman, mais il ne pratique pas sa religion et ne connaît même pas son âge; il a soit dix ans, soit quatorze ans. Pendant plusieurs années son chèque mensuel de trois cents francs est toujours arrivé sans faute; mais ils ont cessé brusquement d'arriver il y a trois ans.

La plupart des autres locataires de l'immeuble sont des Noirs, telles que la famille Sarakollé et la famille Toutcouleurs. Il y a aussi Madame Lola, l'androgynisme du quatrième étage qui vient du Sénégal et exerce son métier au Bois de Boulogne, les membres de la famille Zoum qui vivent plusieurs dans une chambre, Monsieur Waloumba "qui a laissé toutes ses femmes en Afrique et [qui est venu] en France pour la balayer" (170) et monsieur N'Da Amédée qui vient du Niger et qui est "le plus grand maquereau de tous les Noirs de Paris [...] portant costume et bagues diamantaires aux doigts" (45). Le reste du boulevard de Belleville est habité par des Juifs et des Arabes (12).

Momo n'a jamais appris à parler correctement le français; il emploie ainsi un langage approximatif qui est mêlé de l'argot des bas-fonds de Paris, mais qui néanmoins

paraît naturel dans la bouche d'un petit enfant des rues. Par exemple parlant de monsieur N'Da Amédée il dit proxynète au lieu de proxénète.

Si Momo est le narrateur du roman et lui en donne le ton, Madame Rosa en est le personnage principal. Tout le roman est construit autour d'elle, autour de ses hantises, de ses expériences et de sa maladie. Durant sa jeunesse, Madame Rosa s'est prostituée pour survivre; elle a aussi connu Auschwitz. Elle souffre d'une sclérose cérébrale qui est en train de la tuer lentement. D'après elle, son plus grand malheur résulte du fait qu'elle a survécu aux chambres à gaz d'Auschwitz (119) ce qui la force maintenant à vivre dans les conditions dégradantes de la vieillesse et de la maladie.

Madame Rosa ne veut surtout pas finir ses jours dans un hôpital et s'abandonner aux recherches des médecins. Elle croit que les médecins, comme les pires nazis, la forceraient à vivre en légume: "ils ne veulent pas nous donner le droit de mourir" (182). Elle est convaincue que son malheur découle du fait qu'elle est née juive, que les Juifs sont nés pour souffrir. C'est pour cela qu'elle garde toujours un portrait d'Hitler sous son lit, pour s'en souvenir. Momo fait l'expérience de la vie à travers les expériences de Madame Rosa.

Un jour Momo apprend la véritable histoire de sa vie à lui, une histoire que Madame Rosa a toujours voulu lui cacher de peur de traumatiser cet enfant sensible. Youssef Kadir, un ancien proxénète, sort d'un asile d'aliénés où il avait été interné et vient voir Momo chez Madame Rosa. Momo apprend alors que ce Youssef Kadir est

son père et que lors d'une crise de jalousie, il a assassiné Aïcha la mère de Momo parce qu'elle "se faisait jusqu'à vingt passes par jour" (186) comme prostituée. Il apprend aussi que c'est un oncle d'Aïcha qui jusqu'à sa mort envoyait des chèques pour sa pension. Après l'aveu de Monsieur Kadir, Mohammed se rend compte du fait qu'il ne lui reste personne d'autre au monde à aimer que Madame Rosa.

Le lecteur suit petit à petit l'affection grandissante de Momo pour Madame Rosa, la seule personne qui lui tienne lieu de famille et qui continue de s'occuper de lui, même lorsque ses chèques n'arrivent plus. Momo qui a toute sa vie devant lui voit mourir Madame Rosa lentement. Momo aide Madame Rosa à mourir loin de tout hôpital dans son "trou juif". Après sa mort, Momo va vivre avec Mademoiselle Nadine, une jeune femme du quartier qui voudra peut-être l'adopter.

Dans ce roman, c'est Momo qui évoque la vie quotidienne des marginaux, des déshérités et des épaves de la société des grandes villes. Le roman présente aussi des problèmes d'actualité tels que la ségrégation des pauvres (souteneurs, androgynes, prostituées de la Goutte d'or), des malades et des vieux (Madame Rosa), des immigrants déshérités (Arabes, Juifs polonais, Africains sans papiers, Vietnamiens), qui, privés de droits juridiques et en l'absence de la protection de la police doivent se débrouiller tout seuls.

L'amour et la tendresse fleurissent malgré tout. Dans ce monde brillent des êtres au coeur d'or: Madame Lola s'occupe de Madame Rosa quand celle-ci est atteinte de

gâtisme; l'aîné de la famille Zoum apporte de la nourriture aux enfants lorsque Madame Rosa ne peut plus s'occuper d'eux (170) et Monsieur Waloumba vient "cracher du feu" (171) devant Madame Rosa afin de lui remonter le moral. C'est un monde où tous ces marginaux sont solidaires.

Les autres romans qui sont inclus dans ce travail sont *L'Étranger* (Camus), *Le Noeud de Vipères* (Mauriac), *La Porte étroite* (Gide) et *L'Immoraliste* (Gide). Pour un aperçu de l'action de ces romans, je renvoie au *Dictionnaire des oeuvres du XXe siècle; littérature française et francophone* de Henri Mitterrand.

3.5. Le choix des textes

Les textes mentionnés ci-dessus ont été choisis pour plusieurs raisons. J'ai tenu, dans un premier temps, à me concentrer sur des textes qui appartiennent au même genre, à savoir le genre romanesque. J'ai aussi choisi les romans de Gary publiés durant la période Ajar qui n'a duré que cinq ans, cela me permet de saisir la langue de Gary à un moment précis de son évolution.

La carrière littéraire de Romain Gary débute en 1945 avec *l'Éducation Européenne* et se termine avec *Les Cerfs-Volants* en 1981. Les stylisticiens s'accordent pour dire que l'écriture d'un auteur évolue au cours d'une longue carrière littéraire, ce qui implique que le style de Gary a pu se modifier avec le passage du temps. Les quatre romans de Gary/Ajar qui figurent dans cette étude ont été publiés entre

1974 et 1977, ce qui écarte la possibilité que le style de Gary ou d'Ajar ait pu subir de grandes modifications durant une si courte période.

Ce qui a aussi influencé le choix de ces textes, c'est que ces quatre romans sont presque tous de longueur semblable. Ceci veut dire que toute différence stylistique qu'on pourrait observer éventuellement ne découle pas du fait que le nombre des vocables n'est pas le même dans ces textes. La dernière raison pour le choix de ces romans est qu'ils sont présentés selon la perspective d'un narrateur *homodiégétique*, c'est-à-dire que dans chacun d'eux le narrateur est présent dans l'histoire qu'il raconte.

Le prochain chapitre examine le subterfuge Ajar en s'attardant sur les éléments les plus importants de cette affaire.

Chapitre 4

L'affaire Émile Ajar ou "Je est un autre"

Accéder à l'authenticité en partant d'une imposture, avouez que ce serait assez beau. (*Tête Coupable* 56)

Au cours d'une conversation entre Romain Gary et Dominique Desanti durant les années soixante-dix, Gary mentionne Elsa Triolet que tous les deux connaissent très bien. Après avoir reçu le Prix Goncourt pour le roman *Le premier accroc coûte 200 francs* en 1944, Triolet s'est butée contre le mutisme des milieux non-communistes en France. Elle a confié à Gary qu'elle a alors voulu faire publier son prochain roman *L'Inspecteur des ruines* sous le nom du narrateur Antonin Blond, et qu'elle souhaitait recevoir un deuxième Goncourt pour ce roman. Ne pouvant compter sur l'appui de quiconque, elle a été obligée au dernier moment de publier le roman sous son propre nom (Desanti 95).

Desanti postule que c'est probablement à l'époque où a lieu la conversation entre elle et Romain Gary que ce dernier songe sérieusement à se métamorphoser en Émile Ajar, car pendant cette conversation Romain Gary lui tient les propos suivants:

Imagine quelqu'un qui en aurait assez, comme moi et d'autres, de cette mesquinerie des critiques qui étiquettent les auteurs une fois pour toutes [...]. Imagine qu'un écrivain crée une écriture qui ne ressemble plus à celle qu'on lui attribue [...]. Mais qu'il aille jusqu'au bout [...].

Attention! L'écueil, ce serait Frankenstein. (Desanti 96-9)

Ce chapitre se propose d'élaborer les diverses tentatives de Romain Gary de se camoufler sous de nombreux pseudonymes. Dans un premier temps il décrit ses évasions avant la période Ajar. Ensuite, il élabore les principales étapes de l'aventure Ajar elle-même, en commençant par les événements entourant la rédaction et la publication de *Gros-Câlin* et, il termine par l'annonce officielle que Gary est Ajar.

4.1. La Raison profonde du faux

Le pseudonyme Ajar est loin d'être la première tentative de Romain Gary de fuir ses racines parentales. Dans *La Promesse de l'aube* il raconte que lorsqu'il était adolescent et habitait avec sa mère à l'hôtel Pensions-Mermonts à Nice, celle-ci l'encourageait à s'entraîner quotidiennement à son futur métier d'écrivain. Pour faire plaisir à sa mère, il rédigeait alors des histoires dans un cahier, sans les signer de son

propre nom; il s'inventait plusieurs pseudonymes tels que François Mermonts et Lucien Brûlard (174).

Plusieurs années plus tard, alors que Romain Gary faisait ses études de droit à Paris, il informe sa mère (qui habite à Nice) qu'il a signé des contrats d'écriture avec plusieurs revues parisiennes tel que *Gringoire*. Il lui explique néanmoins qu'étant donné que les directeurs de ces journaux réclament des nouvelles d'une qualité si bassement commerciale, il n'ose les signer de son propre nom de peur de compromettre sa réputation littéraire. Il préfère signer ces articles de pseudonymes divers. Quelle que soit la vérité sur cette prétendue collaboration entre Gary et les directeurs de journaux, elle montre que Romain Gary songe à se camoufler derrière un pseudonyme depuis sa jeunesse.

Sa première vraie tentative de s'évader de lui-même ne débute qu'en 1958 lorsque Gary publie *L'Homme à la colombe* sous le pseudonyme de Fosco Sinibaldi. Gary est alors diplomate respecté à Los Angeles, aussi bien qu'un auteur de renom en France. *L'Homme à la colombe* raconte l'histoire de Johnny, un cow-boy texan de vingt-cinq ans qui s'est battu en Corée, et qui après les deux guerres mondiales, a cru fermement en la Charte des Nations-Unies, et à un monde où régnerait la colombe de la paix. Mais il s'est vite rendu compte de son erreur. Maintenant pour se venger de ses illusions perdues, il s'habille en Gandhi et passe ses jours à filer de la laine dans une chambre secrète au quartier général des Nations-Unies, accompagné d'une

colombe. Il prétend faire la grève au nom de la paix et devient l'idole des Américains qui le prennent pour un saint. Vers la fin du roman, il est pris à son propre piège et se laisse mourir de faim avec sa colombe.

L'Homme à la colombe fait une satire féroce de l'O.N.U., la qualifiant de "gigantesque moulin à vent" (22). Cette organisation ressemble à un énorme gratte-ciel où le droit des peuples opprimés se perd dans de grands bureaux, dans des ascenseurs, dans d'innombrables propositions et dans des discours. Gary éprouve ce sentiment envers l'O.N.U tout au long de sa vie, car, plusieurs années plus tard il maintient que "L'O.N.U. c'est le viol permanent d'un grand rêve humain" (*Nuit 54*). On peut comprendre son désir de dissimuler son identité par devoir de réserve lorsqu'il fait publier *L'Homme à la colombe*.

En 1974 Romain Gary publie *Les Têtes de Stéphanie* sous son deuxième pseudonyme de Shatan Bogat. Pour ajouter foi à la vie de Shatan Bogat, Romain Gary lui invente la biographie suivante. Shatan Bogat, âgé de trente-trois ans, est le fils d'un immigré turc. Il a servi pendant quatre ans dans la marine américaine, a pratiqué le métier de journaliste et s'est finalement installé aux Indes. De là, il dirige une compagnie de pêche et de transport maritime dans l'Océan Indien et le Golfe Persique.

L'action du roman se déroule dans le Golfe Persique. L'héroïne du roman, Stéphanie est un mannequin international qui arrive en Perse pour faire des photos de

mode. En même temps des têtes coupées s'accumulent autour d'elle. Alors que *L'Homme à la colombe* est un roman cynique et une attaque implicite des Nations Unies, *Les Têtes de Stéphanie* est un roman d'espionnage pastiche, du genre des romans de Ian Fleming; les exploits de Stéphanie rappellent surtout ceux de James Bond.

Les Têtes de Stéphanie est évidemment un roman sans aucune "prétention littéraire" (Catonné 182), mais il permet à Romain Gary de s'échapper publiquement à lui-même et d'oublier temporairement sa réputation de diplomate et d'auteur respectable. Le livre connaît si peu de succès que Robert Gallimard, dans une émission de *France-Culture* doit expliquer que Shatan Bogat est le pseudonyme d'un auteur connu. Il y laisse assez d'indices pour que les lecteurs puissent reconnaître Romain Gary. C'est à partir de ce moment-là que la vente du livre démarre. Cette anecdote montre que parfois le succès d'un livre dépend non seulement de la qualité intrinsèque de l'oeuvre mais aussi de la réputation de l'auteur.

Quelque temps après la déclaration de Robert Gallimard, Romain Gary explique son désir de se cacher sous le pseudonyme de Bogat de la façon suivante: "J'éprouve parfois le besoin de changer d'identité, de me séparer un peu de moi-même, l'espace d'un livre (*Stéphanie* 1).

Ce besoin de se séparer de lui-même ne surprend pas, car Romain Gary soutient que tout romancier est avant tout un auteur-acteur ou un illusionniste de soi:

“lorsque j’entreprends un roman [...] c’est pour me quitter, pour me réinventer” (*Nuit* 258). Dans un autre roman de Gary, *Les Enchanteurs* (1973) le narrateur, Fosco Zaga déclare: “Il m’arrive de croire que j’invente, que je m’invente, [...]. Je n’existe, ami lecteur, que pour ta délectation et tout le reste n’est que tricherie” (294). On pourrait croire qu’il s’agit là d’un aveu caché de Gary qu’il triche parfois avec ses lecteurs.

C’est durant cette période de sa vie (en 1973) que Romain Gary cherche de nouveau à se réinventer et à trouver une autre identité qui le délivre du carcan où les critiques l’ont enfermé. Il considère alors que tous les moyens sont bons pour effectuer cette métamorphose: il n’hésite pas à tromper ses lecteurs si cela s’avère nécessaire. À peine une année après la déclaration de Fosco Zaga, Romain Gary prend le nom de plume Émile Ajar dans le plus grand secret. Paul Pavlowitch postule qu’il prend ce nom parce qu’Émile est le nom du fils naturel de Paul Gauguin (Pavlowitch 46).

Quoique cette explication semble un peu exagérée, elle contient quand même une part de vérité, car Gary, conscient du fait qu’il est bâtard, est souvent à la recherche d’une identité qui l’authentifie. Le pseudonyme d’Ajar est peut-être sa tentative de recommencer, de faire peau neuve et de vivre à nouveau à l’âge de soixante ans. Il explique à Paul Pavlowitch que cette réincarnation lui donnera peut-être la fraîcheur d’un auteur en début de carrière. Il lui dit aussi que les critiques l’ont marqué, classé, jugé et condamné depuis longtemps et que s’il publie un autre roman

sous son propre nom, ils vont “descendre le livre immédiatement” (Pavlowitch 35). Étant donné qu’il s’est démasqué sous l’imposture de Shatan Bogat, il est légitime de postuler qu’il pense que cela le protégera de l’imposture Ajar.

L’aventure Ajar se produit à une époque de tourments, d’angoisses et de contradictions dans la vie de Romain Gary. Jean Seberg et lui viennent de divorcer; il atteint la soixantaine qu’il redoute tant; souffre d’hypocondrie et n’est plus qu’un fonctionnaire retraité du Quai D’Orsay. Lorsqu’il avait rédigé les deux romans sous le pseudonyme de Shatan Bogat et de Fosco Sinibaldi, il s’était confié à son éditeur de la maison Gallimard; lorsqu’il écrit sous le nom d’Émile Ajar, il le fait sans la complicité ni l’autorisation de son éditeur.

4.2. Pierre Michaud et *La Solitude du python*

Lorsque Romain Gary rédige des romans, il les dicte habituellement à sa secrétaire Martine Carré qui tape les manuscrits à la machine. De mars à décembre 1973, il écrit de sa propre main un roman qui s’intitule *La Solitude du python*. Il le donne ensuite à sa secrétaire pour qu’elle le tape à la machine. Par mesure de précaution et pour “preuve de l’authenticité de sa création” (Bona 319), il recopie ce même manuscrit minutieusement de sa main sur un cahier noir de comptable.

Ayant terminé le manuscrit de *La Solitude du python*, Gary l’envoie à son ami Pierre Michaud (un industriel nouvellement établi au Brésil) et le charge de l’envoyer

de Rio de Janeiro à la maison Gallimard. Michaud doit affirmer que ce roman a été rédigé par un certain Émile Ajar, dont lui, Michaud est l'intermédiaire. Sous les directives de Gary, Michaud spécifie aussi au comité de lecture chez Gallimard qu'Émile Ajar est né à Oran en Algérie française durant les années trente et a même connu Albert Camus.

Pour ajouter foi à cette histoire, Gary invente une histoire selon laquelle Émile Ajar était un étudiant en médecine en France et qu'il a pratiqué un avortement clandestin sur une jeune femme qui est morte par la suite. Il a dû s'enfuir à Rio de Janeiro. Comme il est maintenant un repris de justice, il a intérêt à garder l'anonymat le plus complet. Romain Gary charge Michaud de dire au *Mercur* de France que lui, Michaud, a fait la connaissance d'Émile Ajar, une espèce de vagabond dénué de tout scrupule "un salaud, un cochon [...] un type pas fréquentable" (Pavlowitch 54) qui demeure toujours un avorteur clandestin et un souteneur à Rio de Janeiro et qui vit dans un état d'indigence assez avancé. Étant donné son passé douteux, il est impossible de faire sa connaissance.

Lorsque le manuscrit de *La Solitude du python* est distribué aux lecteurs de la maison Gallimard, quelques-uns d'entre eux jugent que le roman est "trop fou" (Bona 322) pour être publié, tandis que d'autres lecteurs jugent qu'il leur a procuré "un plaisir violent et composite" (Bona 322). Raymond Queneau décide quand même de ne pas le faire publier, soupçonnant qu'il s'agit d'un canular: "L'auteur de *Pierrot*, de

Zazie, ne pouvait qu'identifier dans Ajar un mauvais frère, un artiste aguerri en matière de fantaisie et de burlesque, un routier de longue date lui aussi" (Bona 322).

Claude Gallimard, le directeur de la maison d'édition, reconnaissant le talent de ce nouvel auteur, envoie le manuscrit au *Mercure de France*, alors dirigé par sa femme Simone. Les éditeurs du *Mercure de France*, y compris Michel Cournot, sont immédiatement séduits par le ton et le langage du roman, et décident de le faire publier. Ils réclament néanmoins que les cinquante dernières pages du roman soient rayées du manuscrit, et, que le roman soit intitulé *Gros-Câlin*. Lorsque le roman paraît au mois de septembre 1974, il attire l'attention du public parisien et le séduit par l'originalité de son sujet.

Ajar est proclamé comme l'une des plus grandes découvertes de l'année. Les critiques et les lecteurs admirent surtout le style du roman, les jeux verbaux et le narrateur qui s'exprime dans un langage curieux:

Je dois donc m'excuser de certaines mutilations, mal emplois, sauts de carpe, entorses, refus d'obéissance, crabismes, strabismes et immigrations sauvages du langage, syntaxe et vocabulaire [...], je tiens donc à donner au langage employé dans le présent traitement une certaine indépendance et une chance de se composer autrement que chez les usagés. (9-10)

Les jugements critiques sur *Gros-Câlin* sont très favorables. Jacqueline Piatier affirme dans *Le Monde* qu'il s'agit d'une "satire de l'homme du XXe siècle" (Bona 325). Dans *Le Parisien libéré*, Christine Anorthy parle de "nouvelle psycholittérature" et déclare: "Ajar c'est Gogol, Ajar, c'est le Pouchkine des ténèbres de Paris [...]" (Bona 87). Dans *L'Express*, Matthieu Galley loue le roman en ces termes:

Le langage façon Queneau y est neuf et impertinent autant que le raisonnement, farfelu jusqu'à la poésie canularsque, mais rigoureux à sa manière et d'une naïveté à laquelle il ne faut pas trop se fier [...].

L'auteur est [...] un mandarin déguisé en Douanier Rousseau. (Bona 88)

Dans les salles de rédaction, tous cherchent à deviner qui se cache derrière Ajar. Quelques critiques soupçonnent qu'Émile Ajar est le pseudonyme d'un auteur célèbre. *Le Nouvel Observateur* désigne Louis Aragon ou Raymond Queneau comme Ajar. Quelques années après la parution de *Gros-Câlin*, Émile Ajar fait lui-même allusion à Queneau dans un autre roman: "Mon cul, dis-je en pensant à *Zazie dans le métro*, de Raymond Queneau" (*Pseudo* 33). D'autres soupçonnent Michel Cournot, le directeur littéraire du Mercure de France, d'être le vrai Ajar (Bona 326).

Quelques journalistes fouillent dans les archives de la police pour découvrir l'identité d'un médecin qui a pratiqué des avortements criminels. Ils découvrent un certain Hamil Raja, terroriste libanais et concluent qu'Ajar est l'anagramme de Raja.

Ce faisant ils authentifient l'existence de ce Hamil Raja/Émile Ajar. Personne ne fait de rapport entre Émile Ajar et Romain Gary. Seulement quelques proches amis de Gary (son avocat, sa secrétaire, Jean Seberg, Michaud et Robert Gallimard) connaissent la véritable identité d'Ajar. Aucun critique ne devine encore qu'Ajar est une autre invention de Romain Gary.

Plusieurs mois après la publication de *Gros-Câlin*, les prix littéraires sont accordés aux meilleurs romans de l'année. En France, le Prix Théophraste Renaudot est l'un des plus prestigieux pour la littérature: il est accordé le même jour que le Prix Goncourt. Cette année-là, le jury du prix Renaudot veut l'attribuer à *Gros-Câlin*. Étant donné que cet honneur forcerait Romain Gary à se démasquer, il envoie deux lettres signées Émile Ajar aux membres du jury Renaudot dans lesquelles il refuse le prix d'avance. Le jury, ahuri et blessé dans son orgueil, accorde le prix à un autre candidat, Georges Bourgeaud, pour son roman *Voyage à l'étranger*. Cette même année (1974) le Prix Goncourt est attribué à Pascal Lainé pour son roman *La Dentellière*.

4.3. *La Tendresse des pierres*

Lors de sa rencontre avec Dominique Desanti, Romain Gary lui avait indiqué que si jamais il publiait un roman sous un pseudonyme, il créerait sa propre Galatée mais ne la rendrait pas autonome. Son stratagème consisterait à "prendre un solitaire,

qui ait horreur des mondanités, vive hors Paris, se laisse difficilement approcher (Desanti 97).

Au printemps de 1975, Romain Gary termine le manuscrit d'un nouveau roman *La Tendresse des pierres* qu'il veut à nouveau faire publier sous le nom d'Émile Ajar. Il demande alors à Paul Pavlowitch d'incarner le rôle d'Émile Ajar. À la manière de Pygmalion, il crée sa propre Galatée en transformant Paul Pavlowitch en Émile Ajar. Il n'aurait pu trouver mieux pour participer dans une telle machination. D'après Bona, Paul Pavlowitch, avec son air de hippie et de gauchiste, a horreur des bourgeois; il a aussi une silhouette de marginal et de gitan, un style à la fois vulgaire et poétique. Avec ses cheveux noirs et ses yeux noirs "il a une tête de l'est, russe ou yougoslave, et le même faciès de métèque [...] il a [...] la dégaine du hors-la-loi" (335).

De plus, les liens de parenté entre Romain Gary et Paul Pavlowitch sont très étroits. Paul Pavlowitch est le fils de Dinah, la cousine germaine de Romain Gary qui figure dans *La Promesse de l'aube*; il est ainsi le petit-cousin de Gary et non pas son neveu ou plutôt il est un neveu à la mode de Bretagne. Pavlowitch a trente-deux ans, étant né le cinq février 1942 à Nice. Romain Gary lui a servi de père et de protecteur et a même payé ses études, dont six mois à Harvard.

Paul Pavlowitch habite dans le même immeuble que Romain Gary, au 108 rue du Bac, avec sa femme Annie et sa petite fille, dans deux chambres de bonne

converties en appartement. Il est sans métier fixe mais intellectuel à sa manière, il adore la littérature: "C'est un poète sans oeuvre" (Bona 335). D'ailleurs, il admire beaucoup son oncle. Gary demande à Pavlowitch de prendre contact avec le Mercure de France par téléphone et d'indiquer qu'il est Émile Ajar et qu'il leur téléphone de Rio de Janeiro afin de négocier un nouveau contrat avec eux.

Lors de ce premier contact entre Pavlowitch et Michel Cournot, ce dernier est convaincu qu'il parle à Émile Ajar. Paul Pavlowitch consent à rencontrer Michel Cournot et Simone Gallimard en Suisse et fixe un rendez-vous avec eux à l'Hôtel des Bergues à Genève. Mais, pris de panique au dernier moment, Pavlowitch ne va pas au rendez-vous. Ceci intensifie la rumeur selon laquelle Ajar fuit la foule et la publicité.

Quinze jours plus tard (mai 1975), Pavlowitch reprend contact avec le Mercure de France et consent de nouveau à rencontrer Michel Cournot dans un studio que possède Romain Gary dans un faubourg de Genève. Cette fois-ci, il est au rendez-vous comme prévu. Lorsque Michel Cournot le rencontre pour la première fois, il n'a aucune peine à croire qu'il parle au fameux Émile Ajar:

Je n'ai pas douté une seule seconde que c'était "lui" [...]. Cet homme ne pouvait être qu'Émile Ajar. Il avait [...] la tête de l'emploi. C'était lui, tel que je l'avais exactement imaginé [...]. À la fois bandit et fou, bizarre, inquiétant [...] cet homme avait le physique, la voix et le style d'Émile Ajar. (Bona 337)

Michel Cournot ajoute: "chaque fois Ajar avait une réponse prête comme un vrai écrivain. Il semblait connaître son texte sur le bout des doigts [...]. Il m'a paru très aguerri, très 'pro' " (Bona 338). Au cours de cette première rencontre Michel Cournot demande à Paul Pavlowitch de faire plusieurs retouches au manuscrit.

En septembre 1975, Paul Pavlowitch rencontre Simone Gallimard à Copenhague. Simone Gallimard est convaincue elle aussi qu'elle est en face du célèbre Ajar: elle lui lance un retentissant "Bonjour Émile Ajar" (Bona 340). Elle passe deux jours en sa compagnie à Copenhague. Toujours à Copenhague Paul Pavlowitch accorde une longue interview à Yvonne Baby, une journaliste du *Monde*, et lui déclare dans un langage qui sied à Ajar "Les mots, c'est fait pour ça. C'est une urgence pour moi, je ne cherche pas à démontrer le sens" (Pavlowitch 138).

Loin de Romain Gary, son mentor, qui rédige tous ses scénarios, Paul Pavlowitch improvise: il mélange sa vie et celle de Romain Gary, livre des demi-vérités et des détails superflus sur la vie d'Ajar. Il indique que, lui Ajar/Raja est non seulement né en Pologne mais aussi en Russie. Il précise que Wilno est sa ville d'origine et non pas le Liban (où Hamil Raja est censé être né); il parle de son enfance à Nice, de sa mère et de sa grand-mère et fait mention de sa double ascendance juive et slave. Or Hamil Raja est censé être libanais. Pour Yvonne Baby qui ignore les liens de parenté entre Paul Pavlowitch et Romain Gary, il n'y a plus de doute: cet homme qui confond et mélange tout, conformément à sa réputation, c'est bien cet

Hamil Raja, le célèbre Émile Ajar. Lorsque *Le Monde* publie l'entrevue entre Yvonne Baby et Émile Ajar le 10 octobre 1975, cet article prouve qu'Émile Ajar existe.

Si Paul Pavlowitch joue impeccablement le rôle d'Émile Ajar, il est souvent saisi d'angoisse et la comédie commence à être insoutenable (Pavlowitch 117). Romain Gary demande à Paul Pavlowitch d'envoyer sa photographie au *Mercure de France* et de la signer "Émile Ajar". Pavlowitch envoie donc une photo très floue où on peut à peine discerner ses traits. Pour la sortie du livre qui porte alors le titre *La Vie devant soi*, Pavlowitch livre aussi aux médias la notice biographique fictive suivante:

Émile Ajar a trente-cinq ans. Dans sa première enfance (il était fils unique), il a habité Nice, avec sa maman qui était Yougoslave et qui a fait de la Résistance. Ils sont allés ensuite à Paris où Ajar a fait ses études secondaires, puis quatre ans d'étude de médecine. Pendant sa médecine, il a été moniteur dans un centre pour prédélinquants. Il interrompt ses études en 1964 et part pour l'Espagne. Il séjourne ensuite au Maroc.

Il rentre en France, prend un travail d'assistant monteur de cinéma. Il repart cette fois pour les Antilles (Saint-Martin), d'où il passe aux États-Unis, puis au Brésil où il a vécu ces dernières années. (Catonné 105)

À la parution de *La Vie devant soi*, Paul Pavlowitch devient Émile Ajar aux yeux de tous. Il lui prête son corps, son visage, son passé et son allure. Pour qu'il n'y ait jamais de soupçons sur la véritable identité de cet Hamil Raja/Émile Ajar, Romain Gary donne de faux papiers à Pavlowitch, (y compris un faux permis de conduire pour Hamil Raja) aussi bien que des papiers d'assurance de Lloyds:

À l'aide de *lettraset* et d'une photo d'identité, je fis des transformations [...]. J'avais maintenant un permis de conduire et des papiers d'assurance au nom de Hamil Raja, domicilié Rue de la Cygne, Genève. (Pavlowitch 87)

Le stratagème de Romain Gary réussit si bien que les médias sont convaincus qu'Émile Ajar existe réellement. Lorsque *La Vie devant soi* est publiée en 1975, il est vendu à plus d'un million d'exemplaires, et la critique parisienne le considère comme un chef-d'oeuvre. Le langage argotique et imagé qui reflète le langage du français immigré et de l'enfant de la rue séduit tout le monde. Max-Pol Fouchet compare ce roman au *Voyage au bout de la Nuit* de Céline, et, Yvan Androuard, dans un article paru dans *Le Canard Enchaîné* déclare qu'il s'agit là "du meilleur roman de la saison" (Bona 343).

Entre-temps Paul Pavlowitch joue admirablement le rôle que Romain Gary lui a attribué et, malgré les moments d'angoisse, il signe de nombreuses dédicaces pour ce roman. Malgré le succès de *La Vie devant soi* auprès des critiques et du public, ni

Romain Gary ni Pavlowitch ne pense qu'il sera nommé pour le Prix Goncourt; ils songent plutôt au Prix Médicis qui est moins prestigieux. Quelques critiques commencent à avoir des doutes sur l'identité de ce Hamil Raja. Pour comble de malheur un éditorialiste du journal télévisé *Le Point*, Pierre Billard, rajoute aux noms des mystificateurs qui pourraient se cacher sous le nom d'Ajar, celui de Romain Gary.

Étant donné le succès littéraire grandissant de *La Vie devant soi* Romain Gary vit dans l'inquiétude que le public apprenne la véritable identité d'Ajar. Au lieu de se réjouir du succès de *La Vie devant soi* Romain Gary plonge plutôt dans le plus grand désarroi:

[Lorsqu'il] devint Ajar, Romain commençait à avoir de sérieux doutes sur le sort que la postérité lui réserverait. Il s'en souciait, mais la postérité ne lui semblait qu'une création démesurément surévaluée d'une époque qui cherche à se prolonger. (Pavlowitch 160)

Romain Gary décide que si jamais son roman reçoit le Prix Goncourt, il le refuserait, comme l'avait fait par principe, Julien Gracq pour son roman *Le Rivage des Syrtes* en 1951. Le 17 novembre 1975, on annonce à la radio que le Prix Goncourt a été décerné à Émile Ajar pour son roman *La Vie devant soi*. Romain Gary devient ainsi le premier auteur à obtenir deux fois le Prix Goncourt. S'il se réjouit de cet honneur, il est angoissé en même temps car il a une réputation impeccable de gaulliste et de vieux combattant qu'il ne veut compromettre à aucun prix.

Ajar, quant à lui, demeure introuvable. Romain Gary demande alors à Gisèle Halimi, le conseiller juridique d'Ajar de déclarer au jury Goncourt qu'Ajar refuse le prix et qu'il fuit la publicité. Mais Hervé Bazin de l'Académie Goncourt riposte que ce prix est décerné à un livre et non à un candidat et "ne peut ni s'accepter ni se refuser, pas plus que la naissance ou la mort" (Catonné 106).

Après s'être vu décerner son deuxième Goncourt, Romain Gary vit dans la hantise d'être découvert; il menace même de se suicider. Saisi de panique, il est persuadé que les journalistes sont à ses trousses. Il téléphone à Annie, la femme de Paul Pavlowitch, et dans un langage incohérent et décousu, il lui fait la déclaration suivante:

Mais tu ne comprends pas [...] ils sont sur le palier [...] ils attendent que je me montre [...]. -- Voilà [...] j'ai mon revolver [...] c'est foutu [...] je vais en finir. [...] ils sont de plus en plus nombreux. [...] C'est un scandale terrible [...] il n'y a plus rien à faire [...]. Romain au téléphone, son revolver en main, prêt à disparaître au milieu d'un vaudeville [...]. Pour lui c'était une faillite totale [...]. Il garda toute la soirée le trente-huit en main. (Pavlowitch 168)

Ce dont Romain Gary a surtout peur, c'est d'être humilié et dégradé en public dans la Cour des Invalides et d'avoir les épaulettes et la rosette de la légion d'honneur arrachées.

4.4. Paul Pavlowitch

Dans l'interview que Paul Pavlowitch avait accordée à Yvonne Baby, il avait indiqué qu'il avait fait ses études de médecine à Toulouse. Jacques Bouzerand, un journaliste du journal télévisé *Le Point*, décide d'aller à la recherche d'Émile Ajar. Faisant une enquête auprès d'anciens étudiants toulousains et leur montrant la photo floue de Pavlowitch (qui est censée être celle d'Émile Ajar/Hamil Raja), il apprend que c'est une photo de Paul Pavlowitch.

Il en déduit que derrière ce célèbre Émile Ajar se cache Paul Pavlowitch et non pas Hamil Raja. Lorsqu'il apprend par ailleurs que Pavlowitch habite alors à Caniac-de-Causse, près de Cahors il prend contact avec lui. Pavlowitch accepte de lui accorder un interview, mais demande en échange que son identité ne soit pas révélée au public. Tout en lui consacrant l'article s'intitulant "*Énigme: Qui est Ajar?*" dans *Le Point*, Bouzerand cache néanmoins son identité. Cependant le 21 novembre 1975, un autre journaliste de *La Dépêche du Midi*, qui voit lui aussi que la photo envoyée au *Mercure de France* est celle de Paul Pavlowitch, révèle qu'Émile Ajar est le pseudonyme de Paul Pavlowitch et non celui de Hamil Raja.

Les journalistes se mettent alors à la recherche de Paul Pavlowitch étant convaincus qu'il est le véritable Ajar. Lorsque, par inadvertance, ils découvrent que ce dernier est le neveu de Romain Gary, ils assiègent l'appartement de Romain Gary la rue du Bac dans l'espoir de tomber sur Paul Pavlowitch. Celui-ci, suivant le conseil

de Romain Gary, dans un geste qui sied à son image d'Ajar, lance des coups de poing aux journalistes: "L'un des photographes dansa près de moi. Je lui lançai un crochet malhabile tandis que l'autre prenait photo sur photo" (Pavlowitch 175).

Dès lors, Romain Gary doit faire taire la rumeur qu'il a aidé son neveu à écrire ses romans. Il plaide auprès de la presse de ne pas harceler ce neveu hystérique et hypersensible. Lorsque quelques critiques soupçonnent qu'Émile Ajar est le pseudonyme de Romain Gary, ce dernier donne sa parole d'honneur à Claude Gallimard qu'il n'est pas Ajar. Il affirme aussi, dans une entrevue accordée à la journaliste Jacqueline Piatier, que "Je ne suis pas Émile Ajar et je n'ai collaboré en aucune manière aux ouvrages de cet auteur" (Bona 350). Avec condescendance il déclare qu'il a pu influencer son neveu et que même si ce neveu est coupable du plagiat, on devrait l'excuser étant donné que c'est "un jeune auteur, je n'ai pas l'intention de protester [...]. Il est normal qu'un écrivain de ma stature influence les jeunes" (Bona 347).

Suivant le conseil de Romain Gary, Paul Pavlowitch envoie une photo plus nette à la presse et accorde des interviews aux quatre coins de la France. Tout le monde est maintenant convaincu qu'Émile Ajar est le pseudonyme de Paul Pavlowitch. Pour tous, Paul Pavlowitch (Émile Ajar) est le souffle neuf et puissant de la jeunesse. Il est un meilleur écrivain que Romain Gary.

Pendant toute la période Ajar, Romain Gary ne néglige pas pour autant de faire publier d'autres romans sous son propre nom: *Au-delà de cette limite votre ticket n'est plus valable* (1975), *Clair de femme* (1977), *Les Clowns lyriques* (1979) et *Les Cerfs-Volants* (1980). Il rédige aussi en 1977 *Charge d'Âme*, qui est la traduction française de son roman anglais *The Gasp*, et en 1979, la pièce de théâtre *La Bonne moitié* une adaptation de son roman *Le Grand Vestiaire*.

4.5. Un Extraordinaire Ballet de masques: *Pseudo*

En 1965 Romain Gary avait élaboré une théorie du roman où il avait indiqué qu'un roman "devrait même, de préférence, avoir plusieurs 'vérités' intérieures, contradictoires, en conflit" (*Sganarelle* 467). En 1976 Romain Gary met en pratique cette théorie du roman. Obéissant à "un puissant besoin psychologique" (Pavlowitch 178), Romain Gary compose en quinze jours, dans son studio à Genève, le premier jet du manuscrit *Pseudo* (*Vie et mort* 21). Ce roman passe pour être la confession d'Émile Ajar/Paul Pavlowitch, dans lequel Paul Pavlowitch explique son angoisse existentielle qui découle du fait que la rumeur attribue la paternité de ses oeuvres à Gary. Ce dernier, oncle abusif veut traîtreusement lui voler son oeuvre, son Goncourt (*Pseudo* 191).

Gary ne fait rien pour démentir les bruits qui courent qu'Émile Ajar est Romain Gary lui-même. Pavlowitch, un personnage hystérique, veut faire taire les

soupçons qu'Émile Ajar est en réalité Romain Gary. Ainsi, Pavlowitch annonce-t-il à tout le monde:

Je suis Émile Ajar! [...] Le seul, l'unique! Je suis le fils de mes propres oeuvres et le père des mêmes! Je suis mon propre fils et mon propre père. [...] Je suis authentique. Je ne suis pas un canular! Je ne suis pas pseudo-pseudo. (192)

Dans cette invective lancée à Gary par Pavlowitch, c'est comme si Pavlowitch voulait détruire une fois pour toutes la rumeur que son oncle Gary aurait pu enfanter Ajar. Romain Gary apparaît dans ce roman comme un personnage ridicule qui persécute son neveu, tout en étant risible à cause des nombreux cigares qu'il fume. Il apparaît sous les traits caricaturaux de Tonton Macoute "un salaud [...] vicelard comme pas un. Se taper une sainte, ça devait être le rêve de sa vie" (10-29). Cet être orgueilleux et cynique ne s'intéresse à rien, sinon à la peur de grossir. C'est peut-être même le père de Paul Pavlowitch étant donné qu'il a sans doute couché avec Dinah, sa mère.

Dans ce roman qui est écrit à la première personne du singulier, "je" joue un triple rôle, se référant à Paul Pavlowitch, à Émile Ajar et à Romain Gary. C'est une "superposition d'identités, du créateur à la créature, en passant par le personnage-acteur" (Bona 358).

Lorsque Romain Gary termine la rédaction de *Pseudo*, il demande à Paul Pavlowitch de taper le manuscrit à la machine au fur et à mesure qu'il le réécrit. Ainsi Paul Pavlowitch peut-il lire son autoportrait fictif où il est présenté comme un être dépressif et drogué, un fou à lier qui se fait soigner dans une clinique psychiatrique à Copenhague. Paul Pavlowitch n'a jamais été atteint de troubles mentaux et ne s'est jamais fait soigner dans un asile d'aliénés. Son oncle, remuant toujours ses ficelles, lui demande de ne pas censurer le manuscrit même si le contenu de *Pseudo* est blessant et repoussant.

Paul Pavlowitch en veut à Romain Gary d'avoir mis toute sa famille dans le texte, en commençant par Dinah, sa mère qui selon Romain Gary "est morte malhonnêtement, d'une lente sclérose cérébrale" (144), ce qui explique la folie d'Ajar, pour finir par son grand-père Ilya Ossipovich que Gary prétend s'être ruiné au jeu. Il n'épargne pas, non plus, les tantes de Paul Pavlowitch. Le soi-disant Ajar/Pavlowitch raconte que l'une d'elles, à dix-sept ans a contracté la syphilis la nuit de ses noces, et que l'autre s'est fait violer par un cosaque étant donné que "Les Russes ont le goût du drame" (*Pseudo* 143).

En lisant les demi-vérités que Gary écrit sur la famille de Paul Pavlowitch, ce dernier ne peut maîtriser sa haine pour Gary. Les conflits entre eux commencent à partir de ce moment. L'une des fonctions de Pavlowitch comme secrétaire de Gary est de taper à la machine le manuscrit de *Pseudo*. C'est une tâche très lourde à exécuter,

d'autant plus que l'écriture de Romain Gary est illisible: "Toutes les lettres se ressemblaient. Les *t* n'avaient pas de barre, les *f* semblaient des *b*, les *a*, les *o* les *g*, les *l* étaient écrits de la même manière chaotique et inachevée" (Pavlowitch 190).

De plus en plus excédé, nerveux et angoissé, Paul Pavlowitch commence à boire et à prendre des amphétamines. Il ne se rase plus et néglige tellement sa toilette qu'il commence à ressembler de plus en plus au modèle Émile Ajar, tel que Romain Gary l'avait conçu. Furieux contre son oncle, il tape "sur la vieille Olympia comme si c'était sur sa figure" (Pavlowitch 197).

Pour se venger de Romain Gary Paul Pavlowitch commence à lui jouer des tours. Travaillant dans un appartement qui avoisine celui de Gary, il sonne à la porte de celui-ci, retourne furtivement à son bureau et s'assoit en vitesse comme s'il n'en avait pas bougé. Romain Gary ouvrant la porte n'y trouve personne et, ahuri, retourne dans son appartement. Ce manège se répète plusieurs fois jusqu'au point où Gary croit qu'il devient fou. Il demande à Paul Pavlowitch:

- Tu as entendu la sonnette? (...)

- Non. Je n'ai rien entendu. J'étais là, j'aurais dû, je t'ai bien entendu rentrer. Qu'est-ce qui se passe? Il me regarda. J'avais réussi.

Maintenant il doutait de lui. Il avait la trouille. (Pavlowitch 220)

Lorsque Michel Cournot reçoit le manuscrit de *Pseudo* au mois de janvier 1976, il a peine à croire ce qu'il lit, mais accepte l'authenticité de Paul Pavlowitch.

Jugeant que *Pseudo* serait une entrave à la carrière littéraire d'Ajar, il suggère à Pavlowitch de retarder la publication de ce texte en attendant de faire publier un autre vrai roman digne du succès de *La Vie devant soi*. Lorsque Romain Gary entend les propos de Michel Cournot, il refuse de rédiger un autre roman sous le pseudonyme d'Ajar avant que *Pseudo* ne soit publié.

Simone Gallimard est convaincue, elle aussi, que *Pseudo* calomnie Romain Gary et s'inquiète de l'image de Romain Gary dans ce roman. Elle demande à ce dernier de lire le manuscrit afin d'y apporter les coupures qui s'imposent avant sa publication. Mais Gary lui envoie une lettre qui dégage Simone Gallimard et le Mercure de France de toute responsabilité concernant ce roman, et renonce aussi à toute poursuite judiciaire. Étant donné les conditions établies dans cette lettre, le Mercure de France accepte de publier *Pseudo* en novembre 1976.

À sa parution le roman est mal reçu par la critique, mais il établit une fois pour toutes que Paul Pavlowitch est bien Émile Ajar, quoiqu'il soit dépeint comme un Ajar instable et peu sympathique. Après avoir lu *Pseudo*, aucun lecteur n'aurait soupçonné que l'auteur Ajar qui s'acharne tant à ruiner la réputation de Romain Gary est Romain Gary lui-même. Entre-temps Paul Pavlowitch signe des contrats importants, en particulier ceux des droits cinématographiques. Les romans d'Ajar, qui sont presque tous des succès de librairie intéressent les réalisateurs de cinéma. Les divergences

entre Romain Gary et Paul Pavlowitch s'accroissent à tel point qu'une rupture s'annonce d'un moment à l'autre (Pavlowitch 215).

Si Romain Gary écrit *Pseudo* sous la peur d'être découvert (ce qui compromettrait sa réputation de gaulliste et de Chevalier de la Légion d'honneur) la publication de ce roman convainc la critique que Paul Pavlowitch est l'authentique Émile Ajar. Même si Pavlowitch est représenté comme un être fou, délirant et peu sympathique, il est représenté comme Ajar tout de même. La presse ne met plus en doute l'authenticité d'Émile Ajar et au mois de décembre 1976, *Les Nouvelles littéraires* font un nouveau reportage sur Paul Pavlowitch (l'authentique Ajar) et mettent sa photo sur la couverture de la revue.

Un jour après la sortie du livre *Pseudo*, Romain Gary se trouve dans un bar avec quelques connaissances. Ils commencent à parler de *Pseudo* et Gary joue à l'indigné. Il exprime sa rancune envers Paul Pavlowitch et verse son venin contre ce neveu ingrat qui l'a tant calomnié dans *Pseudo*. Jouant la comédie jusqu'au bout, il laisse même couler des larmes en murmurant: "Me traiter de tonton-macoute, moi! moi! Dire que je le maintenais dans les mansardes, alors que je le suppliais d'habiter dans l'appartement!" (Desanti 98). Après une telle performance personne ne peut soupçonner qu'Émile Ajar est Romain Gary lui-même.

Comme le fait remarquer Paul Pavlowitch, les romans qui suivent *Pseudo*, qu'ils soient signés Ajar ou Gary, laissent tous des traces d'Ajar: "Le style du livre

subit nécessairement l'intoxication Ajar [...]. Romain invente des justifications scéniques afin de faire passer son trop-plein d'ajarismes" (250). Entre-temps Ralph Menheim, le traducteur de Brecht, de Céline et de Grass entreprend la traduction anglaise de *La Vie devant soi*. Romain Gary, qui s'y connaît en argot américain, fait des retouches à la traduction de Menheim par l'intermédiaire de Paul Pavlowitch.

Ralph Menheim a traduit la dernière phrase du livre *Il faut aimer* par *It takes love*. Étant convaincu que cette traduction de Menheim est un contresens, Gary exige qu'elle soit retraduite par *It takes loving* ou mieux, *It takes lovin*. Il n'accepte pas non plus la traduction de *trou juif* par *jewish holes*, mais propose *Black holes as we call them in the neighbourhood*. En 1977, la version cinématographique de *La Vie devant soi* sort aux États-Unis et connaît un grand succès. L'actrice française Simone Signoret interprète le rôle de Madame Rosa.

4.6. Revendication de la vieillesse et des racines juives

En janvier 1979 paraît le dernier roman signé Émile Ajar, *L'Angoisse du roi Salomon* qui d'après Catonné est "le chef-d'oeuvre d'Émile Ajar" (115). Ce roman raconte l'histoire d'un ancien roi du prêt-à porter, Salomon Rubinstein. C'est un vieillard de quatre-vingt-quatre ans, qui, à l'encontre de Romain Gary son auteur, refuse de se laisser abattre par son âge. En effet, il est plus alerte qu'un jeune homme.

M. Salomon ne craint pas la mort mais la vieillesse; il ne croit pas non plus à sa fin prochaine car, à quatre-vingt-quatre ans, il consulte une voyante pour voir ce que l'avenir lui réserve; de plus, le matin de son quatre-vingt-cinquième anniversaire il demande à Jeannot, le jeune chauffeur de taxi: "Emmenez-moi chez les putes" (315). Pour lutter contre l'angoisse de la mort et calmer d'autres vieillards angoissés et désespérés, il fond un groupe de jeunes bénévoles, sous le nom de SOS amitié.

Un jour, se liant d'amitié avec Jeannot, M. Rubinstein l'engage pour aller livrer des cadeaux aux gens malades ou à ceux qui vivent seuls. Jeannot est envoyé à Nice pour livrer des fruits à Mademoiselle Cora Lamenaire, une ancienne chanteuse de cabaret. Cora Lamenaire, âgée de soixante-trois ou soixante-quatre ans, comme vestige sans doute de son passé théâtral, met "trop de produits sur son visage (116). Celle-ci s'attachant à Jeannot veut le lancer dans le spectacle. Ce n'est qu'après bien des péripéties que le jeune Jeannot comprend la véritable histoire de M. Salomon Rubinstein et de Mademoiselle Cora Lamenaire. Ils se sont aimés autrefois, mais durant l'occupation allemande, il a dû se cacher dans une cave près des Champs-Élysées. Elle ne l'a pas dénoncé aux Allemands mais l'a quand même abandonné dans son trou.

Ce dernier roman signé Ajar, plus que tout autre roman de Romain Gary peut-être, valorise les racines juives de son auteur. Comme le remarque Bona ces racines, jadis enfouies remontent de plus en plus à la surface à mesure que vieillit Romain

Gary: "Parmi tous ses portraits romanesques, le Juif aura toujours représenté l'homme qui souffre et qui a souffert" (387). Romain Gary n'écrit pas un livre sur l'holocauste, mais traite la souffrance des Juifs avec humour, l'humour de M. Salomon.

Étant donné les liens de parenté entre Pavlowitch (le célèbre Ajar) et Romain Gary, Claude Gallimard recommande à Simone Gallimard d'engager Paul Pavlowitch comme conseiller littéraire au Mercure de France. Elle le met sous la tutelle de Paul Léautaud, contre la volonté de Gary: "Cela ne lui plaisait pas du tout, mais il n'y pouvait rien", avoue Pavlowitch (302). Comme il s'acquitte mal de cette fonction, Simone Gallimard s'inquiète sur ses capacités d'éditeur et le congédie après une année et demie. Elle lui suggère de se consacrer à sa vocation d'écrivain. Romain Gary qui est de moins en moins maître de la situation Ajar, vit dans une grande frayeur que, par inadvertance, Paul Pavlowitch dévoile son secret au public.

Pour comble, Paul Pavlowitch réclame des droits à des pourcentages plus élevés que ceux que lui accorde Romain Gary. Il déclare plus tard que "l'argent devait servir de faux prétexte" (295) à leur rupture. D'après Pavlowitch, Romain Gary lui a envoyé une lettre sèche dans laquelle il a réclamé "l'état des comptes" en la signant "j'attends". Le gouffre entre eux s'approfondit et désormais la rupture est inévitable:

Nous étions chacun sur nos positions, sauf que je ne savais pas quelles étaient les miennes. Il ne m'impressionnait plus autant. Son valet encore, de moins en moins son vassal. (Pavlowitch 286)

En mars 1979 Romain Gary rédige *Vie et mort d'Émile Ajar*. Le 8 septembre 1979 la police découvre le corps de Jean Seberg. Romain Gary accuse le F.B.I. d'avoir causé sa mort par ses calomnies. À la fin du mois de mars 1980 Romain Gary a aussi des soucis auprès du contrôle fiscal qui l'interroge sur les droits de ses romans à l'étranger. Il ne s'agit que d'une bagatelle étant donné que "le contrôle a d'ailleurs été réglé sans trop de difficulté" (Bona 393), mais cette affaire qui se produit peu de temps après le suicide de Jean Seberg prend des proportions énormes aux yeux de Romain Gary. Il est fort probable que pour la première fois Gary commence aussi à avoir des doutes sur Émile Ajar, pseudonyme qui a si bien réussi. Le 2 décembre 1980, il se suicide.

4.7. Retentissement dans le milieu littéraire

Lorsque le testament littéraire de Romain Gary *Vie et mort d'Émile Ajar* est publié six mois après son suicide, il établit une fois pour toutes la parenté de l'oeuvre signée Ajar, et Gary rejoint toute une légion de mystificateurs. Les critiques restent incrédules et acceptent mal cette double mystification.

Dans son testament littéraire qui est comme une sorte de confession, Romain Gary explique les raisons profondes qui l'ont poussé à entreprendre cette mystification: "J'étais las de n'être que moi-même. J'étais las de l'image Romain Gary qu'on m'avait collée sur le dos" (*Vie et mort* 28). Il explique, par ailleurs, qu'il voulait à tout prix s'éloigner des idées préconçues que la critique lui réservait et cherchait à se libérer de "la gueule qu'on m'avait faite" (*Vie et mort* 23). Avec Émile Ajar, dit-il: "Je recommençais. Tout m'était donné encore une fois. J'avais l'illusion parfaite d'une nouvelle création de moi-même par moi-même" (*Vie et mort* 30).

Lors de la parution des *Racines du ciel* en 1956, plusieurs erreurs typographiques avaient paru dans les cinq cents exemplaires imprimés lors de la première édition de l'oeuvre. Elles avaient été corrigées dans les réimpressions, mais cet incident fâcheux a nui à la réputation littéraire de Gary. Les critiques trouvaient que le style du roman était touffu, qu'il y avait trop de lourdeurs et de répétitions. Ils accusaient Gary de ne pas maîtriser la langue française. Le bruit courait aussi que la maison Gallimard avait engagé Albert Camus et Jacques Lemarchand pour retoucher le texte. Jean Giono s'était mêlé à cette affaire et avait déclaré avec mépris: "Dommage que Gary n'écrive pas toujours en français" (Catonné 56).

D'autres critiques accusaient Gary de plagiat; une rumeur courait que Romain Gary avait trouvé le manuscrit du roman *Éducation européenne* sur le corps d'un aviateur polonais de la RAF et s'en était accaparé.

Ceci montre que plusieurs critiques ont cherché à dénigrer la valeur littéraire de l'oeuvre de Gary. En 1977 Poirot-Delpêche établit une sorte de bilan de la littérature française contemporaine dans la revue *Le Monde*; Romain Gary n'y figure pas mais Ajar est cité deux fois, comme "sang neuf" (Pavlowitch 286). Cette anecdote confirme l'opinion de Gary que le monde littéraire parisien et les critiques l'enferment dans un carcan. Elle aide à comprendre pourquoi Romain Gary n'a jamais avoué être Émile Ajar de son vivant, et pourquoi il l'a si catégoriquement répudié.

Chapitre 5

Romain Gary et Émile Ajar: Ressemblances

[Les gens] sont tout autre chose. Mais personne ne s'en doute, ils simulent, ils font pseudo-pseudo huit heures par jour et on leur fout la paix (*Pseudo* 21).

Plusieurs années après l'aveu posthume de Gary qu'Émile Ajar est bien son pseudonyme à lui et non pas le pseudonyme de Paul Pavlowitch, les nombreuses similarités entre l'oeuvre signée Gary et celle signée Ajar sautent aux yeux, car Gary a laissé des traces indéniables de son style dans les romans publiés sous le nom d'Ajar. Comme le remarque Catonné, Ajar "avec son ironie introvertie, son parler décalé, sa philosophie parodique" (144) est partout présent dans l'oeuvre de Gary.

Dans le *Pseudo* d'Ajar, le narrateur Pavlowitch/Ajar fait une allusion à peine voilée à sa véritable identité. Atteint de schizophrénie, il se fait traiter dans une clinique à Copenhague; dans un moment de complicité avec ses lecteurs, il leur fait la

confession suivante: “Le Docteur Christianssen [...] savait que je simulais [...] car la seule chose que je ne voulais pas être, c’était moi-même” (*Pseudo* 174). On peut alors postuler qu’il s’agit là d’un aveu indirect de l’auteur Gary, qu’Ajar (le soi-disant pseudonyme de Pavlowitch) n’est pas ce qu’il semble être.

Ce chapitre se propose d’examiner les ressemblances entre l’oeuvre de Romain Gary et celle d’Émile Ajar. Dans un premier temps, il souligne les similarités entre le langage de Gary et d’Ajar. Ensuite il élabore les thèmes, les personnages et les motifs qui sont communs aux oeuvres de Gary et d’Ajar.

5.1. Les Ressemblances linguistiques

Ajar et Gary sont tous deux des noms d’emprunt de Roman Kacew. Dominique Bona explique qu’Ajar signifie “braise” en russe tandis que “Gary” signifie “brûle” à l’impératif en russe aussi (317). Le chapitre 2 a montré que Gary, séparé dès son plus jeune âge de la Russie, éprouve une grande nostalgie pour cette patrie perdue. Par ailleurs, d’après Pavlowitch, pour Gary “La Russie [...] est une terre romantique et populaire, c’est avec Pouchkine la terre de tous. [...] Quand Romain parlait de Pouchkine, les larmes lui venait aux yeux” (45).

Dans *La Promesse de L'aube* Gary affirme que sa mère lui faisait apprendre à parler et à écrire le russe durant son enfance à Wilno et à Varsovie. Il explique aussi, que sa mère et les autres Russes expatriés à Nice, parlaient en russe entre eux,

(*Promesse* 78). Lui-même, à Nice, parlait souvent le russe avec ses voisins le colonel Aprelev, ancien officier de l'armée de Tzar, et avec la Princesse Cantacuzène (Bona 13). La langue russe fait partie intégrante de sa culture, et explique en partie le choix de ses deux noms russes.

Soit par manque d'attention, soit par désinvolture, Romain Gary reprend dans ses romans les mêmes noms de rues, les mêmes phrases et les mêmes titres, qu'ils soient signés Gary ou Ajar. Dans *Gros-Câlin*, par exemple, le narrateur fait mention d'un bordel qu'il fréquente situé dans la rue "Roy-le-Bien". Le nom de cette rue avait déjà paru quelques années auparavant dans un autre roman publié sous le nom de Romain Gary.

Par ailleurs, lorsque Romain Gary termine le manuscrit de son deuxième roman signé Ajar, il le nomme par inadvertance, *La Tendresse des pierres*. Ceci est le titre d'un roman fictif écrit par l'un des personnages dans le roman *Adieu Gary Cooper* que Gary publie sous son propre nom en 1969. Lorsque la femme de Pavlowitch signale cette erreur à Gary, il change immédiatement le titre du roman en *La Vie devant soi* (Catonné 78). Gary ne peut pas toujours se défaire de certains aspects de son écriture lorsqu'il assume le nom d'Émile Ajar.

Gary utilise un style différent, nouveau, familier et humoristique dans les romans signés Ajar; ce style se manifeste déjà dans le premier roman *Gros-Câlin*,

comme on le constate dans le passage suivant, lorsque Cousin explique au commissaire de police comment un garçon de bureau l'a insulté:

Le garçon de bureau [...] m'a dit que [...] je devais ramper hors de mon trou et me dérouler librement au soleil sur toute ma longueur [...].

Je fais des noeuds tout le temps à cause de ma démarche intellectuelle, je colle simplement à mon sujet. (41-3).

Cousin poursuit son discours de digression en digression au point d'imiter la manoeuvre du serpent. Ce style nouveau est en rupture avec les oeuvres passées: il représente une espèce de "désarticulation" (Bayard 65). Le ton du langage est différent de celui de Gary et correspond, peut être, au côté gitan et bohémien de la personnalité de Gary. Par contre, le style des romans signés Gary évoque l'homme respectable, le diplomate et le Chevalier de la Légion d'honneur.

5.2. Thèmes et leitmotive communs aux oeuvres de Gary et d'Ajar

Dans l'oeuvre de Gary on retrouve souvent des thèmes, des revendications, des obsessions et des personnages semblables. Le même cynisme, les mêmes idées philosophiques et politiques et la même vision pessimiste et ironique du monde caractérisent les textes Gary et les textes Ajar.

L'un des thèmes les plus importants est la vieillesse. Ce thème se manifeste d'une manière prédominante dans *La Vie devant soi* (Ajar), *Au-delà de cette limite*

votre ticket n'est plus valable (Gary) et *L'Angoisse du roi Salomon* (Ajar). Pour les personnages principaux dans ces romans, la vieillesse est à redouter.

Dans *Au-delà de cette limite votre ticket n'est plus valable* le thème de la vieillesse est abordé par le sujet délicat de l'impuissance sexuelle des sexagénaires. Le narrateur du roman, approchant la soixantaine, associe son manque de vitalité sexuelle à la vieillesse qui, d'après lui, préfigure la mort. Cette hantise de vieillir risque de détruire sa liaison avec Laura; dans son angoisse obsédante, il s'imagine que les hommes plus jeunes et plus virils que lui se moquent de lui à cause de sa liaison avec une si jeune femme.

À l'époque de la publication du roman, les critiques ont cru lire l'aveu indirect de Gary (qui approchait la soixantaine) de ses propres défaillances sexuelles. Avec le recul, il est légitime de lier ce thème à l'impuissance de l'auteur Gary devant son alter-ego Ajar: "sous le masque de la faiblesse virile, il nous parle d'une impuissance beaucoup plus grave, celle qui menace le créateur au soir de son oeuvre" (Catonné 88). Dans une telle perspective, le personnage de Ruiz que crée Jacques Rainier ne symbolise pas la virilité du jeune mâle, mais constitue plutôt l'aveu que Gary "a besoin d'un autre nom qui le fasse changer d'identité, comme Rainier a besoin de Ruiz [...]. *Au-delà de cette limite* nous donne alors une des clés de l'enfantement d'Ajar par Romain Gary" (Catonné 188).

Le thème du vieillissement est repris dans *La Vie devant soi* avec le personnage de Madame Rosa. Comme dans *Au-delà de cette limite votre ticket n'est plus valable*, la vieillesse est à craindre car elle entraîne la déchéance physique, la maladie, le gâtisme, et annonce la mort. Madame Rosa, ayant dépassé la soixantaine, est condamnée à vivre dans des conditions qui la font associer sa dégradation physique à un prolongement d'Auschwitz.

Ce thème figure aussi dans *L'Angoisse du Roi Salomon* signé Ajar. Dans ce roman, la vieillesse et non la mort est à redouter lorsqu'on a la vie derrière soi. La vieillesse offre la réalité du changement et d'être "quelqu'un d'autre alors qu'on est toujours le même" (74). Contrairement à Mme Rosa et à Jacques Rainier qui se sentent accablés par la vieillesse, Monsieur Salomon, personnage principal du roman, octogénaire, qui a échappé aux chambres à gaz des Nazis lutte contre la mort: "Je n'ai pas échappé à l'holocauste pour rien, mes petits amis. J'ai l'intention de vivre vieux, qu'on se le tienne pour dit" (156). À l'encontre de Mme Rosa, M. Salomon est décidé à ne pas mourir du tout: "Il respirait l'élégance de la dernière heure et on sentait tout de suite que ce n'était pas un homme à se laisser mourir facilement" (10). Ce qui sauve ce personnage de l'angoisse de la mort, c'est son optimisme, sa foi humaniste et son amour du prochain.

Un autre thème commun aux romans de Gary et d'Ajar est celui de la maladie. Il est très évident dans *La Vie devant soi*. Madame Rosa souffre de gâtisme et de

sénilité: elle est atteinte d'une sclérose cérébrale dont elle finit par mourir. Elle demande à Momo de lui assurer la fin solitaire qu'elle a choisi, loin de tout hôpital, en la laissant mourir dans son "trou juif".

Le thème de la maladie figure aussi dans *La Promesse de l'aube* (Gary); Dinah, la cousine du narrateur et la mère de Paul Pavlowitch est elle aussi atteinte de gâtisme et de sénilité, qui la tuent à la fin. Dans ce même roman, Nina Kacew, la mère du narrateur est gravement malade: elle souffre du diabète et en meurt à la fin du roman.

Le thème de la maladie est aussi important dans l'oeuvre signée Gary que dans celle signée Ajar. Quoiqu'il existe une différence de ton au niveau du style entre les deux, il y a néanmoins des échos de détails.

5.3. Dépassement de soi et dédoublement

Un autre thème commun aux oeuvres de Gary et d'Ajar est celui du dépassement de soi qui trouve sans doute ses racines dans la vie de Gary elle-même. Il peut être retracé dans les expériences de Roman Kacew à Wilno où, enfant, il apprend une leçon qui le marquera pour le reste de sa vie. Il raconte dans *La Promesse de l'aube*, comment âgé de huit ans et pauvre, il rencontre le fils de l'épicier qui, riche, se gave de pastèque et n'en offre point à Romain Gary. Cette "pastèque" est restée tout au long de sa vie le symbole de l'inaccessible:

je sentirai toujours la morsure du regret dans mon coeur et [...] toutes les pastèques du monde ne me feront pas oublier celles que je n'ai pas mangées, à huit ans, lorsque j'en avais le plus envie [...] la pastèque absolue continuera à me narguer jusqu'à la fin de mes jours, toujours présente, pressentie, et toujours hors de portée. (114)

C'est à partir de ce moment décisif de sa vie qu'il commence sa quête de se surpasser. Il s'entraîne à jongler avec huit oranges au lieu de sept mais n'arrive jamais à dépasser la septième: "Le chef d'oeuvre demeurerait inaccessible, éternellement latent, éternellement pressenti, mais toujours hors de portée. [...] la dernière balle est restée à jamais hors d'atteinte (*Promesse* 187).

Gary exploite le thème du dépassement de soi dans plusieurs romans. Dans *Les Mangeurs d'étoiles* (la version française de son roman *Talent Scout*), les personnages évoqués par le titre sont d'abord les Indiens pauvres d'Amérique latine qui se nourrissent de champignons hallucinatoires afin de dépasser leur existence. Il y a aussi les mangeurs d'illusions tels que Charlie Kuhn, le "talent scout" américain, dénicheur de vedettes et prospecteur de talents qui sont à la recherche d'actes de prouesse.

Kuhn arrive en Amérique latine à la recherche de numéros de cirque par les ventriloques, les contorsionnistes et les hypnotiseurs qui puissent faire croire à la transcendance et au dépassement de soi, ou à une façon d'échapper aux lois de la

nature. Il y rencontre un violoniste virtuose qui, dépassant Paganini, joue un concert en faisant le poirier; un surhomme sexuel capable de faire l'amour un nombre incalculable de fois, et un homme obsédé qui veut "[se] surpasser et jongler avec treize balles, au lieu de mes maudites douze balles habituelles" (34).

Dans un autre roman, *Les Têtes de Stéphanie*, roman d'espionnage que Gary publie en 1974 sous le pseudonyme de Shatan Bogat, le thème de la quête de l'impossible est symbolisé par un Gurkha, célèbre pour avoir décapité deux ennemis d'un seul coup de sabre. À la fin du roman, il atteint le sublime lorsqu'il fait rouler trois têtes d'un seul coup de sabre. Comme ce Gurkha, Gary veut lui aussi devenir un maître de l'impossible, trancher plusieurs têtes à la fois et jongler "avec sept balles au lieu de six" (*Promesse* 187). Il devient ce maître de l'impossible lorsqu'il crée son double Émile Ajar.

Dans plusieurs romans de Gary, le thème du dépassement ou du dédoublement est intimement lié à celui du *dibbouk* (à savoir de l'âme d'un défunt qui n'ayant pas trouvé de repos prend possession d'un vivant). Ce thème est souvent rattaché au thème d'un psychisme scindé en deux par la dichotomie entre l'imagination et la réalité.

5.4. Le Dibbouk

Le thème du dibbouk fait partie du folklore juif. Il paraît dans l'oeuvre dramatique *Le Dibbouk* que Chalom Anski publie en français en 1957. Dans cette pièce Hanan un jeune homme vivant dans un village polono-lithuanien, est amoureux de Léa, la fille d'un riche commerçant mais ne peut l'épouser, car Léa est fiancée à un autre jeune homme plus riche que lui. Lorsque Hanan a recours à des pratiques cabalistiques pour obtenir l'amour de Léa, il en meurt. Au moment où Léa est sur le point d'épouser son fiancé, elle se tord dans des convulsions et se met à parler d'une voix d'homme. Le dibbouk ou l'âme de Hanan a pris possession d'elle.

Romain Gary, qui est juif et qui connaît sans doute la culture juive, s'empare du thème du dibbouk et l'introduit dans plusieurs de ses romans. Dans *Europa* (1972) par exemple, l'un des personnages du roman, le docteur Jarde, dit qu'il faut "s'arranger avec les deux mondes, le *faux* qui existe, et le *vrai* qui n'existe que comme une déchéance [...]. D'une manière plus générale [...] la dichotomie, la schizophrénie [...] est l'état physiologique *normal* de notre cerveau, une réalité biologique" (361-2). À la fin du roman Danthès, ambassadeur de France à Rome finit dans la schizophrénie "ce qu'on appelle en anglais Split personality" (*Nuit* 72).

Ce thème figure de manière proéminente en 1967 dans le roman *La Danse de Gengis Cohn*, une farce surréaliste qui raconte l'histoire de Gengis Cohn (un alliage du Juif Moïché Cohn et le Mongol Gengis Khan) un homme à deux vies. Gengis

Cohn est un comique juif qui a travaillé dans les cabarets yiddish. Il a exercé son métier à Berlin, a vécu dans les ghettos de Varsovie et a finalement été déporté à Auschwitz. Au moment de son exécution par les Nazis, il a pris refuge dans la tête de Schatz, un Nazi. Après la guerre, Schatz est devenu un honnête citoyen de la République Fédérale Allemande, et le commissaire de police d'une bourgade. Gengis Cohn le hante et le persécute pendant plus de vingt ans. Maintenant victime et bourreau sont inséparables.

Gengis Cohn a plusieurs revendications; pour se venger de l'ancien Nazi Schatz, il lui fait manger des rahat-loukoums et lui fait respecter les fêtes juives:

L'autre jour[...] c'était la fête de *Hannukah* et Schatz, qui connaît nos fêtes sur le bout du doigt, m'avait cuisiné quelques-uns de mes plats *kosher* favoris. [...] Sa logeuse, Frau Müller, entra à ce moment-là et la vue de commissaire de police Schatz, offrant d'un air suppliant un plateau de *tcholnt* et de *gefillte fisch* à un Juif qui n'était pas là lui fit tellement peur qu'elle se trouva mal. (*Danse* 13)

Persécutant Schatz, Gengis Cohn exige qu'il parle en yiddish, une langue que Schatz ne connaît pas. Parfois il se manifeste en personne devant Schatz qui se met à lui parler. Comme personne autre que Schatz ne peut voir Cohn, tous les proches de Schatz sont convaincus qu'il est fou. Gengis Cohn ne sait plus si c'est lui le dibbouk

de Schatz ou si c'est ce dernier qui le hante, mais il a le sentiment que quelqu'un l'a créé:

une chose en tout cas est sûre; je ne suis pas chez moi là où je suis [...].

Je ne sais plus si je pense ou si je suis pensé, si je souffre ou si je suis souffert, si je hante ou si je suis hanté. Bref, je me sens *possédé*.

(*Danse* 168)

Pour embrouiller davantage les pistes, le personnage principal et narrateur du livre n'est pas Schatz mais Gengis Cohn, un fantôme, un personnage qui n'existe pas. Ce n'est qu'à la fin du roman que le lecteur se rend compte que Cohn aussi bien que Schatz ne sont que des personnages fictifs qui proviennent de l'inconscient de Gary lui-même qui lors d'une visite dans les ghettos à Varsovie, est pris d'évanouissement et d'hallucination. De ce point de vue "les deux personnages sont présentés comme des instances conflictuelles du psychisme de Gary, tout droit sorties, [...] d'un texte de Freud sur la topique de l'appareil psychique" (Bayard 88), car ce dernier n'est pas un véritable fantôme, mais la mauvaise conscience de Schatz.

Romain Gary présente ici sa propre version du dibbouk dans le personnage de Gengis Cohn: "[Il] participe d'une immense parabole comique où se lisent à la fois l'influence de l'humour hassidique et celle des conteurs russes à la Gogol" (Bayard 83).

Gengis Cohn, le dibbouk, réapparaît en 1968 dans un autre roman de Romain Gary, *La Tête coupable*. Cette fois-ci, Gengis Cohn se réincarne en Marc Matthieu, un savant réfugié à Tahiti. Se sentant responsable de toutes les horreurs que l'humanité s'inflige à elle-même, il veut "se fuir d'une identité à l'autre" (248). D'après Pierre Bayard (87) et Jean-Marie Catonné (169), Gengis Cohn évoque déjà en 1967 et 1968 le futur romancier Ajar qui lui aussi est fictif. Gengis Cohn ressemble à plusieurs personnages figurant dans les romans d'Ajar et de Gary (tel que le narrateur de *Pseudo*).

Le thème du dibbouk apparaît de manière moins manifeste dans *Les Cerfs-Volants* de Gary, publié en 1980. Ce roman raconte l'histoire du jeune orphelin Ludo, né en Normandie et élevé par son oncle Ambroise Fleury, "le facteur timbré". À l'âge de dix ans Ludo rencontre une petite polonaise blonde aux yeux bleus dans la forêt de Normandie. C'est Lila, la fille du comte Bronicka qui passe ses vacances avec sa famille dans son château de Normandie. À partir de ce moment, elle devient l'unique amour de Ludo. Mais la famille Bronicka regagne la Pologne peu après.

Quelques années plus tard, au moment des accords de Munich, Ludo est invité en Pologne par la famille Bronicka. Il essaie de se rendre digne de Lila, étudie et devient même le secrétaire du comte Bronicka. Il souffre aussi de jalousie et se bat en duel avec un beau soupirant de Lila, l'Allemand Hans von Schweden. Quelque temps

après, lorsque la Deuxième Guerre mondiale éclate, Ludo quitte la Pologne et Lila pour rentrer en France.

Ludo se joint à la Résistance en Normandie, et à travers toutes les épreuves de la guerre et de l'occupation allemande, il est soutenu par l'image des cerfs-volants de son oncle qui symbolisent l'aspiration à la liberté. Ludo possède aussi un don unique: celui de la mémoire historique qui est une tare de la famille Fleury. Sa mémoire lui permet de se rappeler des personnages et des événements bien longtemps après que toutes les autres personnes les aient oubliés. Il a aussi la capacité d'héberger en lui des gens disparues. Tout le monde affirme que Lila Bronicka est passée du côté des Allemands, mais Ludo, étant convaincu qu'il va la retrouver après la guerre, conserve en mémoire une Lila fidèle. Entre-temps il la fait vivre en lui:

pendant tout ce temps la présence de Lila à mes côtés prenait une telle réalité physique que mon oncle, ironie ou non, en était venu à mettre un troisième couvert sur la table pour celle qui n'était pas là d'une manière aussi présente. (97)

Le dédoublement qui survient chez Ludo provoque de l'inquiétude chez ses proches: ceux-ci craignent que les autorités, le croient fou et l'enferment parce qu'il parle à une femme absente. À cela il réplique simplement: "Eh bien, qu'on nous enferme. Elle restera avec moi" (157). La confusion entre présence et absence est

l'une des caractéristiques de la notion du dibbouk, c'est-à-dire qu'on est possédé soit par un être qui n'existe pas (Gengis Cohn), soit par un être qui est absent, comme Lila.

Dans le cas précédent, le thème du dibbouk se manifeste par le refus de la séparation. Ce thème se trouve aussi dans *Gros-Câlin* que Cousin présente comme un essai consacré aux reptiles. Dans ce roman Cousin donne son intention de "coller à son sujet" (17). Il s'exprime alors dans un langage curieux et loufoque qui contient beaucoup de glissements et d'imperfections:

Je précise immédiatement par souci de clarté que je ne fais pas de digressions [...]. Cette démarche ne s'effectue pas en ligne droite, mais par contorsions, sinuosités, spirales, enroulement et déroulement successifs, formant parfois des anneaux et de véritables noeuds. [...].

Il faut qu'il se sente chez lui dans ces pages. (*Câlin* 17)

Dans ce passage le pronom "il" se réfère à Gros-Câlin, le python. Mais il se réfère aussi à Cousin qui s'identifie à son serpent et se prend pour le reptile. Cousin s'exprime dans un langage qui imite la manoeuvre des serpents qui s'enroulent et se déroulent. Il se rapproche de son python mal-aimé et minoritaire, ressent "une espèce de mutualité" (11) avec lui, commence à ressembler à Gros-Câlin et à être en symbiose avec lui. Même s'il ne s'agit pas ici d'un cas typique de possession tel que l'on trouve dans les cas de dibbouk, il existe quand même le thème de l'indissolubilité de deux êtres.

Ce thème est repris dans *La Vie devant soi*, d'Ajar. Lorsque Madame Rosa meurt, Momo agit comme si elle était vivante quoique malade. Il s'allonge sur le matelas à côté d'elle et ne la quitte pas du tout, car "elle pouvait se réveiller et croire qu'elle était morte" (265). Pour qu'elle reste fraîche il parfume son corps d'eau de Cologne et la maquille: "Je pensais qu'elle pourrait vivre ainsi encore des années [...]. Je lui ai versé dessus tout le parfum qui restait [...] j'ai acheté des couleurs à peindre" (264-8). Momo refuse d'admettre que Mme Rosa soit morte.

À la fin du roman il déclare très simplement qu'il la fera vivre en lui: "Moi j'ai aimé Madame Rosa et je vais continuer à la voir [...]. Madame Nadine [...] m'a montré comment on peut faire reculer le monde" (269). Mme Rosa vivra toujours en lui, dans sa mémoire et deviendra ainsi son dibbouk.

Dans *La Promesse de l'aube* l'un des principaux thèmes est le refus de séparation entre Nina Kacew et son fils; on y trouve "un croisement d'écriture où la mère rédige un roman dont le fils est le héros" (Bayard 71). Effectivement le "moi" de Gary, dont avait rêvé sa mère Nina Kacew, se réalise après la mort de Nina lorsque Romain Gary deviendra écrivain de renom. Le "moi" de l'auteur de *La Promesse de l'aube* ne se réalise donc que sous la dictée de sa mère défunte. C'est comme si sa mère le possède au delà de la mort; elle agit toujours en lui et le guide plusieurs années après sa mort à elle, ce que Bayard appelle "une prise de personne" (55).

En fait, une fusion s'est faite entre la mère et le fils: la mère possède son fils au-delà de la mort et parle à travers lui. Romain Gary explique comment, lorsqu'il était loin de sa mère et faisait la guerre, il essayait de convaincre ses amis de partir avec lui pour l'Angleterre:

Je gesticulai, je mis une main sur le coeur, je brandis le poing, je pris un air inspiré. Je crois vraiment que c'était la voix de ma mère qui s'était emparée de la mienne [...]. Je crois même que ma voix changea et qu'un fort accent russe se fit entendre clairement alors que ma mère évoquait "la patrie immortelle". (*Promesse* 295)

Tous ces exemples montrent que le thème du dibbouk existe aussi bien dans les romans signés Gary que dans ceux signés Ajar.

5.5. Personnages communs aux oeuvres de Gary et d'Ajar

Dans l'oeuvre de Gary les personnages se ressemblent souvent, qu'ils apparaissent dans des romans publiés sous le nom de Gary ou d'Ajar. Morel, le héros des *Racines du ciel* est un stoïque qui refuse de se soumettre à la douleur ou de renoncer à ses idéaux quelles qu'en soient les conditions et quelles qu'en soient les pénalités. Ce leitmotiv du héros stoïque qui ne renonce jamais, apparaît aussi chez Ajar dans le personnage du vieux M. Salomon. Ce dernier refuse de se soumettre aux conditions dégradantes de la vieillesse.

Janek Twardowski de *l'Éducation européenne* (le premier roman de Gary) et Ludo des *Cerfs-Volants* (son dernier roman) se ressemblent aussi dans leur malheur. L'éducation européenne de Janek se fait à travers les pelotons d'exécution dont il est témoin, les violeurs de jeunes femmes et les bourreaux allemands qui l'entourent. Son éducation européenne, c'est aussi sa prise de conscience du fait que sa mère a été forcée à se prostituer par les Allemands, que ses frères et son père sont morts assassinés ou exécutés par les Nazis. L'éducation de Ludo se fait par ses expériences durant l'occupation allemande de la Normandie et par le truchement de son aspiration à la liberté.

Ce leitmotiv d'un jeune garçon victime des circonstances existe aussi dans les romans d'AJAR. Il y a le petit Banania dont le vrai nom est Touré et dont l'existence est dévalorisée par le fait qu'il est un "fils de pute", le petit Moïse, enfant illégitime, un "blond aux yeux bleus" qui fait honte à sa mère juive (*La Vie* 22). Mais il existe surtout Mohammed dit Momo, le petit algérien musulman dont la lignée généalogique laisse beaucoup à désirer. Tous ces jeunes gens ont en commun le fait qu'ils sont des frères du malheur. Ce qui unit aussi ces jeunes gens c'est l'absence du père.

La mère de ces personnages est souvent une prostituée, soit qu'elle ait été forcée de l'être comme la mère de Janek, soit qu'elle exerce ce métier pour subvenir aux besoins matériels de sa famille, comme la mère de Banania, de Moïse ou de Momo. L'existence de ces jeunes gens est rendue moins malheureuse par la présence

d'une femme tendre qui leur tient lieu de mère, telle que Mme Rosa (*La Vie*), d'une amante comme Zoia (*Éducation*) ou de Lila (*Cerfs-Volants*).

Un autre personnage qui apparaît souvent dans l'oeuvre de Gary et d'Ajar est la prostituée au coeur d'or, souvent une vieille femme accablée mais tenace. Ce personnage apparaît sous les traits de Nina Ivanovna du *Grand Vestiaire* (Gary) et de Mme Baumgartner de *Tulipe* (Gary) qui sont en quelque sorte l'archétype de Mme Rosa. Madame Rosa, ancienne prostituée, vieille femme au coeur d'or est elle aussi, accablée et tenace. Ces trois femmes ressemblent à une autre vieille femme, "Lili Marlène", la vieille maquerelle qui figure dans *Les Cerfs-Volants*. Mais il y a aussi Madame Lola, la travestie qui fournit paix et réconfort à Madame Rosa dans ses pires moments.

Il n'y a rien d'original dans le personnage de la prostituée au coeur d'or. Tout comme les personnages-types de la littérature (Arlequin, Pantalon, Pierrot, le valet complice des amours de son maître au dépens d'un jaloux, le voleur volé, le médecin malgré lui) ce personnage appartient à une longue tradition de personnages-types remontant à l'antiquité et au Moyen-Âge. Mais le fait demeure que la prostituée au coeur d'or apparaît dans les romans signés Gary aussi bien que dans les romans signés Ajar.

Un autre personnage-type qui figure dans l'oeuvre de Gary et d'Ajar est la jeune héroïne qui se sacrifie pour les autres. Zoia (*Éducation*), jeune patriote

polonaise de quinze ans, couche avec plusieurs soldats allemands afin d'apprendre s'ils transportent des explosifs dans leurs camions. Elle ressemble en cela à Minna (*Racines*), une ancienne prostituée allemande sans illusion, qui croit en la cause de Morel et le suit dans la brousse afin de sauver les éléphants.

Lila Bronicka (*Cerfs-Volants*) se prostitue par patriotisme afin de sauver des Nazis, sa famille et les autres patriotes polonais. Le personnage de la jeune prostituée figure aussi dans les romans d'Ajar. Mais à l'encontre de Minna et de Zozia, ces jeunes femmes se livrent à ce métier non pas par patriotisme mais parce qu'en se prostituant elles ont le sentiment d'aider à la survie de leurs proches. Ainsi Mlle Dreyfus n'est pas une patriote en temps de guerre mais une jeune femme des années soixante-dix qui, travaillant au bureau de statistique du Jardin des Plantes à Paris, se prostitue pendant ses heures de loisir parce qu'elle se sent valorisée:

Ici c'est peut-être pas considéré, mais c'est beaucoup plus vivant [...].

C'est plus social, il y a le contact humain, c'est plus personnel [...]. Je soulage [...] j'aide les gens à vivre [...] quand je suis payée je sais que j'ai de la valeur. (*Câlin* 214)

5.6. Les Marginaux

Il existe dans les oeuvres de Gary et d'Ajar l'être opprimé ou marginal qui ne peut jamais se faire entendre. Ce personnage est représenté par les éléphants des

Racines du ciel qu'au Chad tous, sauf Morel et ses alliés, veulent détruire au nom du progrès. Ces éléphants représentent tous les faibles qui sont écrasés par les systèmes totalitaires.

À bien des égards ces éléphants ressemblent à *Gros-Câlin* qui est lui aussi tyrannisé et marginal. Un jour, Cousin se promène dans les rues de Paris avec son python; Gros-Câlin est tellement repoussant et effrayant que la police doit s'en mêler, car il est "interdit de troubler l'ordre public en montrant des bêtes dites dangereuses dans les rues" (*Câlin* 56). Une autre fois, alors que Cousin est en train d'étreindre Gros-Câlin en nouant ses deux bras autour de lui sensuellement, il entend la concierge, Madame Niatte, derrière lui avec son seau et son balai qui les regarde, stupéfaite. Cousin doit alors dénouer Gros-Câlin "par considération pour son incompréhension et ses habitudes" (36). Étant donné qu'il vit avec Gros-Câlin, Cousin ne peut avoir de relation normale avec des femmes.

Dans le cours du roman, Michel Cousin se rapproche de son python et ressent "une espèce de mutualité" (11) avec la bête mal-aimée et minoritaire; il commence à ressembler à Gros-Câlin et tombe lui aussi "dans la catégorie des mal-aimés" (10). Les éléphants des *Racines du ciel* de Gary, Gros-Câlin et Michel Cousin d'Ajar ressemblent en cela aux enfants démunis d'Ajar (Banania, Moïse et Touré). Comme les éléphants des *Racines du ciel* et Gros-Câlin, ils sont mal aimés ou mal compris.

Ces êtres qui vivent à la frontière de ce qu'on pourrait qualifier de normalité se rapprochent aussi d'un autre groupe de marginaux comme Mme Lola "qui travaillait au Bois de Boulogne et avait été champion de boxe du Sénégal" (*La Vie* 16), et M. N'Da Amédée le proxénète dont on retrouvera le corps poignardé dans La Seine (*La Vie* 45).

5.7. Nina Kacew et Madame Rosa

On trouve le thème de la mère dévouée qui se sacrifie pour le bien-être de son enfant dans l'oeuvre de Gary aussi bien que dans celle d'Ajar. Dans *La Vie devant soi*, Madame Rosa, sans doute l'un des personnages les plus émouvants de Gary/Ajar, ouvre un foyer clandestin aux enfants démunis de Belleville. Malgré ses soixante-six ans passés, elle se dévoue pour tous les enfants qu'elle accueille chez elle même lorsque leur cheque n'arrive plus, mais surtout pour Mohammed qu'elle aime comme son fils.

Lorsque, après onze années d'absence, le père de Mohammed vient chez Mme Rosa pour voir son fils, celle-ci craignant que M. Kadir ne vienne le réclamer, lui présente un petit garçon juif qu'elle fait passer pour Mohammed:

- Moïse, dis bonjour à ton papa. [...]

- Pardon? Qu'est-ce que j'ai entendu? Vous avez dit Moïse?

- Oui, j'ai dit Moïse, et alors? [...]

- Moïse est un nom juif [...]. Je vous ai confié un Mohammed [...]. Je ne peux pas avoir un fils juif madame [...]. Je vous ai confié il y a onze ans un fils musulman [...] . Je vous ai donné un fils arabe [...]. Je ne veux absolument pas un fils juif Madame [...]. Je veux mon fils Mohammed Kadir dans l'état arabe où je vous l'ai confié [...]. Je ne veux pas de fils juif. (*La Vie* 191-3)

Madame Rosa explique à Youssef Kadir que par inadvertance, elle a élevé son fils musulman dans la foi juive. Youssef Kadir meurt sous l'effet du choc, mais Madame Rosa est convaincue que, grâce à ce stratagème, elle a réussi à sauver Mohammed d'un sort pire que la mort. Elle explique à Mohammed: "un père qui a été psychiatre, c'est pas du tout ce qu'il te faut, parce que des fois c'est héréditaire" (*La Vie* 201). Comme toute mère protectrice, Madame Rosa a recours au subterfuge pour protéger son enfant.

Le personnage de Madame Rosa avait apparu en germe dans *La Promesse de l'aube* de Romain Gary, sous les traits de Nina Kacew, la mère du narrateur. D'après Catonné, *La Vie devant soi* n'est qu'un "remake ajarien" (144) de *La Promesse de l'aube*. Dans ce roman, Nina Kacew se sacrifie constamment afin de pouvoir assurer le bien-être de son fils. Lorsqu'ils habitent à Varsovie, elle prend soin d'apporter à l'heure du déjeuner, des tartines bien beurrées et du bon chocolat chaud à son fils à

l'école. Le soir, à la maison, elle lui prépare un repas nourrissant en le rassurant qu'elle a déjà pris le sien.

Un jour, en passant près de la cuisine, le jeune Kacew s'aperçoit que sa mère mange les miettes du repas qu'il a laissé. Il se rend alors compte du fait qu'elle se prive de bonne nourriture depuis longtemps, afin de s'assurer que lui au moins soit bien nourri (*Promesse* 78). Plus tard, lorsque Nina Kacew et son fils déménagent à Nice et habitent à l'hôtel-Pension Mermonts, elle cède à son fils la chambre la plus aérée donnant sur le jardin, et occupe la plus petite chambre de l'hôtel.

Le narrateur de *La Promesse de l'aube* raconte aussi comment, lorsqu'il faisait la guerre en Afrique du Nord, la seule chose qui le soutenait dans ses moments de détresse, étaient les lettres que lui envoyait sa mère. Pendant toute la période de la guerre, alors que Romain Kacew est séparé de sa mère, elle le soutient par les encouragements qu'elle formule à distance. Lorsque la guerre est finie, son premier geste est d'aller la voir à l'Hôtel-Pension Mermonts, mais il apprend qu'elle est morte depuis longtemps:

Au cours des derniers jours qui avaient précédé sa mort, elle avait écrit plus de deux cent cinquante lettres, qu'elle avait fait parvenir à son amie en Suisse. Je ne devais pas savoir - les lettres devaient m'être expédiées régulièrement - c'était cela sans doute qu'elle combinait avec amour, lorsque j'avais saisi cette expression de ruse dans son

regard, à la clinique Saint-Antoine où j'étais venu la voir pour la dernière fois. (*Promesse* 386)

L'amour que Nina Kacew (une diabétique) éprouve pour son fils est presque semblable à l'amour que Madame Rosa (gravement malade) témoigne envers Mohammed qui lui n'est pourtant pas son fils. On peut ainsi voir que ces deux vieilles femmes se ressemblent à bien des égards.

Ce qu'on vient de démontrer, c'est que plusieurs personnages qui apparaissent chez Ajar sont souvent des échos des personnages qui sont présents dans l'oeuvre de Gary.

5.8. Motifs communs aux oeuvres de Gary et d'Ajar

Il existe dans *La Promesse de l'aube* et *La Vie devant soi* des motifs identiques tels que l'escalier que doit gravir plusieurs fois par jour Madame Rosa qui habite au sixième étage de son immeuble, et que doit gravir Nina Kacew qui se déplace plusieurs fois par jour pour gérer les affaires de l'hôtel-Pension Mermonts.

Un autre motif commun aux oeuvres de Gary et d'Ajar est le trou souterrain qui sert de refuge aux personnages. C'est l'un des leitmotifs les plus puissants dans ses oeuvres. Gengis Cohn, le jeune Janek et M. Karl de Gary aussi bien que Mme Rosa et M. Salomon Rubinstein d'Ajar ont chacun un "trou juif". Le fameux "trou

juif" où se réfugie Mme Rosa dans *La Vie devant soi* d'Ajar avait déjà apparu dans *Éducation européenne* de Gary publié en 1944.

Dans ce roman, le jeune Janek creuse un trou dans les forêts de Pologne pour se cacher de ses ennemis. Cet endroit le protège: "dans son trou il se sentait à l'abri des hommes" (*Éducation* 9-12). Cette protection n'est qu'illusoire car lorsque Janek allume des brindilles de bois pour se réchauffer, la fumée remplit sa cachette et lui brûle les yeux. À la fin du roman après la fin de la guerre, Janek veut revoir son ancienne cachette; il constate alors que le trou qu'il avait creusé "est comblé comme il convient à une tombe" (*Éducation* 279).

Le motif du trou oppressif apparaît dans une autre oeuvre de Romain Gary, *Gloire à nos illustres pionniers*, dans la section qui s'intitule "l'humaniste". Dans cette nouvelle, Monsieur Karl Loewy, fabricant de jouets à Munich est un homme jovial, optimiste, qui croit "en la nature humaine, aux bons cigares et à la démocratie" (78). Humaniste, il lit Platon, Descartes, Montaigne, Erasme et Heine parce que, ces auteurs plaident en faveur de l'homme. Au moment de l'arrivée au pouvoir d'Hitler en Allemagne, ce vieux juif intégré, germanisé, confiant en la nature humaine, s'installe au sous-sol de sa maison où il compte demeurer jusqu'à la fin de la guerre.

Monsieur Karl se fie à Herr Schutz, son chauffeur et jardinier, et à Frau Schutz, sa cuisinière, pour lui signaler la fin de la guerre. Mais ceux-ci, devenus maîtres de la maison et patrons d'une entreprise prospère, ne lui signalent pas la fin de

la guerre et Monsieur Karl vit dans son trou pendant vingt ans "pour y attendre, bien à l'abri, le retour de la bonne saison" (*Gloire* 76).

Lorsque la bonne saison ne revient pas, il meurt seul dans sa cave. Le sort de Monsieur Karl ne surprend pas tellement. Romain Gary semble indiquer que ceux qui sont forcés par la nature des choses à passer leur vie dans les trous et les ténèbres, sont voués à une existence malheureuse.

Ce thème de l'isolement d'un être dans un espace clos est repris dans *La Vie devant soi* d'Ajar. Dans ses moments de détresse, Madame Rosa prend refuge dans son "trou juif" (*La Vie* 62), sa résidence secondaire. Enfermée dans sa cave, loin des autres, elle se sent en sécurité. Mais cet endroit est loin d'être un lieu douillet: au contraire, il sent "la pisse [...]". Il y avait au milieu un fauteuil rouge complètement enfoncé, crasseux et boiteux [...]. Les murs, c'était [...] des pierres qui sortaient comme des dents" (*La Vie* 38). Le "trou juif" de Mme Rosa ressemble beaucoup au trou creusé par Janek et à la cave de Monsieur Karl. Ce "trou juif" offre une protection oppressive, et, comme la cave de M. Karl et le trou de Janek, donne la claustrophobie et offre une existence étouffante. Madame Rosa meurt dans ce trou.

Un autre leitmotiv commun aux romans signés Gary et à ceux signés Ajar est l'image de la vieille femme qui se farde au moment de mourir. Madame Rosa meurt affreusement à la fin du roman: "Elle n'a plus rien dit du tout et elle est restée là avec un air vide à regarder le mur en face et à chier et à pisser sous elle" (*La Vie* 259). Juste

avant sa mort, dans un de ses rares moments de lucidité et aussi de coquetterie, Madame Rosa veut se faire belle. Mohammed raconte alors: "j'ai dû tenir le miroir pendant qu'elle se maquillait [...] elle a aussi mis sa perruque" (*La Vie* 257).

Momo ajoute plus loin que: "J'ai pris son maquillage et je lui en ai mis sur les lèvres et les joues et je lui ai peint les sourcils comme elle l'aimait. Je lui ai peint les paupières en bleu et en blanc" (*La Vie* 264).

Une scène analogue figure dans le roman *Europa* de Gary publié en 1972. Dans ce roman Malwina Von Leyde, sur le point de mourir, se fait belle:

Elle était assise dans son fauteuil d'infirme, le visage tourné vers la fenêtre, et son visage dans le naufrage total des traits, paraissait un magma informe de blanc, de rouge et de bleu sous une perruque rousse, un visage de vieille maquerele d'un autre âge. (*Europa* 370)

L'image de la vieille femme qui porte une perruque et se farde en attendant la mort apparaît dans l'oeuvre de Gary aussi bien que dans l'oeuvre d'Ajar.

5.9. Le Motif du parapluie et du petit caniche

Malgré le fait que Mohammed soit entouré d'autres enfants dans le foyer tenu par Madame Rosa, il se sent souvent seul et isolé. Pour se protéger de la solitude, il se promène dans les rues de Belleville en portant son ami "Arthur". Arthur est un parapluie habillé en être humain des pieds à la tête:

Je lui avais fait une tête avec un chiffon vert que j'ai roulé en boule autour du manche et un visage sympa, avec un sourire et des yeux ronds, avec le rouge à lèvres de Madame Rosa [...]. Arthur était habillé comme unijambiste avec un soulier de basket bleu et blanc, un pantalon, un veston à carreaux sur un cintre. (*La Vie* 76)

Durant ses randonnées dans les rues de Belleville, Mohammed est accompagné de son ami Arthur. Ce motif du parapluie doté d'une personnalité humaine est très original. Il avait déjà apparu dans un autre roman de Gary *Le Grand Vestiaire* publié en 1948. Dans ce roman c'est une petite fille nommée Josette qui, elle aussi solitaire, se promène avec un parapluie qui lui tient lieu de compagnon.

Momo a un autre ami, un petit caniche gris qu'il a volé d'un chenil de la rue Calfeutre. Il nomme ce chien "Super" et se sent valorisé lorsqu'il se promène dans les rues accompagné de Super: "je me sentais quelqu'un parce que j'étais tout ce qu'il avait au monde" (*La Vie* 25). Mais malgré les liens étroits qui les unissent, Momo vend "Super" à une dame riche pour qu'il trouve chez elle une meilleure vie: "chez Madame Rosa il y avait pas la sécurité et on ne tenait qu'à un fil [...] c'était pas une vie pour un chien" (*La Vie* 26).

Une scène semblable avait apparu encore une fois dans *Le Grand vestiaire*, Dans ce roman Luc, un jeune orphelin est accueilli par le vieux Vanderputte, une sorte de père adoptif, tout comme Madame Rosa est la mère adoptive de Mohammed.

Comme Madame Rosa, Vanderputte, qui lui aussi avait été arrêté par la Gestapo durant la guerre, accueille des orphelins dans son foyer après la guerre. Luc, dont le père a été tué le dernier jour du maquis, vit dans ce foyer, entouré de plusieurs autres jeunes sans parents. Pour lutter contre la solitude, Luc adopte un chien. Mais, un jour, il donne son chien à un G.I. puisqu'il est convaincu que le chien aura une meilleure vie en Amérique, la terre promise.

Par inadvertance ou par manque d'attention Romain Gary reprend souvent les mêmes thèmes, les mêmes motifs et les mêmes personnages lorsqu'il écrit sous le nom d'Ajar. Malgré cela, il a si bien réussi à se camoufler sous le nom d'Ajar que personne de son vivant n'a deviné qu'il était Ajar. Son *alter ego* lui a permis d'exploiter ouvertement le côté farce et même bouffon de sa personnalité. Il a pu exprimer sa nature humoristique, burlesque, satirique et même grossier parfois; il a inventé un style parlé et familier où les jeux de mots, les entorses à la langue et les calembours abondent. Ceci ne figure pas dans les oeuvres avouées de Gary qui sont plutôt graves.

Après ces précisions sur les similarités entre l'oeuvre de Gary et d'Ajar, le chapitre suivant se propose d'aborder un autre aspect de la mystification Ajar, à savoir l'histoire de la mystification littéraire en Europe. Elle démontrera jusqu'à quel point Gary se distingue de tous les autres mystificateurs lorsqu'il invente le pseudonyme Ajar.

Chapitre 6

La Mystification littéraire en Europe

Mais il faut mystifier pour démystifier. Il nous faut une mystification nouvelle [...]. (*Voix de la foule*). À bas la mystification! [...]. Vive la mystification des mystificateurs [...]. Vive la nouvelle mystification.
(Ionesco, *Tueur sans gages*, III, 1649)

Les chapitres précédents ont décrit les événements-clefs de la mystification Ajar. Ce chapitre examine la mystification littéraire en Europe; il trace l'histoire des mystificateurs littéraires et classe par catégorie les divers genres de mystification. Il démontre que Romain Gary n'était nullement le premier auteur en France, ou ailleurs, à mystifier son public et, qu'il n'en est pas le dernier non plus. Il analyse aussi la

mystification Émile Ajar et permet de comprendre ce qui est habituel et ce qui est exceptionnel dans cette affaire.

Les mystifications littéraires abondent en littérature française aussi bien qu'étrangère et ne sont pas une invention nouvelle; elles remontent même à l'antiquité. L'origine du mot *mystification* est très ancienne aussi. Le mot provient du grec "μυστήριον" et du latin *mysterium* voulant dire initier quelqu'un dans des rites secrets. Ainsi "μυστεος" au singulier et "μυστοι" au pluriel désignent un initié. À l'époque gréco-romaine le mot s'appliquait surtout aux cultes religieux secrets qui n'étaient pas reconnus par l'état.

Le mot *mystes* vient aussi du grec "μυσεν" (fermer), qui veut dire "fermer les lèvres et les yeux". *Mystification* suggère ainsi un rite de passage ou une épreuve qui permet à l'initié d'accéder à une connaissance non accessible aux profanes. Une fois cette épreuve subie, l'initié appartient à un groupe secret d'élus et peut devenir, lui aussi, mystificateur.

6.1. Historique de la mystification

Jean-François Jeandillou indique que le mot *mystification* a apparu pour la première fois dans la littérature française en 1768 dans un conte de Diderot qui s'intitule "Mystification ou Histoire des portraits". On le trouve dans la locution "tour joué par une société de mystificateurs" (Jeandillou 98). Au dix-neuvième siècle le mot *mystification* a

pris une tournure toute nouvelle pour signifier entre autres "supercherie littéraire". Les lexicologues indiquent que le mot *supercherie* provient du mot italien *seperchieria*, un "affront excessif" ou une tromperie par rapport aux livres.

Depuis la fin du dix-neuvième siècle le mot mystification comporte un élément de fraude et de tromperie dans le monde des lettres et, effectivement, de nos jours, personne ne met plus en doute que le but essentiel du mystificateur est de tromper ou de duper les autres en abusant de leur crédulité. Quoique les dictionnaires et les éditeurs aient tendance à classer ensemble toutes les fraudes et supercheries littéraires, la mystification existe sous plusieurs formes. Elle apparaît sous les formes du pseudonyme, du canular épistolaire, de l'écrit apocryphe ou anonyme, du plagiat, du pastiche, aussi bien que de la parodie (*Don Quichotte* est une parodie des romans médiévaux).

6.2. Le Pseudonyme

De toutes les formes de mystification, le pseudonyme, "tout nom forgé à plaisir" (Jeandillou 47), est sans aucun doute le plus répandu. Presque tous les artistes qui adoptent des pseudonymes le font souvent par pure fantaisie. Cette pratique qui existe depuis longtemps, s'est surtout développée à la Renaissance lorsque des personnages célèbres prennent des noms autres que le leur.

C'est surtout dans le domaine de la peinture et de la sculpture que les artistes les mieux connus prennent des pseudonymes: Michelangelo Buonarroti se nomme tout

simplement Michelangelo (nom par lequel il sera mieux connu); Joco Robusti est surtout connu comme Le Tintoret, tandis que Raffello Santi ou Sanzio devient tout simplement Raphaël. En adoptant des pseudonymes, ces peintres et ces sculpteurs ne cherchent pas à tromper le public car tout le monde connaît leur véritable identité.

Cette pratique est encore plus répandue dans le monde des lettres: comme les peintres et les sculpteurs, les hommes de lettres voilent leur identité et lèguent un autre nom que le leur aux lecteurs. Des écrivains comme Molière (Jean-Baptiste Poquelin), Blaise Cendrars (Frédéric Louis Sauser), Saint-John Perse (Alexis Saint-Léger Léger), Mark Twain (Samuel Langhorne Clemens) et Voltaire (François-Marie Arouet) sont plus connus par leurs pseudonymes.

Dans chacun des cas cités plus haut, le but des mystificateurs n'est pas simplement de tromper les lecteurs. Ces mystificateurs ont souvent recours au pseudonyme soit par un désir d'amusement, soit par obligation. À l'époque de Molière et de Voltaire par exemple, il n'était pas de bon ton d'avoir une carrière d'écrivain et d'homme de lettres, si on était quelqu'un. Poquelin père était tapissier à la cour de Louis XIV, et Arouet père avait une situation à la cour lui aussi. Donc en adoptant un pseudonyme, Molière et Voltaire tâchaient de protéger le nom de famille. Cependant, ils n'éprouvaient pas le besoin de passer dans un autre corps, de se réinventer, d'oublier leur patronyme et leurs racines. En effet tous les lecteurs connaissent leur vrai noms.

Le Baron de Reiffenberg postule que l'une des raisons qui poussent certains écrivains à cacher leur identité sous un pseudonyme est la crainte (Quérard 24). Reiffenberg a sans doute raison, surtout en ce qui concerne les écrivains qui faisaient de la fiction narrative (le roman). La fiction narrative qui découle du genre "nouvel" (the novel en anglais) est écrite en langue vulgaire. Pendant longtemps, elle n'avait pas le même cachet que la poésie, qui elle, étant la plus ancienne des genres, a eu le temps d'acquérir des titres de noblesse. L'histoire, quant à elle, avait depuis longtemps un certain cachet.

Il arrive alors que des auteurs qui prennent des pseudonymes ne dévoilent pas leur identité, par mesure de précaution. Plusieurs auteurs cachent leur identité comme dans un jeu de devinettes. D'autres encore, surtout ceux en début de carrière (tels que Prosper Mérimée), ont recours au pseudonyme par pudeur ou par timidité. Mérimée se réfugie contre le mauvais goût de livrer au grand public des confidences sur sa personne. Il déclarera plus tard que c'était une erreur de jeunesse.

Même s'il existe un élément de tromperie dans le cas de Mérimée par exemple, les conséquences de sa dissimulation sous son pseudonyme ne sont pas graves, car il a avoué depuis longtemps être Clara Gazul, et aujourd'hui, tout le monde connaît sa véritable identité. J.-M. Quérard, tout en reconnaissant que ce genre de pseudonyme est une tromperie, affirme que les conséquences d'une telle infraction sont moins déroutantes pour la critique littéraire que les autres espèces de mystifications.

Par contre, au dix-huitième siècle il était souvent nécessaire pour des auteurs tel que Beaumarchais (Pierre Augustin Caron) d'écrire sous un pseudonyme pour des raisons de censure.

Quérard fait une distinction parmi les auteurs qui prennent un pseudonyme. Pour lui il faut distinguer entre: (a) les auteurs "apparents", (b) les auteurs "anonymes", (c) les auteurs "apocryphes", (d) les auteurs "supposés", (e) les auteurs "déguisés" et (f) les auteurs "plagiaires" (25-26).

6.3. Les Auteurs apparents et les auteurs anonymes

D'après Quérard, les auteurs "apparents" constituent une variante des auteurs pseudonymes. La différence en est que l'auteur "apparent", tout en dissimulant son identité sous un pseudonyme, ne fait nullement croire en l'existence de son double. Son identité n'est ni révélée, ni voilée. Voltaire appartient à ce groupe de mystificateurs car, souvent, il se crée des pseudonymes par humour ou par simple amusement. Né François-Marie Arouet, il prend le nom de Voltaire, anagramme de Arouet le jeune: ceci ne l'empêche pas pour autant de publier sous d'autres pseudonymes. Par pure fantaisie, il prend des noms farfelus tels que "le Docteur Akakia" (traduction d'un vrai nom), "le Rabbin," "Irénee Aléthès," "Ivan Aléthof" et "Catherine Vadé". Il ne dupe personne.

La pratique de publier sous un pseudonyme se poursuit au dix-neuvième siècle. Honoré de Balzac publie des oeuvres sous les pseudonymes de "Lord R'Hoone"

(anagramme d'Honoré), "Horace de Saint-Aubin" et ainsi de suite. Henri Beyle, plus connu sous son pseudonyme Stendhal, publie sous plusieurs noms d'emprunt tels que "Théodore Bernard," "Louis-Alexandre-César Bombet", "M.B.A.A", "Catonet", "F. de Lagenais" et "Salviati". Plus tard, Max Jacob prend le pseudonyme de "Morven le Gaélique". À l'époque de la publication des oeuvres de ces auteurs, lorsque ces noms fictifs paraissent sur la couverture de ces oeuvres, tout le monde connaît la véritable identité de ces auteurs apparents.

Quant aux auteurs anonymes, Quérard postule que certains auteurs publient des oeuvres anonymes parce que, se croyant médiocres, ils ont peur de se montrer au public. Jean-Guillaume Sigismond Lindner, savant allemand, appuie ce postulat: il affirme lui aussi dans la préface du *Dictionnaire des écrivains pseudonymes allemands* que les auteurs publient des oeuvres anonymes par modestie ou parce qu'ils craignent la persécution religieuse ou politique. D'après Lindner, les auteurs anonymes attendent le moment propice de révéler leur identité au public.

Parfois, il était même avantageux pour certains auteurs de conserver leur nom d'adoption pour des raisons pécuniaires, même si en ce faisant, ils se privaient de l'occasion de se faire connaître dans le monde des lettres.

6.4. L'Apocryphe

De nos jours le mot "apocryphe" s'applique aux livres faussement attribués à des écrivains de l'antiquité ou à des récits de la Bible. Mais J.-M. Quérard suggère que: "tout livre publié par son auteur sous le nom d'un homme de lettres connu, qu'il appartienne à l'antiquité ou à l'époque moderne" (4) est un ouvrage apocryphe. Par ailleurs, le dictionnaire *Robert* affirme qu'apocryphe est tout objet "dont l'authenticité est au moins douteuse". L'apocryphe est très répandue durant la Renaissance lorsque beaucoup d'auteurs se livrent à cette sorte de mystification. Elle débute surtout au seizième siècle.

En 1583 paraît pour la première fois un livre intitulé *La Consolation* qui est censé être de Cicéron. Mais il a été établi depuis ce temps que c'est un ouvrage de Charles Sigonius de Modène, grand savant de l'époque. Sigonius n'avait découvert que quelques fragments de *La Consolation* de Cicéron, au moyen desquels il a recomposé un ouvrage entier et l'a fait passer pour l'oeuvre du grand orateur latin.

En 1789, Simon Despréaux prétend publier en un volume, une oeuvre posthume de Jean de La Fontaine et se présente comme l'éditeur de cette oeuvre. Mais, ce volume ne contient même pas une seule oeuvre de La Fontaine (Quérard 29). Plus tard, au dix-neuvième siècle, une comédie inédite s'intitulant *Le Docteur amoureux* censée être de Molière est représentée au Théâtre Français à Paris. Il a été établi depuis ce temps-là, que cette pièce avait été composée par un certain Ernest de Calouze (Quérard 30).

La pratique de faire passer des oeuvres contemporaines pour des oeuvres de l'antiquité existe depuis longtemps. Les fouilles de Pompéi et de Herculanium fournissent l'occasion à plusieurs personnes de prétendre qu'ils ont découvert sous ces ruines, des fragments d'oeuvres ou même des oeuvres intactes écrites à l'époque classique. Durant les fouilles de Pompéi, une maison où est inscrit le nom de Polybe (le fils de Lycortas qui s'était fait exiler à Pompéi) est découverte dans les ruines. Un certain Monsieur Pierquin, à l'Académie de Bourges, propose alors que les fables de Phèdre (Phaidros) dont s'inspire La Fontaine, sont en fait de ce même Polybe. Il propose que Phèdre est le surnom de Polybe et que ce dernier avait pris le pseudonyme Phèdre pour ses fables (Quérard 67).

Quelques critiques ne font aucune distinction entre un ouvrage supposé et un ouvrage apocryphe: pour Roger Picard, par exemple, toute oeuvre publiée sous un pseudonyme est un travail apocryphe. La mystification Émile Ajar serait donc un apocryphe selon la définition de Picard. Mais pour J.-M. Quérard tel n'est pas le cas: les ouvrages apocryphes sont des oeuvres publiées sous un nom d'auteur connu. Quérard qualifierait d'oeuvre apocryphe *La Consolation* de Cicéron.

6.5. La Supposition d'auteur

La supposition d'auteur est un autre genre de mystification. Dans un tel cas, l'auteur fait passer pour authentique un ouvrage qui est faux. Picard le résume de cette

manière: "La supposition d'auteur consiste à faire croire en l'existence d'un écrivain purement imaginaire dont on donne comme authentiques des livres écrits sous son nom par les mystificateurs" (8).

Les suppositions d'auteurs et d'artistes abondent au seizième siècle. Michel-Ange a recours à ce stratagème lorsqu'il feint d'avoir tiré des fouilles de Rome un torse dont il avait conservé les extrémités. Mais il avoue plus tard que ce torse est de lui. Il attend que l'admiration publique vienne lui en rendre l'honneur (Quérard 18).

En principe, les auteurs qui se font passer pour des auteurs supposés publient des livres sous le nom d'un personnage connu qui est plus ou moins historique, ou sous le nom d'un personnage imaginaire. Quoique l'auteur supposé soit souvent un substitut plausible, jusqu'à la publication de l'oeuvre en question, rien de ce personnage historique n'est venu le faire connaître comme écrivain. Le jeu consiste alors à chercher qui est ce personnage supposé. Voltaire au dix-huitième siècle, Mac Pherson (Ossian) et Mérimée (Clara Gazul) au dix-neuvième siècle font appel à ce stratagème lorsqu'ils publient des oeuvres sous des noms imaginaires.

D'après Quérard, ces auteurs supposés veulent faire accepter des livres d'une incontestable authenticité (4) pour de vrais livres. Cependant, tout en reconnaissant que la supposition d'auteur est définitivement une tromperie pratiquée par certains afin de conférer la dignité d'auteur à un individu dont l'existence est attestée, Quérard remarque

qu'elle est très différente de l'apocryphe, n'ayant pas autant d'inconvénient dans ses conséquences.

Ce qui distingue l'auteur supposé de l'auteur apocryphe, c'est le fait que l'auteur supposé a recours au subterfuge par modestie. D'après Charles Nodier, c'est souvent un auteur qui aime jouir sous le nom d'un autre "du succès de son propre talent" (Jeandillou 67). Roger Picard est lui aussi de cet avis car il affirme que l'auteur supposé veut avant tout garder l'incognito le plus complet. Il ne veut pas imiter ou pasticher un auteur connu, il invente l'auteur lui-même et crée les oeuvres qu'il lui prête (161).

6.6. L'emprunt et les auteurs imaginaires

Jeandillou et Quérard affirment que l'emprunt qui était une pratique normale constituent une autre forme de mystification. Durant les siècles passés, il se pratiquait fréquemment. En Angleterre, William Shakespeare emprunte ses idées des auteurs contemporains ou des siècles passés sans aucune hésitation. Lorsqu'un critique l'accuse d'avoir pris une scène entière de l'oeuvre d'un auteur contemporain, Shakespeare lui répond: "C'est une fille que j'ai tirée de la mauvaise société pour la faire entrer dans la bonne" (Quérard 72).

En France au dix-septième siècle, Racine, Corneille et La Fontaine puisent leur inspiration dans les oeuvres de l'antiquité. Ils empruntent des thèmes à des oeuvres très connues de l'antiquité, mais se font un devoir d'indiquer leur source dans leur préface.

Ainsi Corneille, en indiquant ses sources pour *Le Cid* déclare s'être inspiré de Guillén de Castro "qui a mis ce fameux événement sur le théâtre avant moi" (Corneille 21); il ajoute en outre "je me suis contenté du texte de l'historien" (21).

Racine s'inspire en grande partie de la mythologie grecque et dans ses préfaces indique sa fidélité aux sources. Il prend garde cependant de ne pas se faire un esclave de ces textes. Pour *Iphigénie*, par exemple, il mentionne l'*Agamemnon* d'Eschyle, l'*Electre* de Sophocle, *De la Nature (livre I)* de Titus Lucretius Carus (Lucrèce), les *Satires (livres II, satire 3)* d'Horace et *L'Iliade* de Homère. Mais il explique qu'à partir de ces oeuvres, il a créé une version originale d'*Iphigénie* (*Iphigénie* 33-7).

Pour *Phèdre* Racine déclare s'être inspiré de la pièce *Hippolyte* d'Euripide dont il emprunte l'intrigue et la première scène où Phèdre fait son entrée. Il admet quand même avoir "suivi une route un peu différente de celle de cet auteur pour la conduite de l'action [...] (et) même pris soin de la rendre moins odieuse qu'elle n'est dans les tragédies des anciens." (*Phèdre*,1). Il reconnaît aussi avoir emprunté des idées et des phrases à Sénèque. Mais il existe aussi des échos de la pièce *Hypolyte* de Robert Garnier (1573), et de Gilbert (1647) dans *Phèdre*. Par ailleurs quelques vers semblent s'inspirer de Virgil.

La Fontaine remanie les fables de Phèdre, d'Ésope et de Pilpay. Mais il imite aussi des fabulistes moins connus tels que Babrius et Aphthonius, l'humaniste Abstenius du 15e siècle et Rabelais. Il retranche ou amplifie les fables, change les incidents et les personnages si cela s'avère nécessaire à la composition des fables. Il prend soin,

cependant, de citer ses sources. Pour son second recueil de fables par exemple il explique: "Seulement je dirai par reconnaissance, que, j'en dois la plus grande partie à Pilpay, sage indien" (15). Molière, pour sa part, justifie cette pratique et dans une phrase connue déclare: "Je prends mon bien où je le trouve" (*L'Avare* 21). Pour ces auteurs du dix-septième siècle, puiser son inspiration des oeuvres de l'antiquité est acceptable et même préférable.

Quant à la catégorie d'auteurs imaginaires, Jeandillou distingue surtout le type des auteurs prétendus qui sont hors du cadre narratif. Il ne s'agit pas de mensonge dans ce cas-ci mais de simulation, le nom de l'auteur paraît en couverture, mais celui-ci prétend présenter l'écrit d'un autre auteur. Les romans épistolaires (*Les Lettres persanes* de Montesquieu, *Les Liaisons dangereuses* de Laclos, *La Nouvelle Héloïse* de Rousseau) sont présentés sous cette forme, et sont typiques des romans du dix-huitième siècle. L'extrait qui suit, où le rédacteur de l'oeuvre explique la source de ce roman, peut servir d'exemple:

Cet ouvrage [...] que le public trouvera sans doute trop volumineux, ne contient pourtant que le plus petit nombre des lettres qui composaient la totalité de la correspondance dont il est extrait. Chargé de la mettre en ordre par la personne à qui elle était parvenue [...] j'ai tâché de ne conserver en effet que les lettres qui m'ont paru nécessaires, soit à

l'intelligence des événements, soit au développement des caractères.

(Laclos 15)

Il est évident qu'il ne s'agit pas de mensonge ici mais d'une convention. Le lecteur sait fort bien que cette préface du rédacteur n'est qu'une simulation: ainsi il participe volontairement au jeu de l'auteur.

6.7. Les Auteurs hétéronymes, cryptonymes et polynomes

La catégorie d'auteurs "hétéronymes" consiste à faire croire en l'existence d'un écrivain purement imaginaire dont on donne comme authentiques des livres écrits sous son nom (Picard 8). Dans un tel cas

l'écrivain ne se borne plus à cacher sa personnalité sous un pseudonyme facile à découvrir, il garde l'incognito le plus complet. Il ne se borne pas à imiter ou à pasticher un auteur connu, il invente l'auteur lui-même, crée les oeuvres qu'il lui prête, portant ainsi l'apocryphe à son plus haut degré.

(161)

Les cryptonymes sont une autre variante du pseudonymat. Alcofrybas Nasier est le cryptonyme de François Rabelais, Yourcenar celui de Séraphine Calobarcy, Pablo de Herlañes celui de Paul Verlaine, et Bleucoton celui de Stéphane Mallarmé. Parfois les écrivains préfèrent des noms de terre à leurs noms de famille. Pierre Eyquem de

Montaigne devient Montaigne, Pierre Carlet de Marivaux se nomme Marivaux et Jean de Rond D'Alembert devient D'Alembert.

Quand aux polynymes, il s'agit des multiples pseudonymes que les auteurs se créent, tels les noms dont Voltaire (Arouet) et Henri Beyle se couronnent (voir 6.3). Cette définition s'applique aussi à Romain Gary, car il publie sous plusieurs noms d'emprunt (Shotan Bogat, Émile Ajar et Fosco Sinibaldi).

Comme, il vient d'être démontré, les pseudonymes ne sont pas tous semblables; les auteurs y ont recours soit par un simple désir de changer de peau, soit pour duper le public.

6.8. Histoire de la mystification littéraire en France

Quérard suggère que lorsqu'un auteur ne possède pas un nom harmonieux à léguer à la postérité, il prend souvent un autre nom. En effet, plusieurs noms en France (comme partout ailleurs) ont une origine très ordinaire. Des noms tels que "Legalant", "Petit", "Le Grand", "Noiraud", "Legras" et "L'étourdi" étaient donnés à l'origine selon les qualités et les défauts physiques ou moraux des individus.

D'autres noms tels que "Boulangier", "Le Forestier", "Maçon", "La Tasse", "Brodeur", "Octobre" et "Mercredi" étaient pris des jours de la semaine, des mois, des métiers, ou des choses servant aux besoins de la vie quotidienne. D'autres encore, sont plus grossiers, plus vulgaires tels que "Bordel", "Baudet", "Torchon", "Catin", "Cochon",

“Cornu”, “Fesse” etc. Louis XV a même décrété que le nom de famille de certains de ses sujets soit changé en “Bourreau”, “Cailloux”, “Canard”, “Hareng”, “Piedevache” entre autres.

Quérard remarque que des noms tels que “Fesse”, “Torchon”, “Hareng” et “Cornu” doivent donner lieu à des sarcasmes et sont sans doute ridicules dans le monde quotidien, mais d’autant plus dans le monde des lettres et des sciences. Dès le seizième siècle de tels noms sont jugés malsonnants et même désobligeants. Si on est homme de lettres, il y a aussi nécessité de changer de nom. On le change ou le traduit alors par obligation.

Le seizième siècle témoigne d’une véritable transformation et de falsification de noms lorsque plusieurs écrivains grécisent ou latinisent leurs noms. Le médecin du roi François Premier qui s’appelle “Sans Malice” prend le nom du “Docteur Akakia” (“Sans malice” en grec). D’autres auteurs au seizième siècle choisissent des pseudonymes par pure fantaisie. Durant ce siècle paraissent des noms tels que “Alcofribas”, “Le traverseur des voies périlleuses”, “Un vrai zèle”, “Brédin le Cocu” et ainsi de suite.

Au dix-septième siècle plusieurs écrivains prennent des pseudonymes. C’est une époque marquée par des disputes théologiques entre les “Dignes enfants de Loyola”, les Jansénistes et les Protestants. Ces disputes donnent lieu à d’innombrables écrits et comme chaque auteur veut cacher sa véritable identité, chacun prend un pseudonyme.

Parfois un auteur prend le nom de sa mère si ce nom semble plus harmonieux que celui de son père, et parfois, il a recours à l'anagramme de son nom paternel. Au dix-septième siècle, Thomas Corneille, qui ne veut pas être confondu avec son fameux frère Pierre, prend le nom de "de l'Ile", le nom toponymique de sa mère. Molière se moque de cette manie d'anoblir son nom en introduisant la scène de satire suivante dans *L'École des Femmes*:

Chrysalde: Je me réjouis fort Seigneur Arnolphe.

Arnolphe: Bon!

Me voulez-vous toujours appeler de ce nom!

Chrysalde: Ah! malgré que j'en ai, il me vient à la bouche,

Et jamais je ne songe à monsieur de la Souche.

Qui diable vous a fait aussi vous aviser

À quarante-deux ans de vous débaptiser,

Et d'un vieux tronc pourri de votre métairie

Vous faire dans le monde un ton de seigneurie?

Arnolphe: Outre que la maison par ce nom se connaît

La Souche plus qu'Arnolphe à mes oreilles plaît.

Chrysalde: Quel abus de quitter le vrai nom de ses pères

Pour en vouloir prendre un bâti sur les chimères!

De la plupart des gens c'est la démangeaison,

Et sans vous embarrasser dans la comparaison,

Je sais un paysan qu'on appelait Gros-Pierre,
 Qui n'ayant pour tout bien qu'un quartier de terre
 Y fit tout à l'entour, faire un fossé bouleux,
 Et de monsieur de l'Ile en prit le nom pompeux.

(École des femmes, I, ii)

D'autres auteurs prennent des pseudonymes pour se donner une origine aristocrate. Michel Eyquem change son non d'origine marchand en se nommant Montaigne, le nom de terre acheté par son arrière grand-père; Charles-Louis de Secondat se donne le nom du baron de Montesquieu, Georges Louis Leclerc devient le comte de Buffon; Marie Jean Antoine Nicholas de Carital devient le Marquis de Condorcet.

Au dix-huitième siècle Voltaire (Arouet) est sans doute le plus célèbre de tous les auteurs qui prennent des pseudonymes. Ayant gagné beaucoup d'argent, il est le premier à vivre de sa plume. Parfois les auteurs qui sont criblés de dettes tels que Balzac, vendent et revendent deux ou trois fois le même manuscrit sous des titres différents et des noms différents. Ils ont recours au pseudonyme parce que c'est "un préservatif contre les créanciers" (Quérard 55).

Dans certaines régions d'Europe, le mouvement intellectuel des grandes villes n'arrive que bien tard. Au dix-huitième siècle, au fond des provinces en France, il y a beaucoup de familles nobles ruinées, étant donné que le patriarche était obligé par Louis

XIV de vivre à la cour à Paris. Comme, il fallait suivre la mode en tout, ils négligeaient la terre, source de toutes leurs fortunes.. Si on est écrivain dans ces endroits retirés de la France, il faut cacher ce fait ou courir le risque de se faire ridiculiser. Faire des romans, c'est se faire déconsidérer, surtout si on est noble, car c'est un travail rémunéré.

Durant la Révolution française de 1789, les nobles qui veulent rester en vie cachent leur blason et dissimulent leurs titres. Par ailleurs, les hobereaux (gentilshommes campagnards de petite noblesse vivant sur leurs terres) se glorifient de ne savoir ni lire, ni écrire, et par conséquent ne risquent pas d'être guillotинés.

6.9. Le Dix-neuvième Siècle: anoblissement des roturiers

Sous l'Empire (1804-1815) les nobles qui n'ont pas trouvé la mort s'empresment de déterrer leurs titres. Par ailleurs Napoléon crée une nouvelle noblesse. Avec la Restauration (1814-1830) et la Monarchie de juillet (1830-1848) plusieurs émigrés et d'autres personnes fidèles à la cause royaliste sont anoblis. Louis-Philippe crée une nouvelle noblesse parmi les hommes politiques et l'administration, moyennant une bonne somme d'argent. Ces pseudo-nobles adoptent facilement la particule de noblesse ou des traits-d'union à leurs noms. À l'exemple des hommes politiques ou administrateurs, les hommes de lettres fondent entre eux une noblesse qui se dispense de l'agrément du roi.

Le Journal des Débats d'octobre 1847 traite de ridicule cette manie des fils de marchands et de paysans (à savoir les hommes politiques, les administrateurs et les

hommes de lettres) de vouloir s'anoblir. Ce journal établit qu'à cette époque trois types de noblesse se côtoient: l'ancienne noblesse (ceux dont le titre de noblesse date d'avant la Révolution), la noblesse impériale (ceux qui ont été anoblis sous l'Empire) et la noblesse spontanée (ceux qui se sont anoblis tels que les hommes de lettres).

Le dix-neuvième siècle témoigne d'une véritable floraison d'anoblissement spontané. Des écrivains prennent le nom de villes, de villages et de hameaux et les impriment sur le titre de leurs livres. Aussi Ducrotay, né à Banville, prend-il le nom de "de Banville"; Ros dit Beaupin se nomme "de Beauvallon"; le frère de Pétrus Borel essaie de rattacher le nom aux Borel d'Hauterive. Balzac, pour sa part, ajoute la particule à son nom et devient Honoré de Balzac. À l'aide de la particule tous ces auteurs revendiquent un nom noble. Pour J.-M. Quérard, qui lui est bibliomane, et non pas homme de lettres, cet état de pseudonymie est une supercherie qui n'a d'inconvénient que pour l'histoire littéraire.

Au dix-neuvième siècle, Henry Beyle adopte plusieurs noms sans pour autant ajouter la particule à son nom pour s'anoblir. Mais il ne signe jamais ses oeuvres de son véritable nom. Il est tour à tour "Cotonet" "Salviati", "Stendhal" (le pseudonyme le plus connu), Viscontini etc. Voulant rester inconnu, il crée une biographie fictive pour chaque pseudonyme, leur donnant parfois une qualité, parfois une autre: il se déclare ainsi officier de cavalerie, douanier, marchand de fer etc.

Presque à la même époque, soit en 1829 Sainte-Beuve fait publier une prétendue biographie posthume dans un ouvrage intitulé *Vie, poésies et pensée de Joseph Delorme*, sans laisser figurer son propre nom. Ce volume raconte la vie de Joseph Delorme; il est censé avoir été rédigé par un ami qui voulait rester anonyme. On sait maintenant que cette oeuvre est de Sainte-Beuve.

6.10. Une amusante supercherie: *Clara Gazul*

En 1825 Prosper Mérimée entre dans la littérature par la porte de la mystification lorsqu'il publie un livre intitulé *Le Théâtre de Clara Gazul, comédienne espagnole*; sur la couverture se trouve le titre *Collection des théâtres étrangers*. À l'époque, Clara Gazul est censée être une comédienne espagnole qui publie ses écrits. Il existe aussi une notice d'un certain Joseph L'Estrange, le soi-disant traducteur de Clara Gazul (aujourd'hui reconnu comme personnage fictif). Cette notice biographique est imprégnée d'une atmosphère historique réaliste.

Par ce livre Mérimée crée une dramaturge imaginaire vraisemblable. En 1827 Mérimée continue dans la mystification lorsqu'il publie *La Guzla, Choix de poèmes illyriques* d'un prétendu barde serbo-croate Hyacinthe Maglanowich, recueillis dans la Dalmatie, la Bosnie, la Croatie et l'Herzégowine. Encore une fois Mérimée ajoute un élément de réalisme au soi-disant poète de *La Guzla*.

Comme le signale Trahard, Mérimée ne veut pas seulement créer une mystification: il y ajoute un cadre historique et donne une réalité au personnage (168). Dans l'introduction du *Théâtre de Clara Gazul*, Maxime Revon suggère, qu'étant donné que cette pièce a une allure de pamphlet antireligieux et fait indirectement l'éloge du libéralisme espagnol que la monarchie française vient de combattre, Mérimée a intérêt à cacher son identité pour échapper aux persécutions de la Congrégation qui était alors puissante.

En mai 1891, paraissent *Les Cahiers d'André Walter*, la présumée oeuvre posthume d'un jeune écrivain disparu prématurément. Cette oeuvre porte la signature d'un certain "Pierre C" un ami d'André Walter qui fait publier ses oeuvres. Lors de sa parution, les critiques soupçonnent immédiatement qu'il s'agit d'une supercherie. Augustin Filon déclare dans *La Revue Bleue* du 16 mai 1891: "Il n'est pas difficile de deviner que ce dénouement, comme la personnalité même d'André Walter, est de pure fantaisie" (Jeandillou 115). Le 10 octobre 1891, un critique anonyme dans *Le Magazine Littéraire* annonce "André Walter s'appelle André Gide, ce mort est bien vivant" (Jeandillou 115). Ce jeu de Gide ne surprend pas en effet, car son oeuvre romanesque est riche en journaux intimes tels que le carnet d'Alissa qui apparaît dans la dernière partie de *La Porte étroite*.

Par contre, il existe d'autres écrivains qui se travestissent en femmes. Au début du vingtième siècle, Apollinaire prend le nom de Louise Lalanne et contribue à un journal

sous la rubrique "littérature féminine". Alphonse Daudet et Paul Arène publient *Les Poèmes idiots*, une prétendue oeuvre posthume de Myriam Mester, femme poète; il existe même une introduction et des notes de Gaston Picard. Durant la deuxième partie de ce siècle Raymond Queneau fait publier *Les Oeuvres complètes* d'une certaine Sally Mara chez Gallimard. Cette oeuvre contient une préface de l'Irlandaise Sally Mara elle-même, dont l'oeuvre a été traduite par Michel Presle. Michel Presle est un autre pseudonyme de Queneau. Cette oeuvre est réimprimée en 1981 dans la collection "L'Imaginaire". Ce à quoi on a affaire ici, c'est une double supposition d'auteur et de traducteur aussi bien qu'à un changement de sexe.

Les exemples qui ont été mentionnés montrent que la mystification littéraire est bien vivante en France. En fait, elle continue jusqu'à l'heure actuelle. Le cas le plus connu de mystification en France ces dernières années, est celui d'Erik Orsenna, un auteur renommé qui publie *Loyola's blues* en 1974, *La vie comme à Lausanne* en 1977 (pour lequel il reçoit le Prix Roger-Nimier) et *Une comédie française* en 1980. En 1988, il obtient le Prix Goncourt pour son roman *L'Exposition coloniale*. Or, Erik Orsenna est le pseudonyme d'Eric Arnoult, ex-conseiller littéraire du Président François Mitterrand, haut fonctionnaire de l'État français et conseiller de Roland Dumas au Quai d'Orsay. Il se justifie de la manière suivante: "J'ai publié mon premier roman en même temps que ma soutenance de thèse d'économie. Donc, on m'a conseillé de dissocier nettement mes deux carrières" (Lane 42).

6.11. Les Mystificatrices

Au Moyen-Âge, un petit groupe de femmes de la noblesse, telles qu'Aliénor d'Aquitaine, protectrice des lettres, Marie de France, la poétesse du XIIe siècle, Christine de Pisan, femme de lettres du XVe siècle, ont des droits très étendus et se distinguent dans la poésie et la musique. Mais jusqu'à fort récemment les femmes-écrivains étaient assez rares: le monde des sciences et des lettres était réservé aux hommes. C'est surtout pour cela que souvent des écrivains-femmes telle que la duchesse de Duras n'inscrivent pas leur nom à leurs travaux, de peur d'être prises pour des pédantes.

En fait, jusqu'au début du vingtième siècle, si on est femme-écrivain, il est difficile de se faire prendre au sérieux dans un monde où les hommes règnent en maîtres. Si on a du talent littéraire et qu'on est femme, la meilleure façon de se faire reconnaître est de se travestir en homme. D'ailleurs il peut être très dangereux de se déclarer femme-écrivain. Au dix-neuvième siècle, une femme entreprend la rédaction de son autobiographie et la signe "Pérégrinations d'une paria". Elle reçoit pour sa peine une balle près du poumon, tirée par son mari (Desanti 93).

C'est l'une des raisons pour lesquelles les femmes-écrivains prennent souvent des pseudonymes pour écrire. En 1662, lorsque Madame de la Fayette publie son premier roman, *La Princesse de Montpensier* elle cache son identité sous le nom de Segrais. En 1670, elle publie le roman précieux *Zaïde*, sous ce même nom. Ce n'est qu'en 1678 qu'elle fait paraître le roman d'analyse *La Princesse de Clèves* sous son propre nom.

C'est pour cela que de temps en temps, on conteste l'attribution de *La Princesse de Clèves* à Madame de la Fayette, car elle a nié elle-même en être l'auteur. Cependant, à cette époque, les dames nobles ne pouvaient admettre qu'elles écrivaient des romans. C'est pour cette même raison qu'au dix-neuvième siècle, Aurore Dupin se travestit en homme pour devenir George Sand et que Marie d'Agoult prend le nom de plume masculin, Daniel Stern.

Au dix-neuvième siècle, en Angleterre une jeune fille à marier doit se comporter de manière modeste. Lorsque Jane Austen rédige des romans, elle ne peut trouver de maison d'édition qui veuille publier des romans écrits par une jeune femme, on l'accuse même d'exhibitionnisme. Par ailleurs, à cette époque, les femmes du milieu bourgeois ou aristocrate ne travaillent, ni n'écrivent des romans. Ainsi Jane Austen publie-t-elle ses quatre premiers romans *Sense and Sensibility* (1811), *Pride and Prejudice* (1813), *Mansfield Park* (1814) et *Emma* (1816) en cachant son identité et en les signant du nom anonyme "A Lady" (une dame). Ce n'est que deux ans après sa mort avec la publication de *Northanger Abbey* et *Persuasion* (1818) que son identité est dévoilée au public anglais.

Quelques années plus tard, en 1846, les soeurs Brontë (Emily, Anne et Charlotte) publient à leurs propres frais un volume de poésie sous les pseudonymes masculins de Ellis, Acton et Currer Bell. En 1847, Emily et Anne Brontë publient ensemble leurs romans *Wuthering Heights* et *Agnes Grey* sous les pseudonymes masculins Ellis et Acton

Bell. Lorsque leur soeur Charlotte publie son roman *Jane Eyre* en 1847, elle se déguise sous le pseudonyme masculin de Currer Bell.

En 1859, Mary Anne Evans Cross, l'un des meilleurs romanciers anglais et l'une des femmes-auteurs les plus importantes, se travestit en homme et prend le nom de plume George Eliot lorsqu'elle publie ses romans *Adam Bede* (1859), *The Mill on the Floss* (1860), *Silas Marner* (1861) et *Middlemarch* (1871-1872) entre autres.

En France, au début du vingtième siècle paraît le livre de Sidonie-Gabrielle Colette *Claudine à l'école* écrit en collaboration avec son mari Willy (le pseudonyme de Henry Gauthier Villars) et publié sous le nom de Willy. Dans la préface du roman, Willy affirme qu'il publie le journal d'une jeune fille. Durant les trois prochaines années, Colette écrit trois autres romans sur "Claudine", publiés tous les trois sous le nom de Willy. Colette revendique ce livre plus tard comme le sien. Aujourd'hui le livre paraît sous la double signature de Colette et de Willy.

6.12. La Mystification Ajar

Les pseudonymes font légion en littérature française et étrangère et ne sont nullement une invention du 20e siècle; ils sont vieux comme le monde. Romain Gary se joint ainsi à une longue tradition de mystificateurs lorsqu'il adopte le pseudonyme d'Ajar. Dans la plupart des cas, les auteurs mystifient le public soit par nécessité (comme les femmes qui de peur d'être ridiculisées dans leur milieu bourgeois ou aristocrate n'avaient

aucun autre recours que de se cacher sous un nom de plume masculin), soit par simple jeu (tels Mérimée, Voltaire et Mark Twain), soit par convention (comme le soi-disant éditeur des *Liaisons dangereuses*). On doit alors déterminer à quelle catégorie Gary/Ajar appartient.

Quérard postule que les pseudonymes découlent souvent d'un "besoin d'euphonie [...] du besoin d'amplifier son nom à la mesure de l'être que l'écrivain sent en soi" (91). Ceci implique que souvent les auteurs changent de nom voulant avoir un nom dont la sonorité est plus agréable ou un nom plus exotique que le leur sur la couverture de leurs livres. D'après Philippe Lane, Erik Orsenna confirme cette constatation. En parlant de son pseudonyme Orsenna déclare:

Ce patronyme me plaît de plus en plus, car on m'a dit qu'il venait d'un nom étrusque. Il semble qu'il y ait eu un Dieu du nom de Porsenna; ce qui ne sonne pas très bien, mais, lorsqu'on supprime le "p", ça devient beaucoup plus joli. [...] Je me sens très bien dans ce nom là. (Lane 42-3)

Ces notions peuvent expliquer quelques aspects de la vie de Romain Gary. Juste avant la Deuxième Guerre mondiale, son nom civil de Roman Kacew, loin d'être malsonnant, empêche sa carrière militaire de démarrer. Ce nom est sans doute jugé trop étranger en France à cette époque, mais lorsqu'il adopte le nom moins étranger de Gary, il connaît le succès dans le corps diplomatique, aussi bien que dans le monde des lettres. Il semble que dans un cas pareil, le choix d'un nom influence le succès d'un auteur.

Quérard affirme aussi que parfois les auteurs prennent un pseudonyme lorsque leur propre nom ne produit plus d'effet dans le milieu littéraire (41). Ceci s'applique encore une fois à Romain Gary. Vers le début des années soixante-dix, les critiques jugent que Gary est un auteur passé. Romain Gary tente alors de se renouveler en se travestissant en Bogat, Sinibaldi et finalement en Ajar.

Chaque fois que Gary change de nom, ce n'est pas pour des raisons d'euphonie (Shotan Bogat, Fosco Sinibaldi et Émile Ajar ne sont guère plus harmonieux que Gary) mais plutôt pour déjouer les critiques qu'il accuse surtout de mal lire. Il se travestit afin d'être jugé à sa juste valeur.

Selon J.-M. Quérard le cas d'Émile Ajar est typique de la fraude littéraire; pour lui tous les pseudonymes sont des impostures, même si l'auteur ne songe nullement à faire accepter un livre supposé pour un livre historique, mais veut seulement faire accepter un nom qui est faux. Il est incontestable que Romain Gary veut faire passer le nom d'Émile Ajar (un nom qui est faux) pour un vrai nom.

À l'opposé de Voltaire, Mérimée, Jane Austen et d'autres auteurs qui agissent souvent seuls, Romain Gary a ses complices: Paul Pavlowitch, Michel Cournot, Claude Gallimard et son avocat. Ensemble, ils forment un petit groupe d'élus qui font partie du complot, trompent les lecteurs et les critiques et abusent de leur crédulité. Un autre aspect de la mystification Ajar qui d'emblée la rend différente des autres, est que souvent, des auteurs tels que Prosper Mérimée, prennent un pseudonyme par timidité en début de

carrière ne voulant pas s'afficher ouvertement. Mérimée n'a que vingt-trois ans lorsqu'il se travestit en Clara Gazul. Mais Romain Gary prend le nom d'Émile Ajar, change de peau et fuit ses racines parentales et culturelles à la fin de sa carrière.

D'après les critiques, les pseudonymes ne constituent qu'une supercherie bénigne car, par un simple désir d'amusement, l'auteur ne demande au lecteur que de participer à un jeu, il ne cherche nullement à l'induire en erreur. Il ne crée pas, non plus, une biographie fictive pour son pseudonyme. Mais, en créant Émile Ajar, Romain Gary l'affecte non seulement d'une personnalité qui est très différente de la sienne, mais lui donne le visage et le corps de Paul Pavlowitch.

Il existe un autre aspect de l'affaire Ajar qui la rend différente du comportement des autres auteurs mystificateurs. D'habitude, l'auteur d'un canular attend le moment de gloire pour se dévoiler au public, afin de mieux se venger de ses adversaires. Gary s'est privé de ce bonheur, en désignant son neveu Pavlowitch comme Ajar. De son vivant, il n'a jamais voulu se faire connaître comme Ajar.

Quérard qualifie d'hétéronyme les pseudonymes tels que Clara Gazul ou Émile Ajar. Dans ce cas, l'auteur fait croire en l'existence d'un écrivain purement imaginaire et donne comme authentiques des livres écrits sous pseudonyme. Ceci s'applique à presque tous les auteurs qui sont mentionnés dans ce chapitre. Ce qui rend Romain Gary encore plus différent d'autres auteurs hétéronymes, c'est, qu'en créant le pseudonyme d'Émile Ajar, Gary lui donne non seulement une vie (fictive bien sûr), mais il lui donne un corps et

un visage. Il crée une aventure sans précédent dans le monde de la mystification. Cette étape ne se produit jamais chez les autres mystificateurs: Émile Ajar se distingue ainsi de ces derniers.

L'auteur hétéronyme peut même présenter sa vie et ses oeuvres par le biais d'une autobiographie (fictive) ou d'interviews journalistiques. Sous la tutelle de Romain Gary, Paul Pavlowitch accorde plusieurs interviews aux médias en se faisant passer pour Ajar. Ainsi à l'encontre de Clara Gazul par exemple, la mystification Émile Ajar ne consiste pas seulement à créer une personnalité d'écrivain fictive (la notice biographique sur *Clara Gazul*), mais à conférer le statut d'auteur à un individu dont l'existence est attestée, c'est-à-dire, de donner corps à Ajar en la personne de Pavlowitch. Par souci de vraisemblance Gary affecte Ajar d'une personnalité différente de la sienne et lui donne une véritable vie et un corps. Gary est un gaulliste respectable et un Compagnon de la Légion d'Honneur tandis qu'Émile Ajar se présente comme un repris de justice.

La section 6.6 de ce chapitre est consacrée aux auteurs supposés. En ce qui concerne la mystification Émile Ajar, cette catégorie correspond à l'étape où Pavlowitch se présente aux lecteurs comme auteur authentique. L'interview que Paul Pavlowitch accorde à la journaliste Yvonne Baby à Copenhague et à d'autres journalistes permet à Pavlowitch, l'Ajar supposé, de présenter la vie et les oeuvres d'Ajar comme authentiques. Le cas d'Ajar présente ainsi une situation extrême de supposition d'auteur.

6.13. La Raison profonde de la mystification

Paul Lacroix postule que la mystification est une stratégie d'un auteur qui consiste à "mettre les rieurs de son côté" (3). Ceci semble être conforme à l'aveu posthume de Romain Gary dans son testament littéraire lorsqu'il déclare: "Je me suis bien amusé. Au revoir et merci" (*Vie et Mort* 43). Pour d'autres critiques comme Gérard Genette, le choix d'un pseudonyme peut influencer sur le succès d'une oeuvre (*Figures II* 107). En effet, lorsque Louis Farigole, agrégé de philosophie, prend le nom plus euphonique de Jules Romains, il connaît du succès dans le monde littéraire. Ceci s'applique aussi au pseudonyme Ajar.

François Caradec décrit ainsi la stratégie du mystificateur: "Le mystificateur s'introduit clandestinement dans le jeu des rites, des coutumes et des habitudes sociales de ses victimes. Celles-ci ne savent pas qu'elles jouent et qu'elles observent des règles abstraites" (108-9), ce qui décrit admirablement la stratégie de Romain Gary lorsqu'il crée le personnage Ajar, sans que personne ne soupçonne qu'Ajar est un personnage fictif.

Gaston Paris postule que les raisons suivantes encouragent un auteur à mystifier son public: "l'intérêt, la vanité, la religion et le singulier amour-propre" (Jeandillou 34). Lorsque Gary crée son double Émile Ajar, il n'est question ni d'intérêt, ni de religion. Par contre la vanité et l'orgueil l'ont sans aucun doute poussé à mystifier son public et les critiques. Voyslav Yovanowich, quant à lui, explique que la mystification découle souvent du "besoin d'écrire d'un fou ou d'une génie bizarre [ou de] l'amusement méchant

d'un esprit moqueur" (264). En fait Romain Gary avoue, dans son testament littéraire, qu'il s'est moqué de ses critiques et qu'il s'est vengé d'eux en leur faisant croire en l'existence d'Émile Ajar.

La mystification Émile Ajar n'a duré que cinq ans. Elle est courte mais intense. Pierre Bayard présente les étapes de cette mystification de la manière suivante. La première étape se retrouve dans toutes les mystifications littéraires, dont le principe est de mettre sur la couverture du livre un nom différent de celui de l'auteur. Cette étape se produit aussi chez Mérimée (Clara Gazul), Henri Beyle (Stendhal), et plusieurs autres auteurs.

Romain Gary la suit trois fois: la première fois en 1958 lorsqu'il publie *L'Homme à la colombe* du nommé Fosco Sinibaldi, la deuxième fois, peu de temps avant l'affaire Ajar, lorsqu'il publie des *Têtes de Stéphanie* sous le pseudonyme de Shotan Bogat, et, la troisième fois avec *Gros-Câlin* lorsque le nom Émile Ajar paraît sur la couverture du roman.

La deuxième étape de cette mystification est celle de l'innovation. Ceci se produit lorsque les auteurs, tels que Mérimée et Henri Beyle, créent des biographies fictives pour leurs pseudonymes. Romain Gary suit cette étape trois fois, la première fois en 1958, lorsqu'il invente la fausse biographie de Fosco Sinibaldi, et la deuxième fois en 1974, lorsqu'il invente la biographie du soi-disant Libanais Shotan Bogat. Mais c'est sa troisième tentative de pseudonyme qui est la plus remarquable. Lors de la production de

La Vie devant soi, Romain Gary demande à Paul Pavlowitch de se faire passer pour Ajar, de lui prêter seulement son corps et non pas sa voix. Il est à remarquer que cette étape ne se produit presque jamais chez les autres mystificateurs. C'est aussi à cette époque de la mystification Émile Ajar que Paul Pavlowitch se présente comme Ajar auprès des médias; il joue impeccablement ce rôle fictif et ne se contente plus d'être une marionnette de Romain Gary.

À la troisième étape de cette mystification, Paul Pavlowitch devient Émile Ajar et fournit à la presse une biographie fictive, contrairement aux désirs de son oncle. Il commence à vivre cette existence fictive d'Émile Ajar avec ses tics langagiers et ses manières de bohémien. Il devient Ajar jusqu'à la révélation de la supercherie.

Cette mystification est sans doute celle qui a le plus réussi dans l'histoire de la littérature moderne, même si Romain Gary a mis tant d'obstination à nier sa parenté avec Ajar. Gary seul a su créer un personnage qui entre dans la légende.

Chapitre 7

L'Analyse de style statistique

Le génie, à force de faire sortir l'homme de lui-même pour le faire entrer dans autrui, peut faire que l'artiste se perde un jour lui-même [...]. Il est des âmes hantées, comme des vieilles maisons, par les fantômes qu'elles ont trop longtemps abrités. (Guyau 105)

L'analyse statistique du style littéraire est une discipline relativement nouvelle. Mais elle jouit d'une grande popularité auprès des chercheurs contemporains. Ce chapitre fait un bilan de cette analyse en débutant par l'antiquité, ce qui de prime abord paraît contradictoire. Il sera néanmoins démontré sous peu que l'analyse statistique du langage remonte à plusieurs millénaires. La deuxième partie du chapitre décrit le genre de recherche qui se fait de nos jours dans ce domaine (tels que la compilation des dictionnaires de fréquences, des index et des concordances).

Un certain nombre de termes techniques est utilisé dans ce chapitre. Aussi, avant qu'on ne puisse entrer dans le vif du sujet, serait-il utile de définir la terminologie employée. Comme dans la langue courante le terme *mot* prend plusieurs significations, le terme *vocabulaire* est utilisé pour signifier le *mot* dans son sens le plus large et le plus général. À l'exemple d'Engwall, le terme *occurrence* est aussi utilisé; il désigne "une série de signes délimités par des blancs, des signes de ponctuation, des apostrophes ou des traits d'union, c'est-à-dire qu'elle est pratiquement identique à l'unité graphique" (*Fréquence*, 27). L'*occurrence* est donc le graphisme, unique ou non, paraissant à tel endroit dans un texte; il est possible de compter les graphismes après avoir coupé un texte en unités graphiques, sans autre traitement.

Les termes *type* et *lemme* sont utilisés aussi. Pour la définition à donner au mot *type*, on retiendra celle que propose Engwall: elle comprend "toutes les occurrences graphiquement identiques" (28). Le *type* représente souvent plusieurs occurrences dans le texte. Pour le compter, il est nécessaire de comparer chaque occurrence à toutes les autres afin d'identifier les formes uniques. Le terme *lemme* veut dire l'ensemble de toutes les formes flexionnelles que peut prendre un mot, par exemple, le verbe *devoir* peut avoir les formes flexionnelles *dois*, *dû*, *dues*, *devaient*, *doivent* et ainsi de suite. Le comptage des lemmes exige l'analyse morpho-syntaxique de toutes les occurrences.

La phrase suivante peut servir d'exemple aux notions déjà suggérées: "La femme est celle qui est venue de l'est". Cette phrase contient 10 occurrences (*la, femme, est, celle, qui, est, venue, de, l', est*); 8 types (*la, femme, est, celle, qui, venue, de, l') est* apparaissant trois fois; elle contient aussi 8 lemmes qui ne sont pas identiques aux types: *le* (article défini), *femme* et *est* (substantifs), *être, venir* (verbes), *celle* et *qui* (pronoms), *de* (préposition).

7.1. Les Précurseurs

L'analyse statistique de la langue remonte à plus de trois mille ans à l'ère védique des Hindous. À cette époque l'Inde antique connaît l'éclosion des premiers linguistes de la langue sanskrite. C'est durant cette période que le grammairien Panini codifie le livre de grammaire *L'Astadhahyi* (les huit livres). Ce livre est constitué de quatre mille "sûtras" ou aphorismes, dont l'enchaînement et la formulation font un ensemble d'une rigueur algébrique. De plus, Panini constate qu'il existe un aspect mathématique dans la chaîne du discours.

Durant cet ère des Sûtra (entre 500 et 200 ans avant notre ère) des grammairiens du sanskrit effectuent des dénombrements lexicaux sur d'autres oeuvres. Ils codifient dans le *Anuvakana-kramani* le nombre de chants, de vers et de syllabes contenus dans le grand poème religieux des Hindous le *Rig-Veda*. Ils trouvent que cette oeuvre contient 1.017 chants, 10.580,5 strophes, 153.826 occurrences et 432.000 syllabes (Francis 39).

Dans l'ancienne Alexandrie, entre 200 et 100 ans avant notre ère, les savants analysent des chefs-d'oeuvre grecs tels que *L'Iliade* et *L'Odyssée*, et comptent systématiquement le nombre de fois que les auteurs utilisent des *aiekechretai* (les occurrences les plus souvent utilisées), des *spanios legomena* (les occurrences les plus rarement utilisées) et des *hapax legomena* (les occurrences utilisées une seule fois) dans le corpus.

Au dixième siècle, les Masorètes (dont un membre éminent est le célèbre littéraire Aaron Ben-Asher) déterminent le nombre de vers, d'occurrences, et de lettres contenus dans la Bible des Hébreux. Au treizième siècle, les Dominicains compilent la première concordance de la Bible latine, et au quinzième siècle, Robert Estienne (1499-1594) imprime la première concordance de La Bible latine, hébraïque et grecque.

Plusieurs siècles plus tard, en 1711, dans un domaine plus prosaïque, Jean Chardin, après avoir visité deux fois la Perse, l'Inde et le Caucase, explique dans son *Journal de voyage* que la langue arabe contient 12.350.142 vocables dont 1.000 termes pour *chameau*, 500 pour *lion* et un très grand nombre pour *palmier* (Dugast 11).

Ces exemples montrent que, depuis très longtemps, l'être humain s'intéresse au dénombrement de vocables dans un corpus, même si quelques fois le comptage est effectué de manière plutôt arbitraire ou "impressionniste" comme dans le cas de Chardin.

Le comptage des vocables dans un corpus est repris vers la fin du dix-neuvième siècle, (1897) lorsque l'allemand F. W. Kading dans le livre intitulé

Häufigkeitswörterbuch der deutschen Sprache, destiné à la sténographie rationnelle, effectue un dénombrement exhaustif des mots contenus dans un texte allemand. Dans ce corpus de 10.910.777 occurrences, Käding relève 258.200 types et 11.173 lemmes (Dugast 8).

Au début de ce siècle, dans le but de faciliter l'enseignement de la sténographie, Jean-Baptiste Estoup dénombre toutes les occurrences rencontrées dans un corpus français. Il découpe le corpus en trente tranches de mille vocables chacune, compte le nombre de types différents et nouveaux rencontrés dans chaque tranche successive. Il s'aperçoit ainsi que le type le plus fréquent est l'article défini (*le, l', la, les*, cumulés) qui apparaît 1.949 fois dans ce corpus de 30.000 occurrences et que 922 types sont des *hapax legomena*. Après avoir fait quelques calculs, Estoup note que pour un total de 60.000 occurrences, on ne doit compter que 3.500 types différents (Dugast 9).

Le but principal du travail d'Estoup est de fournir aux professeurs de sténographie des gammes basées sur la fréquence des occurrences. Mais il montre aussi le nombre de fois que sont répétés les vocables les plus fréquents, et à quel point la chaîne du discours comprend des types différents.

Le travail d'Estoup est néanmoins criblé d'erreurs. Il regroupe en un même type les diverses occurrences de *qui* et *que* (pronom relatif et interrogatif sujet et complément); de *et, est* et *si* (conjonction de coordination, verbe *être* conjugué à la troisième personne du singulier et conjonction); de *ce* et *se* (adjectif démonstratif et pronom personnel

réfléchi à la troisième personne); de *son* et *sont* (adjectif possessif à la troisième personne et verbe être à la troisième personne du pluriel) et ainsi de suite. Son livre est destiné aux sténographes et non pas aux linguistes ou aux grammairiens. Malgré les problèmes de regroupement dans le travail de Jean-Baptiste Estoup, il reste le premier à montrer les deux axes de l'organisation du vocabulaire dans le discours: l'axe horizontal ou l'axe des abscisses (l'accroissement du vocabulaire au long du discours) et l'axe vertical ou l'axe des ordonnées (la distribution des fréquences des occurrences pour une certaine longueur de discours).

Plus tard, G. K. Zipf développe les idées d'Estoup sur la fréquence des occurrences. Il classe toutes les occurrences dans un corpus par ordre décroissant des fréquences, ce qui veut dire que l'occurrence qui apparaît le plus souvent dans le corpus occupe le premier rang; l'occurrence qui est la plus fréquente lorsqu'on fait abstraction de la première est au deuxième rang et ainsi de suite.

Selon Zipf, de l'occurrence la plus fréquente à l'occurrence la plus rare, une distribution croissante du nombre d'occurrences est équilibrée par la fréquence elle-même. Selon la loi de Zipf, dans un corpus, la fréquence d'une occurrence est inversement proportionnelle à son rang, et le produit de la fréquence d'une occurrence multipliée par son rang est plus ou moins constant.

Il existe des cas où cette formule s'applique admirablement; le cas cité le plus souvent est celui offert par Zipf; il se rapporte au vocabulaire de l'*Ulysse* de James

Joyce. La table 7.1 démontre que le produit du rang et de la fréquence (indiqué dans la colonne R x F) est presque toujours le même, variant entre 25.000 et 26.530, qu'il s'agisse de la dixième ou de la cinq millième occurrence.

Table 7.1

***L'Ulysses* de James Joyce**

<i>Rang</i>	<i>Fréquence</i>	<i>RxF</i>
10e occurrence	2.653,00	26.530
20e occurrence	1.311,00	26.220
30e occurrence	883.33	26.500
1,000e occurrence	26,00	26.000
5,000e occurrence	5,00	25.000

Source: Zipf, *Human Behaviour* 134

Les experts de la littérature anglaise signalent cependant, que *l'Ulysses* de James Joyce n'est pas un roman ordinaire; il a révolutionné le roman anglais du vingtième siècle. Il est surtout reconnu pour son style peu conventionnel. Aussi le phénomène observé à la table 7.1 se réfère-t-il à un cas particulier et non pas aux cas généraux. Si la loi de Zipf ne dit rien sur la fréquence de telle ou telle occurrence, elle établit néanmoins que cette occurrence possède un certain rang. Par exemple, si on consulte la table se rapportant à *L'Étranger* (appendice 10.1.5), on observe que la préposition *sur* se situe au rang 43, tandis qu'elle occupe le rang 48 dans *Le Noeud de Vipères* (appendice 10.1.6), le rang 57 dans *La Porte étroite* (appendice 10.1.8), le rang 47 dans *Au-delà de*

cette limite votre ticket n'est plus valable (appendice 10.1.2) et ne figure même pas dans la liste des soixante types les plus fréquents dans *La Vie devant soi* (appendice 10.1.4).

À partir de ces observations, il est légitime de déduire que le type *sur* est plus fréquent dans *Au-delà de cette limite votre ticket n'est plus valable* que dans *La Porte étroite*. Effectivement *sur* apparaît 3,99 fois sur chaque mille mots dans *Au-delà de cette limite votre ticket n'est plus valable* tandis que dans *La Porte étroite*, *sur* apparaît 2,95 fois sur mille en moyenne. Il n'est pas *a priori* absurde qu'une relation existe entre le rang d'une occurrence et sa fréquence. Quant aux occurrences rares, il y a plusieurs types qui apparaissent le même nombre de fois chacun dans un texte, l'ordre dans lequel ces types sont rangés entre eux est alors indéterminé; il est permis de les ranger de façon tout à fait arbitraire.

La nature statistique de la langue n'est donc pas un concept hors de sens: en fait elle préoccupe les grammairiens, les linguistes et les pédagogues depuis longtemps et continue à jouer un rôle important dans le domaine de la linguistique et de la littérature.

La prochaine section aborde un autre aspect de l'analyse statistique de la langue, c'est-à-dire le genre de travail que les compilateurs des index, des concordances et des dictionnaires des fréquences entreprennent depuis le début du vingtième siècle dans le but de faciliter l'apprentissage de la langue.

7.2. Index, dictionnaires des fréquences, concordances etc.

Après les sténographes ce sont les pédagogues qui s'intéressent à la fréquence des occurrences dans un corpus. Au début du vingtième siècle, des pédagogues américains compilent des listes de vocables dans le but de faciliter la tâche des professeurs de langue.

Se basant sur plusieurs textes anglais, Thorndike compile une liste de 4.500.000 occurrences qu'il regroupe en 10.000 types et classe par fréquence décroissante. En 1921, il publie le résultat de ses recherches dans *The Teacher's Word Book*. Thorndike ne distingue pas entre les homonymes (les vocables de prononciation identique mais de sens différents) et entre les emplois nominaux ou verbaux des mots.

Peu de temps après la publication du livre de Thorndike, la Ligue classique américaine (*American Classical League*), sous la direction de V. C. A. Henmon, entreprend une étude de la langue française dont l'un des buts est de déterminer en quelle mesure une connaissance préalable du latin influence l'apprentissage du français. Se concentrant sur plusieurs textes hétérogènes constitués de la manière indiquée dans la table 7.2, Henmon calcule le nombre d'occurrences dans ces textes. Il obtient un corpus de 400.000 vocables qu'il classe en types décroissants.

En 1924, Henmon publie ses données dans *A French Word Book Based on a Count of 400,000 Running Words* et établit la première liste d'occurrences de la langue française aux États-Unis.

Table 7.2

Constitution du corpus de Henmon

Genre de texte	Nombre d'occurrences
Textes scolaires	80.000
Prose philosophique	20.000
Articles de revues	50.000
Articles de journaux	30.000
Textes d'écrivains français modernes	140.000
Textes de poésie	15.000
Essais et articles militaires et politiques	25.000
Lettres et télégrammes	15.000
Fiction populaire	25.000
Nombre total d'occurrences	400.000

Source: Juilland, *Frequency Dictionary of French Words* xviii

Ce travail a quelques faiblesses. D'après Juilland (*Frequency Dictionary of French Words*) les lemmes sont souvent regroupés arbitrairement: les locutions telles que *parce que*, *tandis que*, *quant à*, *chef-d'oeuvre*, *rez-de-chaussée* sont traitées comme un seul type chacune (*parce* et *que* ne sont pas traités séparément, par exemple). Les occurrences masculines et féminines telles que *amoureux* et *amoureuse* sont regroupées ensemble. Les homonymes *mort* (adjectif et substantif) *or* (substantif et

conjonction) sont aussi regroupés ensemble sans tenir compte du fait qu'il s'agit de deux vocables différents.

Les pronoms aussi sont regroupés de façon arbitraire. *Je, me, moi, tu, te et toi* sont mis ensemble, tandis que *nous* et *vous* sont traités séparément. Malgré ce problème de classement, Henmon indique la fréquence de tous les vocables: elle varie entre 27.749 (*le, la, l'* cumulés) et 5 que partagent 400 occurrences (Juilland xvii). Les occurrences sont aussi présentées deux fois; une fois par ordre décroissant des fréquences et une deuxième fois par ordre alphabétique.

Plusieurs années après la publication du livre de Henmon, des commissions d'enquête établies par des Canadiens et des Américains, sous la direction de Vander Beke, continuent le même travail. Pour aider Vander Beke dans cette entreprise de nombreux volontaires se concentrent sur des textes répartis sur plusieurs disciplines et publiés au 19^e et au 20^e siècles (voir table 7.3).

Ils obtiennent une liste de 1.147.748 occurrences réparties en 19.253 types et calculent leur fréquence. Vander Beke classe les pronoms d'après le système de Henmon, mais regroupe ensemble les substantifs tels que *chat* et *chatte*, *lecteur* et *lectrice*. Il regroupe aussi en un seul type des mots composés tels que *bien-être*, *chef-d'oeuvre* et *là-bas*. Par contre, les homographes tels que *livre* (masculin et féminin) sont indiqués par genre tandis que *celui* et *ceux* sont traités séparément. Vander Beke

distingue aussi entre les homographes tels que *devoir* (verbe et substantif), *livre* (masculin et féminin).

Table 7.3

Constitution du corpus de Vander Beke

Genre de textes	Nombre de passages
Fiction	33
Pièces de théâtre	13
Textes scientifiques, philosophiques et religieux	09
Textes d'histoire, de biographie, de civilisation et de critique	16
Textes provenant des journaux, des revues et des périodiques	14

Source: Juilland, *Frequency Dictionary of French Words* xx

Vander Beke apporte une amélioration sensible au travail de Henmon. L'intérêt du travail entrepris par G. E. Vander Beke et E. L. Thorndike est qu'il démontre que les mots grammaticaux (les mots-outils) sont les mots les plus fréquents dans un corpus.

Peu de temps après le travail de Vander Beke, Frédéric D. Cheydleur entreprend une étude de la langue française en se concentrant sur des idiotismes français utilisés à la fin du 19^e siècle et au début du 20^e siècle. Dans un premier temps, trois universitaires compilent une liste de 2.700 idiotismes tirés uniquement d'oeuvres provenant des sources indiquées dans la table 7.4.

Table 7.4

Constitution du corpus de Cheydleur

Genre de textes	Nombre de passages
Romans et contes	30
Pièces de théâtre	23
Journaux	13
Textes d'histoire, de biographie, de critique et de civilisation	13
Textes scientifiques, philosophiques et religieux	08

Source: Juilland, *Frequency Dictionary of French Words* xxi

Durant la deuxième phase de l'entreprise, cent vingt enseignants comptent le nombre de vocables figurant dans ces textes et en compilent une liste de 1.183.000. En 1929, Cheydleur publie ses recherches dans *A French Idiom List Based on a Running Count of 1,183,000 Words*; livre destiné surtout aux étudiants de linguistique et de pédagogie (Juilland xxi).

D'après Juilland la grande faiblesse du livre de Cheydleur est qu'il traite les locutions figées telles que *avoir l'intention de*, *à la fin*, *être pressé* etc. comme des idiotismes (xxi).

7.3. L'Élaboration du français fondamental et *Le Français fondamental*

Malgré ce qui vient d'être démontré, l'analyse statistique du vocabulaire français ne débute véritablement que durant les années cinquante, lorsque le Ministère de

L'Éducation en France se propose d'établir un programme de français de base qui puisse faciliter l'apprentissage de la langue française pour les Français d'outre-mer, les immigrants et les étrangers qui vivent en France et manient mal la langue. G. Gougenheim se donne alors la tâche de compiler un tel manuel. Il se concentre sur la langue parlée.

Dans la première phase de cette entreprise, Gougenheim choisit un groupe de 163 locuteurs représentant la francophonie dans son ensemble. Ces locuteurs viennent de diverses régions de la France, du Canada, de la Suisse, de l'Algérie et de la Guyane française et de milieux sociaux différents. Ensuite, Gougenheim leur pose des questions sur des sujets très variés tels que le travail, la famille, les amis, le voyage, la santé, la maladie, le transport, la vie quotidienne, la littérature, l'art, le spectacle, le sport, l'habitation, le ménage, la vie estudiantine et la nourriture. Il enregistre leurs réponses sur bande magnétique. À partir de ces conversations, Gougenheim compile une liste de 312.000 occurrences qu'il regroupe en 8.000 lemmes. La table 7.5. contient les trente-huit lemmes les plus utilisés.

En 1956, Gougenheim publie *L'Élaboration du français fondamental. Étude sur l'établissement d'un vocabulaire et d'une grammaire de base (L'Élaboration du français fondamental)*. Ce livre démontre que le vocabulaire de base est constitué d'environ 40% de substantifs, 25% de mots grammaticaux, 22% de verbes et 11% d'adjectifs.

Table 7.5.

Liste des Lemmes par fréquence décroissante

<i>Rang</i>	<i>Lemmes</i>	<i>Fréquence</i>	<i>Rang</i>	<i>Lemmes</i>	<i>Fréquence</i>
1	être	14.083	20	qui (relatif)	3.096
2	avoir	11.552	21	oui	2.935
3	de	10.503	22	alors	2.854
4	je	7.905	23	une (article)	2.780
5	il (s)	7.515	24	mais	2.768
6	ce (pronom)	6.846	25	des (art. ind.)	2.646
7	la (article)	5.374	26	elle(s)	.462
8	pas (négation)	5.308	27	en (préposition)	2.405
9	à (préposition)	5.236	28	dire	2.391
10	et	5.082	29	y	2.391
11	le (article)	4.957	30	pour	.076
12	on	4.266	31	dans	2.066
13	vous	4.202	32	me	2.014
14	un (article)	4.188	33	se	1.993
15	ça (pronom)	3.972	34	aller	1.876
16	les (article)	3.815	35	bien (adv.)	1.697
17	que (conj.)	3.537	36	du	1.658
18	ne	3.283	37	tu	1.536
19	faire	3.174	38	en (pro-adv)	1.501

Source: Gougenheim 69-70

Trente-deux de ces trente-huit lemmes, soit 44%, sont des mots-outils ou des vocables "vides" (articles, prépositions, conjonctions et ainsi de suite). Dans cette recherche, 2.700 des lemmes sont des *hapax legomena*, 1.170 en sont des *hapax dislogomena* (des vocables qui n'apparaissent que deux fois dans un corpus) et 694 apparaissent trois fois. Ceci veut dire que plus de 50% des lemmes ne sont pas utilisés

plus de trois fois. Les trente-huit lemmes les plus fréquents constituent 50% du corpus, indiquant qu'un petit nombre de vocables reviennent constamment dans la langue.

En 1954 Pierre Guiraud, en collaboration avec Joshua Whatmouth, publie la *Bibliographie Critique de la statistique linguistique*. Ce livre qui contient environ 2.500 titres identifie des concordances et des index pour différentes langues.

Dans la plupart des cas mentionnés jusqu'ici, le dépouillement des textes a été fait à la main, sans l'aide de machine électronique (sauf dans le cas de Juilland). Aujourd'hui les compilateurs d'index, de dictionnaires et de concordances font le comptage de vocables avec l'aide précieuse de l'ordinateur et peuvent entrer dans un domaine jusqu'ici inaccessible.

Grâce à l'ordinateur, on parvient aujourd'hui à compiler des dictionnaires de fréquences dans plusieurs langues. Alphonse Juilland en a fourni pour l'espagnol, le français, l'italien et le roumain (Voir *Frequency Dictionary of French Words*, *Frequency Dictionary of Italian Words*, *Frequency Dictionary of Spanish Words* et *Frequency Dictionary of Rumanian Words*). D'autres chercheurs entreprennent des travaux sur la fréquence des vocables dans un corpus. L'exemple le plus connu à propos du français est sans doute le *Trésor de la langue française*.

7.4. *Le Trésor de la langue française*

Durant les années cinquante une controverse règne autour de la destinée du dictionnaire *Littré*; il est question de savoir s'il faut remanier ce dictionnaire ou le réimprimer. Les experts choisissent de le laisser intact et de le réimprimer tel quel, étant donné qu'il représente le travail d'un homme et d'une époque.

En novembre 1957, lors d'une réunion des experts de la langue française à un colloque international qui a lieu au Centre de Philologie Romane de la Faculté des Lettres (Université de Strasbourg), des lexicographes français envisagent la création d'un nouveau dictionnaire de la langue française, aux dimensions synchroniques et diachroniques, fait à partir d'un univers de textes. Le but principal de ce nouveau dictionnaire historique est de donner un tableau aussi complet que possible du vocabulaire français allant depuis la Révolution française (1789) jusqu'aux temps modernes.

Le comité décide de le faire à partir du recensement informatique des textes (surtout littéraires), publiés pendant cette période. Ils donnent alors le nom de *Trésor de la langue française (TLF)* à ce projet. Avec l'appui budgétaire du CNRS (Conseil National de la Recherche Scientifique), le Centre de recherche pour un *Trésor de la langue française* à Nancy commence la mise en place du dictionnaire sous la direction de Paul Imbs.

Le comité exécutif choisit une documentation d'exemples constitués de 100 millions d'occurrences et provenant d'oeuvres littéraires et techniques. En ce qui

concerne la sélection des textes, le comité exécutif décide de représenter chaque période de l'histoire par des oeuvres diverses réparties dans chaque genre littéraire (roman, poésie etc.) bien diffusées et bien reçues à époque de leur publication. Pour accomplir cette tâche, les chercheurs consultent d'abord les ouvrages de l'histoire de la littérature (y compris une étude de Gustave Lanson). Un autre groupe d'universitaires vérifient cette liste et y apportent les modifications qui s'imposent. Après d'autres modifications qu'apporte le comité exécutif, la liste est enfin approuvée.

En décembre 1960 un laboratoire est créé, et la perforation des textes commence en mai 1963. Les traitements informatiques sur un ordinateur "Bull Gamma 60" sont entrepris à partir d'octobre 1964, mais la banque de données n'est pas encore accessible au public. En octobre 1971 le premier tome du *Trésor de la langue française (T.L.F.)* est publié sous la direction de Paul Imbs. La table 7.6 montre les textes dans le *Trésor de la langue française* par période et par genre.

Le dictionnaire s'adresse surtout à une élite et vise en premier lieu "l'universitaire de haute culture confronté avec les textes généralement littéraires" (Imbs 6), aussi bien que le public francophone cultivé et à la recherche d'un ouvrage de référence lui garantissant "l'exacte interprétation de ce qu'il entend ou lit" (Imbs 6).

Table 7.6.

Textes et Fréquences dans la banque de données du *Trésor de la langue française*

Période	Romans	Théâtre	Poésie	Prose non-littéraire	Nombre total de textes	Nombre total d'occurrences
1789-1815	11	8	10	46	75	6.268.717
1816-1832	19	5	10	39	73	5.279.621
1833-1841	25	32	18	26	101	5.172.959
1842-1849	15	2	2	21	40	4.376.412
1850-1859	20	5	11	22	58	4.440.320
1860-1869	18	8	11	20	57	4.756.318
1870-1879	22	7	12	15	56	4.136.780
1880-1892	37	6	8	13	64	5.057.974
1893-1907	31	18	11	33	93	5.050.345
1908-1918	35	12	17	18	82	4.725.829
1919-1926	39	12	8	24	83	5.146.014
1927-1932	34	14	0	24	72	4.076.780
1933-1937	30	7	1	15	53	4.497.834
1938-1945	27	22	5	33	87	5.085.310
1946-1964	16	25	5	39	85	5.659.375
Totaux	379	183	129	388	1.079	73.730.588

Source: La Banque de données *ARTFL* <URL: <http://humanities.uchicago.edu/ARTFL/ARTFL.html>>

Durant les années soixante-dix, des chercheurs de l'Université de Chicago créent la banque de données *ATFL* (American and French Research on the Treasury of the French Language) pour rendre le corpus du *Trésor de la langue française* accessible aux chercheurs nord-américains. Aujourd'hui le corpus *ARTFL* est constitué d'environ 2.000 textes comprenant des oeuvres littéraires et de prose non-littéraire.

Cette banque de données qui est l'une des plus grandes au monde fournit les listes de fréquences pour *L'Étranger*, *Le Noeud de Vipères*, *La Porte étroite* et *L'Immoraliste*.

7.5. Gunnel Engwall: *Fréquence et distribution du vocabulaire dans un choix de romans français*

En 1974, Gunnel Engwall publie une liste de fréquences et de distribution des occurrences dans certains romans de langue française. Pour la constitution de son corpus, Engwall se concentre sur des *best-sellers* publiés en France entre les années 1962 et 1968 puisque cette période "englobe à la fois le milieu et la majeure partie des années 1960 qui sont [...] bien représentées" (16). Elle choisit ces *best-sellers* selon les chiffres de ventes publiées dans les périodiques *Les Nouvelles littéraires* et *Le Figaro littéraire*.

À la première étape de la sélection, elle établit une liste de quatre cent titres environ, où figurent entre autres des auteurs tels que Simone de Beauvoir, (*Les Belles Images*, 1966), Marguerite Duras (*Le Vice-consul*, 1966), Françoise Sagan (*La Chamade*, 1965) et Romain Gary (*Lady L*, 1963) (Engwall 111-6).

À la deuxième étape du processus, Engwall établit les critères qui suivent: premièrement l'auteur doit être né en France; deuxièmement l'action du roman doit se dérouler en France à l'époque contemporaine; troisièmement l'auteur doit être parmi les vingt-cinq auteurs les plus jeunes parmi ceux dont au moins un roman est retenu à la

première et la deuxième phase de la sélection. Finalement chaque auteur ne doit être représenté qu'une seule fois. Des auteurs tels que Gary et Duras sont ainsi éliminés à la première étape de la sélection, compte tenu du fait que Gary est né à Moscou et Duras en Indochine. Après l'application de tous les critères de sélection énumérés ci-dessus, le corpus contient les vingt-cinq romans suivants qui sont présentés selon l'ordre de naissance de l'auteur, en commençant par l'auteur le plus jeune:

J.-M. G. Le Clézio, *Le Déluge* (1966); G. Perec, *Les Choses* (1965); F. Sagan, *La Chamade* (1965); C. Etcherelli, *Élise ou la vraie vie* (1967); R.-V. Pilhes, *La Rhubarbe* (1965); Y. Berger, *Le Sud*, (1962); S. Japriso, *La Dame dans l'auto avec des lunettes et un fusil* (1966); C. Aubry, *Sébastien parmi les hommes* (1968); B. Poirot-Delpech, *L'envers de l'eau* (1963); F. Nourissier, *Le Maître de maison* (1968); M. Bataille, *L'arbre de Noël* (1967); C. Paysan, *Les Feux de la Chandeleur* (1966); J.-P. Chabrol, *Je t'aimerai sans vergogne* (1967); M. Droit, *Le Retour* (1964); B. et F. Groult, *Il était deux fois* (1968); J. Husson, *Le Cheval d'Herbeleau* (1965); I. Monési, *Nature morte devant la fenêtre* (1966); J.-P. Ollivier, *Le temps des filles* (1965); R. Sabatier, *Le Chinois d'Afrique* (1966); J. Cabanis, *Les Jeux de la nuit* (1964); C. de Rivoyre, *Les Sultans* (1964); J. Dutourd, *Pluche ou l'amour de l'art*, (1967); P. Moinot, *Le Sable vif* (1963); H.-F. Rey, *Les Chevaux masqués* (1965); R. Vrigny, *La Nuit des Mougins* (1963) (Voir Engwall, xii-xix).

À l'exemple de Juilland, Engwall limite le nombre d'occurrences dans son corpus à 500.000 (*Frequency Dictionary of French Words*). Comme Juilland, elle considère que ce nombre est assez grand pour être représentatif et assez petit pour "être traité par ordinateur sans frais exagérés" (15). Son corpus doit refléter "un vocabulaire simple et constant" (14). Pour obtenir ces 500.000 occurrences, elle extrait 20.000 occurrences de chaque roman "pour éviter à priori qu'un roman soit éliminé faute de texte" (15) et pour donner à chacune la même importance afin de faciliter la comparaison entre eux.

Plusieurs méthodes sont possibles pour obtenir ces occurrences. Un cas extrême serait d'extraire les 20.000 occurrences chacune séparément, mais ceci ne permettrait pas de les considérer dans leur contexte. Une autre méthode serait de tirer au sort une occurrence et de relever les 19.999 autres après la première, mais cette méthode risque de privilégier une seule partie du roman sur les autres.

Pour écarter de telles éventualités, Engwall divise chaque roman en dix sections égales et tire environ deux mille occurrences de chaque partie. Puis, elle tire au hasard un nombre qui détermine la page d'entrée pour chacune des dix sections et pour chaque roman. Chaque occurrence a ainsi une chance égale d'apparaître dans l'échantillon. Si elle arrive à la fin d'une section n'ayant pas encore extrait les deux mille occurrences, elle reprend le comptage au début de cette même section jusqu'à ce que le nombre voulu d'occurrences soit atteint.

Engwall établit une liste des types les plus fréquents dans les romans. Le type le plus fréquent est *de* (fréquence absolue de 18.215 et fréquence relative de 36,43 pour chaque mille vocables), suivi de *la*. La table 7.7 indique les dix premiers types dans le corpus d'Engwall.

Table 7.7

Les Dix Premiers Types dans le corpus d'Engwall

Rang	Type	Nombre d'occurrences	Fréquence relative par tranche de 1.000 vocable
1	de	18.215	36,43
2	la	11.677	23,35
3	le	10.139	20,28
4	et	9.916	19,83
5	il	9.317	18,63
6	à	9.299	18,60
7	les	8.086	16,17
8	je	7.579	15,16
9	l'	7.483	14,97
10	un	6.965	13,93

Source: Engwall 130

La liste des fréquences établie par Engwall sert d'échantillon à la langue contemporaine dans cette analyse du style de Gary/Ajar. Toutes les listes des fréquences mentionnées dans ce chapitre montrent que des constantes existent dans la langue, c'est-à-dire, qu'il suffit d'à peine une trentaine de types pour couvrir presque la moitié d'un texte. Ces listes des fréquences constituent en quelque sorte des dictionnaires définitifs sur un aspect de la langue et jouent un rôle très important dans le domaine de l'analyse statistique du vocabulaire.

Les ouvrages entrepris par Thorndike, Hemnon, Vander Beke, Cheydfleur et Gougenheim démontrent l'importance des dictionnaires de fréquences dans l'apprentissage d'une langue première ou seconde. Mais un autre aspect de leur travail souligne l'importance des mots-outils dans la langue. Il demeure cependant que ces dictionnaires peuvent aussi jouer un rôle important dans le domaine de l'attribution d'auteur aux oeuvres anonymes.

Chapitre 8

Attribution d'auteurs aux oeuvres anonymes et pseudonymes

Rien n'est négligeable dans une oeuvre. C'est peut-être dans la zone souterraine de sa conscience linguistique, et dans les ramifications enfouies de ses habitudes, qu'un auteur cache les traits originaux de son style. (Brunet, 399)

Ce chapitre décrit les méthodes d'investigation employées couramment dans le domaine de l'attribution d'auteurs aux oeuvres anonymes et pseudonymes; il prend soin de mentionner celles qui sont presque marginales. La dernière partie du chapitre est consacrée à une justification de la méthodologie employée dans cette étude.

8.1. La Méthode traditionnelle

La critique d'attribution d'auteur existe depuis longtemps. En France, le modèle le plus connu de cette critique est celui effectué par F. Deloffre en 1970 dans

son oeuvre *Stylistique et poétique françaises* où il examine l'oeuvre de Robert Challes, (un auteur du dix-huitième siècle). La méthode de Deloffre peut être résumée de la façon suivante: après avoir lu attentivement le texte de manière à devenir sensible aux détails formels, les critiques repèrent les thèmes, les motifs, le jeu des personnages, la situation psychologique et les moments dramatiques. Les traits stylistiques utilisés avec un ou plusieurs de ces éléments permettent de déclencher le processus d'identification.

Lorsque les critiques ont repéré plusieurs tics langagiers, ils examinent d'autres ouvrages écrits à la même époque afin de déterminer si le même schéma linguistique y existe. S'ils relèvent ces détails, ils vérifient systématiquement leur apparition. Ainsi, à partir de ces schémas linguistiques, ils pourront attribuer à un autre auteur des oeuvres restées anonymes jusqu'alors.

La méthodologie de Deloffre exige beaucoup de temps à mettre en pratique et demeure subjective et impressionniste. Il existe d'autres méthodes plus objectives et plus efficaces pour attribuer des auteurs aux oeuvres anonymes et pseudonymes, mais le problème est que ces méthodologies ne sont pas uniformes et les points de vue divergent tellement qu'il est difficile de savoir quel système choisir. La section qui suit fait un résumé et un bilan de ces différentes approches, en explique leur "splendeurs" et leur "misères", et souligne celles qui sont les plus prometteuses.

8.2. La Longueur des vocables

En 1851, Augustus de Morgan postule que chaque auteur a tendance à utiliser des vocables de même longueur dans ses écrits, quel que soit le sujet ou la matière traitée. Par contre la longueur des vocables varie d'un auteur à l'autre: "I should expect to find that one man writing on two different subjects agrees more nearly with himself than two different men writing on the same subject" (Lord, 272). De Morgan suggère alors que la longueur des vocables constitue un indice du style d'un auteur.

Vers la fin du dix-neuvième siècle, T. C. Mendenhall applique l'hypothèse formulée par de Morgan, dans l'article "The Characteristic Curves of Composition"; il calcule la fréquence des vocables à une lettre, à deux lettres, à trois lettres etc. dans des oeuvres de Charles Dickens et de William Thackeray, afin de distinguer entre leur style. Quelques années plus tard, Mendenhall se sert de la même méthodologie pour résoudre le problème de la paternité d'une pièce de Shakespeare quelquefois attribuée à Bacon ("A Mechanical Solution of a Literary Problem").

Pour faire cela, Mendenhall se concentre sur les pièces dramatiques de Bacon, de Shakespeare, de Johnson et de Marlowe. Deux assistantes compilent ensuite des listes de vocables provenant des oeuvres de ces quatre dramaturges et obtiennent une liste de deux milliards de vocables. Ensuite, elles calculent le nombre de vocables à deux lettres, à trois lettres, à quatre lettres etc. dans l'oeuvre de chaque auteur.

Lorsque Mendenhall examine ces données, il s'aperçoit que Shakespeare et Marlowe ont tendance à favoriser les vocables à quatre lettres, c'est-à-dire qu'ils ont presque la même distribution de lettres par vocable, et que Bacon et Johnson utilisent plus de vocables à trois lettres. Il attribue alors à Shakespeare la paternité d'une pièce généralement attribuée à Bacon, puisqu'elle contient plusieurs vocables à quatre lettres.

Mendenhall distingue ainsi entre le style de Bacon et de Shakespeare, étant donné la prépondérance des vocables de quatre lettres chez Shakespeare et de trois lettres chez Bacon. Mais sa méthodologie ne lui permet pas de distinguer entre le style de Shakespeare et de Marlowe, compte tenu du fait que chacun de ces deux dramaturges tend à utiliser des vocables de quatre lettres.

En 1963, Claude Brinegar utilise la méthodologie de Mendenhall afin de vérifier si les *Quintus Curtius Snodgrass Letters*, qui sont généralement attribuées à Mark Twain, pourraient être de la plume d'un autre auteur. Brinegar se concentre premièrement sur plusieurs autres oeuvres de Mark Twain, calcule le nombre de vocables qui s'y trouvent, et en obtient un groupe témoin de 11.000 vocables. Ensuite, il regroupe les types d'après leur longueur dans ce groupe témoin.

Puis Brinegar regroupe les types d'après leur longueur dans les *Quintus Curtius Snodgrass Letters*, et, à l'aide de la statistique du χ^2 , compare la fréquence des types dans le groupe témoin à la fréquence des types dans les *Quintus Curtius Snodgrass Letters*. Il

obtient alors un χ^2 de 127,2 à douze degrés de liberté. Cette forte valeur du χ^2 semble indiquer que Mark Twain n'est pas l'auteur des lettres *Quintus Curtius Snodgrass*.

Afin de s'assurer de l'exactitude de ces résultats Brinegar compare d'autres oeuvres connues de Mark Twain au groupe témoin, notamment, *Roughing it* (3.200 vocables) et *Following the Equator* (2.750 vocables). Il regroupe les types d'après leur longueur dans ces deux oeuvres et les compare au groupe témoin. Cette fois-ci, il obtient des valeurs non significatives de 12.4 et 8.5, à douze degrés de liberté. Ces résultats du χ^2 démontrent que Mark Twain (l'auteur du groupe témoin) est effectivement l'auteur de *Roughing it* et de *Following the Equator*. Compte tenu de la haute valeur du χ^2 obtenue pour les *Quintus Curtius Snodgrass Letters*, il semblerait que Mark Twain n'en est pas l'auteur.

Il existe une autre étude d'attribution d'auteur qui est aujourd'hui très célèbre. Il s'agit de l'analyse que Frederick Mosteller et David Wallace ont entreprise en 1964 sur la paternité des *Federalist Papers*. Ces lettres qui sont au nombre de soixante-dix-sept, ont été publiées sous le pseudonyme "Publius" dans cinq journaux new-yorkais entre 1787 et 1788. Les auteurs de ces articles voulaient persuader les habitants de New York d'approuver la ratification de la nouvelle constitution des États-Unis. Mais la ville de New York a voté contre cette ratification. Les articles ont alors été publiés dans un livre portant la signature "Publius". Le livre contenait les soixante-dix-sept articles parus dans les journaux aussi bien que huit articles nouveaux.

L'identité de "Publius" n'était connue de personne. De temps en temps, le bruit courait que le général Alexandre Hamilton était l'un des auteurs. Lorsque celui-ci est mort en 1804, à la suite d'un duel, le contenu d'une lettre qu'il avait laissée à ses proches a été porté à la connaissance du public. Dans cette lettre, Hamilton avouait être l'un des auteurs de ces lettres, mais il désignait en même temps, comme coauteurs, John Jay et James Madison. Madison qui était alors président des États-Unis a gardé le silence.

En 1818, lorsque James Madison n'était plus Président, il a reconnu être l'auteur de douze de ces articles, à savoir les articles 49 à 58 et les articles 62 et 63. Hamilton s'était déjà présenté comme l'auteur de ces même articles. Ainsi Madison et Hamilton déclaraient tous deux avoir écrit ces douze articles.

Il a été établi depuis ce temps que John Jay est l'auteur des articles 2, 3, 4, 5, et 64. Tous les autres articles (sauf les douze en question) ont été attribués définitivement soit à Madison, soit à Hamilton. Pendant cent cinquante ans environ, la paternité de ces douze articles était arbitrairement attribuée à Hamilton ou à Madison, selon la popularité de l'un ou de l'autre à l'époque, jusqu'au moment où Mosteller et Wallace se sont intéressés à cette affaire.

D'abord, Mosteller et Wallace recherchent des indices dans le corpus qui puissent distinguer le style de Hamilton de celui de Madison. Par l'intermédiaire du test du χ^2 , ils comparent la longueur des vocables chez Hamilton et chez Madison et constatent que:

When Hamilton is compared with Madison, his average χ^2 is almost the same as that obtained when Madison is compared with himself. When Madison is compared with Hamilton, the average χ^2 is 10 units higher than for Hamilton versus Hamilton, but four of the 7 Madison papers have χ^2 that are smaller than the largest obtained from Hamilton - Hamilton comparisons. Hamilton's paper No. 72 agrees about as well with the Madison total as does any of Madison's papers. (*Inference 67*)

Il est impossible de différencier entre le style de Hamilton et de Madison en prenant la longueur des vocables comme un indice de style.

Depuis assez longtemps une controverse règne autour de la paternité des pièces de Shakespeare et donne lieu à beaucoup de conjectures. Plusieurs experts littéraires soutiennent que Shakespeare n'est pas l'auteur de son oeuvre, et postulent que celle-ci peut être attribuée à Bacon, à Fletcher, à Spencer ou à Marlowe. En 1955 Calvin Hoffman, un journaliste américain, affirme dans *The Murder of the Man Who Was Shakespeare* que Christopher Marlowe est en fait le véritable auteur de l'oeuvre de Shakespeare.

Plusieurs années après l'observation de Hoffman, M. W. A. Smith entreprend une étude du poème *Hero and Leander*, parfois attribué à Shakespeare et parfois à Marlowe, afin de l'attribuer définitivement soit à l'un soit à l'autre dramaturge. Il se concentre sur la longueur des vocables, et cela malgré l'échec de Mosteller et Wallace dans ce domaine.

Cette étude lui permet à peine de différencier entre le style de ces deux dramaturges: "results [...] were too scattered to permit any definite conclusions to be drawn" ("A Stylometric Analysis of *Hero and Leander*" 132).

Une année plus tard, Smith reprend son analyse du poème *Hero and Leander* en considérant encore une fois la longueur des vocables comme un indice du style. Cette fois-ci il divise le poème en six parties, calcule la longueur des vocables dans chaque partie et compare chaque partie les unes aux autres. Il constate alors que la longueur des vocables dans les deux premières parties du poème est considérablement différente de la longueur des vocables dans les quatre autres parties. Il déclare:

When works which are written of various literary genres, and/or were written during different eras (or in different languages), are compared, the differences observed are likely to exceed greatly any distinguishing characteristics which may reliably identify authors. Furthermore, when works in the same literary form by different contemporaneous authors are compared, their word frequency distributions may appear so similar that they seem to have been written by the same hand [...]. Mendenhall's method now appears to be so unreliable that any serious student of authorship should discard it. ("Recent Experience and New Developments of Methods for the Determination of Authorship" 73).

Smith démontre que la méthode de Mendenhall contient bien des pièges. Il est vrai qu'une telle méthodologie a permis à Mendenhall de distinguer entre l'oeuvre de Shakespeare et celle de Bacon, et à Brinegar de prouver que Mark Twain n'est pas l'auteur des lettres *Quintus Curtius Snodgrass*. Mais, Mosteller et Wallace aussi bien que Smith démontrent que la longueur des vocables ne sert pas toujours à identifier le style d'un auteur. La méthodologie proposée par Mendenhall a plusieurs faiblesses et elle est à déconseiller.

8.3. Le Nombre de syllabes par vocable

En 1952, W. Fucks, postulant sans doute que le nombre de syllabes contenu dans un vocable constitue un indice du style d'un auteur, calcule dans des textes de plusieurs auteurs, le nombre de syllabes par vocable, la fréquence relative des vocables à un certain nombre de syllabes, et leur distribution dans ces textes. Il s'aperçoit alors que le nombre de syllabes par vocable diffère selon la langue plutôt que selon les auteurs.

En 1974, Barron Brainerd adopte la méthodologie de Fucks et compare plusieurs textes anglais entre eux; il constate que le style de certains auteurs varie d'après les genres: "there may be gross changes in the overall distribution of syllable counts as style changes from one kind of writing (e.g. narrative) to another (e.g. conversation)" (*Weighing Evidence* 89). Il constate lui aussi que cette méthodologie distingue entre les genres plutôt qu'entre les auteurs.

Cette sorte d'analyse ne produit pas de résultats satisfaisants. Étant donné que les textes qui sont inclus dans cette étude sont tous des romans, l'analyse de la longueur des vocables en syllabes comme indice de style n'est d'aucune utilité.

8.4. La Longueur des phrases

En 1888, Sherman propose que la longueur des phrases constitue un élément essentiel du style, c'est-à-dire qu'elle varie d'un auteur à l'autre mais demeure la même chez un auteur particulier. Par ailleurs, Sherman affirme: "in the heavier sort of composition, as elaborate criticism, [...] it is noticeable that the sentence average rises" ("Some Observations" 124).

En 1938, Udny Yule vérifie l'hypothèse de Sherman en l'appliquant au texte *De Imitatione Christi*, une oeuvre anonyme qui a circulé dans les milieux religieux des Pays-Bas, peu après 1400. Elle était aussi importante que la Bible au seizième siècle et a influencé la Réforme catholique aussi bien que le mouvement protestant. Il existe un manuscrit de l'oeuvre, copié de la main de Thomas à Kempis et portant sa signature.

Plusieurs experts croient que Jean Charlier de Gerson est le véritable auteur de cette oeuvre. Pour trancher la question, Yule entreprend une analyse du texte en se concentrant sur la longueur des phrases. Pour commencer, il calcule la longueur des phrases chez plusieurs auteurs qui sont contemporains de Thomas à Kempis et de Jean Charlier de Gerson et découvre qu'elle est différente pour chaque auteur. Ensuite, il

calcule la longueur des phrases dans des textes écrits par Thomas à Kempis et par Jean Charlier de Gerson, et les compare à la longueur des phrases dans le texte *De Imitatione Christi*. Il constate alors que Thomas à Kempis est l'auteur de cette oeuvre.

La méthodologie adoptée par Yule est assez problématique: il n'explique pas s'il a basé son étude sur une édition ancienne du *De Imitatione Christi* (qui contient la ponctuation originale de l'oeuvre) ou s'il l'a basée sur une édition moderne de l'oeuvre. De nos jours, ce sont les éditeurs qui mettent la dernière main à la ponctuation d'un texte et déterminent la longueur des phrases. Si Yule a basé son étude sur une édition moderne de l'oeuvre, les résultats seront erronés. Cela n'exclut pas pour autant l'analyse de la longueur des phrases comme indice de style, à condition que la ponctuation originale des textes soit respectée.

Au cours des années soixante Andrew Q. Morton entreprend une étude du style des Épîtres de Saint Paul avec la collaboration de James McLeman. Il se concentre sur plusieurs textes classiques grecs et calcule la longueur des phrases, aussi bien que la fréquence des mots-outils les plus communs. Ensuite, il vérifie la longueur des phrases dans la version grecque des *Épîtres* de Paul et conclut que Paul n'a écrit que quatre de ces épîtres, *Romains I*, *Romains II*, *Romains III* et *Corinthiens*, étant donné que la longueur des phrases est presque la même dans ces quatre épîtres et non pas dans les autres. Mais les littéraires, les spécialistes de la Bible et les statisticiens s'empressent de critiquer sa

méthodologie, l'accusant d'avoir passé sous silence la ponctuation originale des textes, ce qui a dû fausser ses résultats.

Durant les années soixante-dix, Geir Kjetsaa publie une analyse du style de l'auteur russe Mikhäil Alexandrowich Cholokhov (1905-1984), en se basant sur la longueur des phrases. Cholokhov est le lauréat du Prix Nobel de littérature en 1965 pour son roman *Le Don paisible*. Après avoir reçu le prix Nobel, Cholokhov est accusé de plagiat; on affirme que Krynkow, auteur cosaque peu connu, mort de typhus en 1920, est le véritable auteur du roman.

En 1974 paraît en France une publication anonyme affirmant elle aussi que Krynkow est l'auteur du *Don paisible*; dans la préface du livre, Alexandre Solzhenitzyn accuse ouvertement Cholokhov de plagiat. Ceci s'explique sans doute par le fait que Cholokhov a toujours été un communiste convaincu et un ami intime de Staline, tandis que Krynkow s'est toujours déclaré anticomuniste, ce qui a dû provoquer un certain parti-pris de la part de Solzhenitzyn contre Cholokhov.

Premièrement Kjetsaa choisit un roman écrit par Cholokhov et un roman écrit par Krynkow. Ensuite, il tire au hasard le même nombre de pages de narration dans ces deux romans, et calcule la longueur moyenne des phrases dans ces textes. En dirigeant ensuite son attention sur le roman *Le Don paisible*, il choisit de nombreuses pages de narration prises au hasard dans ce roman et calcule la longueur moyenne de ces phrases.

Finalement, il compare la longueur des phrases dans le roman de Choloikhov et dans le roman de Krynkow à celle du *Don paisible*.

Il constate alors que la longueur des phrases chez Choloikhov est statistiquement la même que la longueur des phrases chez l'auteur du *Don paisible*, mais qu'il existe très peu de ressemblances entre le style de Krynkow et celui de l'auteur du *Don paisible*. D'autres tests statistiques (tel que le χ^2) confirment ces résultats. Kjetsaa se concentre ensuite sur un autre aspect du style, à savoir, la fréquence des vocables rares dans *Le Don paisible* et dans l'oeuvre de ces deux auteurs. Après l'application de tests statistiques rigoureux, Kjetsaa conclut que Choloikhov est l'auteur du *Don paisible*. Il l'affirme en ces termes: "The language seems to reveal that he [Choloikhov] wrote his own work, in which case the charge of plagiarism is null and void" (248).

Au cours de leur analyse du style de Madison et de Hamilton, Mosteller et Wallace se concentrent eux aussi sur la longueur des phrases. En faisant quelques tests statistiques, ils s'aperçoivent qu'elle est presque identique chez les deux auteurs, telle qu'indiquée dans la table 8.1.

Table 8.1

La Longueur des phrases chez Hamilton et Madison

	Hamilton	Madison
Moyenne	34,55	34,59
Écart-type	19,20	20,30

Source: Mosteller and Wallace, *Applied Bayesian and Classical Inference* 145

En 1984, dans un article s'intitulant "An Investigation of Morton's Method for the Determination of Authorship" M. W. A. Smith continue ses recherches dans le but de distinguer entre le style de Shakespeare et celui de Marlowe; il se concentre alors sur la longueur des phrases comme un indice de style. En calculant la longueur des phrases dans le poème "Hero and Leander" de Marlowe, les poèmes "Venus and Adonis" et "The Rape of Lucrece" et la pièce *Pericles, Prince of Tyre* de Shakespeare, il constate que "the method based on sentence-length was a complete failure" (77).

Comme Mosteller et Wallace, Smith démontre que la longueur des phrases comme indice de style est trop instable et ne permet pas de différencier entre le style des auteurs. Morton est convaincu de la validité de cette méthode. Mais sa méthodologie laisse beaucoup à désirer. Il ne tient pas compte de la ponctuation originale des textes, ce qui a dû fausser ses résultats.

Il est vrai que Udny Yule a pu distinguer entre le style de Thomas à Kempis et celui de Jean Charlier de Gerson, mais on ne sait au juste s'il a travaillé sur une édition moderne de l'oeuvre. Kjetsaa a eu du succès dans ce domaine: il a pu démontrer que Choloikhov est vraisemblablement l'auteur du *Don paisible*.

8.5. La Répartition des éléments grammaticaux

En 1966, H. H. Somers déclare que les différentes parties du discours constituent un élément essentiel du style, c'est-à-dire que la répartition des substantifs et des verbes dans un texte en dit long sur la personnalité de l'auteur. Il s'explique en ces termes:

A more cultivated habit of thinking can increase the amount of substantives used, while a more dynamic empathy and active attitude can be habitually expressed by means of an increased number of verbs. It is also possible to detect a number of idiosyncracies in the use of prepositions, subordinations, conjunctions and articles. (45)

Presque à la même période Morton publie une étude de *La Septième lettre* généralement attribuée à Platon, dans le but d'en déterminer la paternité (Voir "The Authorship of Greek Prose"). Il calcule la distribution des parties du discours dans cette lettre aussi bien que dans *L'Apologie* de Platon et les compare l'une à l'autre. Il constate qu'elle n'est pas la même dans ces deux textes et conclut par cela que Platon n'est pas l'auteur de *La Septième lettre*.

Les experts de l'oeuvre de Platon ripostent qu'il s'agit de genres différents, c'est-à-dire que *L'Apologie* utilise un style oratoire tandis que *La Septième lettre* emploie un style narratif, et que le style des textes varie souvent d'après les genres. D'ailleurs, lorsque l'on compare *La Septième lettre* de Platon aux *Lois* de Platon, on peut observer des ressemblances frappantes entre leur style.

Malgré l'échec de Morton dans ce domaine, Louis Milic publie une analyse du style de Jonathan Swift en 1967, en se concentrant sur les parties du discours. Il part du postulat que ce sont les catégories grammaticales plutôt que le vocabulaire lui-même, qui

constituent l'élément le plus essentiel du style, puisqu'elles demeurent le même au cours de la carrière littéraire d'un auteur, quel que soit le sujet ou le genre employé.

Milic classe les vocables qui figurent dans l'oeuvre de Jonathan Swift et de quatre autres auteurs contemporains d'après leur catégorie grammaticale. Ensuite, il compare la distribution de ces catégories dans les cinq textes. Il peut à peine distinguer entre le style de ces cinq auteurs.

En 1973, Barron Brainerd publie une étude dont le but est de vérifier si la fréquence des parties du discours varie d'un genre à l'autre; à savoir entre les pièces de théâtre, les "belles lettres" (les romans et les nouvelles, ou tout ce qui en littérature a trait aux connaissances littéraires par opposition aux sciences) et les autres oeuvres en prose non littéraire.. En se concentrant sur 59 textes qui englobent plusieurs genres, Brainerd calcule la proportion des parties du discours pour chaque genre. Les pièces de théâtre contiennent beaucoup de pronoms et très peu d'articles, mais les oeuvres en prose contiennent beaucoup d'articles et peu de pronoms ("On the Distinction Between a Novel and a Romance").

Les résultats qu'obtiennent Brainerd, Somers et Milic montrent qu'une analyse statistique des parties du discours n'est pas *a priori* absurde, même si plusieurs autres facteurs sont en jeu dans la création de l'oeuvre littéraire. Mais Milic et Brainerd montrent aussi qu'une telle méthodologie ne distingue pas entre les oeuvres appartenant

au même genre, étant donné que les catégories grammaticales sont uniformes dans ces oeuvres.

Tous les exemples cités plus haut s'appliquent soit à une étude du style anglais (Milic, Somers, Brainerd) ou du style grec (Morton). Si on envisage une telle étude pour un texte français, il faut se rappeler que, comme le souligne Bally, le français, ayant tendance à insister sur la permanence des choses, favorise l'emploi du substantif car "bien loin de rechercher le devenir dans les choses, il [le français] présente les événements comme des substances" (591). Si on envisage une analyse statistique de ces parties du discours en français, on devrait tenir compte de la prépondérance du substantif en français.

L'analyse des parties du discours n'est pas abordée dans cette étude car il aurait fallu un logiciel qui classifie les vocables selon les parties du discours sans frais exagérés. Par ailleurs, comme toutes les oeuvres qui font l'objet de cette étude appartiennent au genre romanesque, l'analyse des parties du discours risque fort de ne pas produire de résultats utiles.

8.6. La Fréquence des lettres

La pièce *Two Noble Kinsmen* a été imprimée en 1634 par John Waterson et porte l'inscription "Written by the memorable Worthies of their time; Mr. John Fletcher, and Mr. William Shakespeare - Gent" (Ledger et Merriam 234). Depuis ce temps plusieurs

experts cherchent à déterminer quelle partie de cette pièce attribuer à Shakespeare et quelle partie à Fletcher.

À l'époque victorienne Harold Littledale et Samuel Hickson (Ledger et Merriam 241) affirment que le personnage du médecin dans la scène iii de l'acte 4 de la pièce *Two Noble Kinsmen* ressemble aux médecins figurant dans les pièces *Hamlet* et *Macbeth* de Shakespeare. Par contre, ils notent que le médecin dans la scène ii de l'acte 5 ressemble plutôt aux médecins des farces grossières de Fletcher. Littledale et Hickson attribuent alors la scène iii de l'acte 4 à Shakespeare, et la scène ii de l'acte 5 à Fletcher.

Depuis ce temps d'autres experts affirment que l'acte 1 et l'acte 5 de la pièce *Two Noble Kinsmen* sont de Shakespeare, tandis que les autres actes sont de Fletcher. En 1970, G.R. Proudfoot distingue deux styles différents et attribue les diverses parties de la pièce à Shakespeare et à Fletcher, tel qu'indiqué dans la table 8.2.

En 1987 Horton attribue, lui aussi, les diverses parties de la pièce *Two Noble Kinsmen* soit à Shakespeare, soit à Fletcher (voir table 8.2). Cette table démontre que Proudfoot et Horton ne s'accordent pas sur la paternité de cette pièce: Proudfoot est de l'avis que Fletcher en a rédigé la plus grande partie tandis que Horton affirme le contraire. Étant donné cette divergence d'opinion, Gerard Ledger et Thomas Merriam entreprennent une étude de la pièce pour déterminer quelle partie attribuer à Fletcher et quelle partie à Shakespeare.

Table 8.2

La Classification d'après Proudfoot et Horton

	Shakespeare	Fletcher
Proudfoot	Acte 1	Acte 2, sc 2-6
	Acte 3, sc 1-2	Acte 3, sc 3-6
	Acte 5, sc 1, sc 3-4	Acte 4, sc 1-2
		Acte 5, sc 2
Horton	Acte 1, sc 1-3	Acte 2, sc 2
	Acte 2, sc 1, sc 3	Acte 3, sc 6
	Acte 3, sc 1	
	Acte 4, sc 3	
	Acte 5, sc 1, sc 3-4	

Source: Ledger et Merriam 234

Pour distinguer le style de Fletcher de celui de Shakespeare, Ledger et Merriam se concentrent sur la fréquence des lettres. Cette méthodologie se base sur le fait qu'en anglais par exemple la lettre *e* est la plus courante et que *x* et *z* sont les plus rares. Ledger et Merriam postulent aussi que "authors have a distinctive and characteristic letter frequency pattern" (Ledger et Merriam 235). À l'aide de la classification hiérarchique ils déterminent quelles lettres sont les plus caractéristiques du style de Shakespeare et de Fletcher. Après de nombreux tests statistiques, ils constatent qu'une grande partie de la pièce *Two Noble Kinsmen* a été composée par Shakespeare. Ils attribuent ainsi à Shakespeare les scènes, ii, iii et iv de l'acte 1; les premières scènes des actes 2, 3 et 4, la scène iii de l'acte 4, et les scènes ii, iii et iv de l'acte 5.

À Fletcher sont attribuées les scènes ii et v de l'acte 2; les scènes iii, iv et vi de l'acte 3. (Voir table 8.3). Cependant huit scènes de la pièce ne peuvent, en fin de compte être attribuées définitivement ni à Shakespeare ni à Fletcher.

Table 8.3

La Classification d'après Ledger et Merriam

Fletcher		Acte 2, sc 2, sc 5	Acte 3, sc 3, 4, 6		
Shakespeare	Acte I, sc 2-4	Acte 2, sc 1	Acte 3, sc 1	Acte 4, sc 1, sc 3	Acte 5, sc 2-4

Source: Ledger et Merriam 244

La recherche entreprise par Ledger et Merriam corrobore celle de Horton. Comme lui, ils affirment que l'acte 1, l'acte 2, et l'acte 5 sont surtout de la plume de Shakespeare et que les actes 2 et 3 sont de Fletcher. L'acte 4 est plutôt problématique. Ils affirment alors "What is quite new is the fact that authorial letter frequency patterns in the sixteenth and seventeenth Century dramatic writing are discernible" (247).

Malgré le succès de Ledger et Merriam, leur méthode ne peut être mise en pratique que difficilement sans frais exagérés. D'ailleurs, étant donné l'étendu de cette étude, il aurait fallu utiliser le logiciel de la classification hiérarchique, méthode qui est assez compliquée.

8.7. L'Analyse des "vocables distinctifs"

Certains critiques attribuent parfois à Shakespeare ou à d'autres dramaturges des oeuvres jusqu'alors demeurées anonymes. Quatre pièces de théâtre *The Double Falsehood*, *The London prodigal*, *Henry VIII* et *The Two Noble Kinsmen* ont ainsi été attribuées à Shakespeare ou à Fletcher au cours des années. Il est maintenant reconnu que la pièce anonyme *The Double Falsehood* attribuée à Shakespeare est une oeuvre de collaboration entre Shakespeare et Fletcher (Taylor, 1987).

The London Prodigal est une autre oeuvre anonyme souvent associée à Shakespeare. Récemment Merriam a démontré que Fletcher en est l'auteur (chapitres 10 et 11). Quant aux pièces *Henry VIII* et *The Two Noble Kinsmen*, on considère qu'elles résultent d'une collaboration entre Shakespeare et Fletcher (Voir Hart, Maxwell, Shoenbaum et Proudfoot).

Robert Matthews et Thomas V. N. Merriam voulant trancher la question, réévaluent la paternité de ces pièces. Ils développent un logiciel (le *multi-layer perceptron* ou *MLP*) qui est capable de traiter rapidement plusieurs pièces de Shakespeare et de Fletcher et de révéler quelques indices de style, telles que la fréquence d'une seule lettre, la proportion entre les vocables *did* et *did + do*, entre les vocables *no* et *no + not* et de distinguer entre le style de Shakespeare et de Fletcher.

Ils peuvent alors faire des distinctions entre le style de ces deux auteurs. Ensuite en dirigeant leur attention sur des pièces qui sont parfois attribuées à Shakespeare et

parfois à Fletcher, ils concluent qu'elles doivent être attribuées de la manière suivante: *The Double Falsehood* et *The London Prodigal* à Fletcher, *Henry VIII* et *Two Noble Kinsmen* à Shakespeare.

En 1994, Matthews et Merriam vérifient si Christopher Marlowe pourrait être l'auteur de la pièce *Henry VI, part 3* (généralement attribuée à Shakespeare). En utilisant le logiciel *MLP*, ils distinguent entre le style de Shakespeare et celui de Marlowe. Ensuite ils construisent un logiciel *MERMAR* (un acronyme de Merriam et de Marlowe) et analysent les pièces de Shakespeare et de Marlowe. Ils découvrent que les pièces *Edward III* de Marlowe et *The Jew Of Malta* (qui est censée être de Marlowe) ressemblent plutôt aux pièces de Shakespeare. Par contre les pièces *The Contention* et *The True Tragedy* aussi bien que la pièce *Henry VI, part 3* de Shakespeare ressemblent le plus au style de Marlowe. (Merriam et Matthews 1994, 6).

Les résultats sont conformes à l'affirmation de Tucker Brooke qui avait postulé en 1912 que la pièce *Henry VI, part 3* de Shakespeare est un remaniement de la pièce *The True Tragedy of Richard Duke of York* de Marlowe. Matthews et Merriam concluent ainsi:

Tucker Brooke's argument for [...] *The True Tragedy* being by Marlowe, and that *Henry VI, part 3* is a Shakespearian revision of a Marlovian original, all deserve renewed consideration. (6)

Ledger, Merriam et Matthews démontrent qu'il est aujourd'hui possible d'effectuer des analyses objectives sur la fréquence des vocables en se servant de l'ordinateur et de la statistique. Leur méthodologie exige une connaissance très approfondie de la statistique et demeure difficile à mettre en pratique pour les littéraires. C'est l'une des raisons pour lesquelles elle ne sera pas utilisée dans cette étude.

8.8. La Fréquence des vocables dans un corpus

De plus en plus, les analyses statistiques du style se concentrent, aujourd'hui, sur la fréquences des vocables. L'analyse du lexique existe depuis assez longtemps, quoique de temps en temps elle soit employée de manière impressionniste. En 1962 Alvar Ellegård adopte une telle méthodologie pour entreprendre un examen stylistique des lettres anonymes les *Junius Letters*.

Les *Junius Letters* sont une série de pamphlets politiques écrits entre 1769 et 1772 et attribués généralement à Sir Philip Francis quoique de temps en temps quelques experts remettent en question leur paternité. Voulant trancher la question, Ellegård se concentre sur la fréquence de certains vocables dans ce corpus comme un indice de style.

Il examine l'oeuvre de plusieurs auteurs importants du dix-huitième siècle, y compris celle de Sir Philip Francis. Ensuite, il constitue un corpus de plus d'un million de vocables. Enfin, il compile la liste des vocables dans les *Junius Letters* et détermine ceux

qui caractérisent ces lettres. Il attribue l'étiquette *plus* aux vocables utilisés le plus souvent dans ces lettres et l'étiquette *moins* aux vocables utilisés le moins souvent. Ellegård vérifie ensuite la fréquence de ces mêmes vocables dans les autres oeuvres, y compris celle de Sir Phillip Francis.

En appliquant des tests statistiques, Ellegård observe des similarités frappantes entre le style de l'oeuvre de Phillip Francis et celui des *Junius Letters*, ce qui indique que Philip Francis est effectivement l'auteur des *Junius Letters*.

En 1975, A. Q. Morton utilise une telle méthodologie dans un milieu médico-légal lorsqu'il tente de démontrer que les dépositions que la police attribue à Steve Raymond (un détenu de la Prison de Brixton) sont frauduleuses. En postulant que la fréquence des vocables dans un texte constitue un indice de style, il analyse le style de Steve Raymond (*Literary Detection*).

Il affirme que les auteurs anglais utilisent des locutions figées telles que *of course*, *as if* à une fréquence constante. Il calcule la fréquence avec laquelle Steve Raymond utilise ces locutions dans ses lettres et la compare à celle obtenue dans ces dépositions. Morton observe que ces locutions figées ne se produisent pas à la même fréquence dans les dépositions et conclut qu'elles ont été rédigées par d'autres personnes. Il convainc la Cour d'Assises The Old Bailey que ces dépositions sont frauduleuses. Raymond est libéré sur le champ. Le vocabulaire qu'emploie un auteur constitue souvent un indice de son

style. Mais la méthodologie employée demeure souvent subjective, quoiqu'Ellegård ait eu du succès dans ce domaine.

8.9. John Burrows et l'Analyse des vocables les plus fréquents

En 1987 John Burrows publie une étude très détaillée de l'oeuvre de Jane Austen. À l'aide de plusieurs logiciels, il étudie la langue des personnages présents dans les romans d'Austen dont le dialogue est constitué d'au moins 2.000 vocables. Il obtient une liste de quarante-huit personnages répartis de la manière indiquée dans la table 8.4.

En premier lieu, cette table indique le nombre absolu des vocables pour tous les personnages dans chaque roman: elle remonte à 63.348 vocables pour le roman *Mansfield Park* par exemple. Ensuite la table regroupe le nombre de vocables constituant le dialogue de chaque personnage. La dernière catégorie qui est regroupée sous la rubrique "divers" comprend le dialogue de tous les autres personnages du roman dont l'idiolecte pris séparément est inférieur à 2.000 vocables.

Table 8.4

Les Personnages dans les romans de Jane Austen dont le dialogue comprend au moins 2.000 vocables

<i>Northanger Abbey</i>	28.930	<i>Persuasion</i>	28.715
Catherine	7.040	Anne	4.336
Henry Tilney	6.149	Mrs Smith	3.985
Isabella Thorpe	5.657	Mary Musgrove	3.733
John Thorpe	2.928	Wentworth	3.386
Eleanor Tilney	1.938	Admiral Croft	2.034
Divers	5.218	Divers	11.241
<i>Sense and Sensibility</i>	52.334	<i>Emma</i>	78.057
Elinor	9.039	Emma	21.501
Marianne	6.580	Mr Knightley	10.112
Mrs Jennings	5.659	Frank Churchill	8.225
Willoughby	5.278	Miss Bates	7.624
John Dashwood	4.462	Mrs Elton	6.415
Lucy Steele	3.977	Harriet Smith	5.274
Colonel Brandon	3.700	Mr Woodhouse	4.639
Mrs Dashwood	3.560	Mrs Weston	4.144
Edward Ferrars	2.359	Mr Weston	3.197
Anne Steele	2.066	Divers	6.926
Divers	5.654		
<i>Pride and Prejudice</i>	55.976	<i>Mansfield Park</i>	63.348
Elizabeth	13.597	Edmund	14.300
Darcy	6.399	Mary Crawford	13.030
Mrs Bennet	6.048	Henry Crawford	6.735
Jane	5.203	Mrs Norris	6.659
Mr Collins	4.444	Fanny	6.117
Mr Bennet	3.248	Sir Thomas Bertram	3.953
Mrs Gardiner	2.865	Tom Bertram	2.995
Lady Catherine	2.342	Lady Bertram	2.162
Lydia	2.214	Divers	7.397
Miss Bingley	2.127		
Wickham	2.040		
Divers	5.449		

Source: Burrows, *Computation*, préface

Par la suite, Burrows regroupe les types les plus fréquents dans l'idiolecte de chaque personnage en ordre décroissant, tel que l'indique la table 8.5.

Table 8.5

Les Trente Vocables les plus fréquents de l'idiolecte des personnages de Jane Austen

Rang	Type	Fréquence	Rang	Type	Fréquence	Rang	Type	Fréquence
1	I	38,75	11	is	14,85	21	for	7,45
2	the	26,45	12	in	13,39	22	me	7,41
3	you	24,97	13	have	12,20	23	was	7,18
4	and	24,64	14	to	11,19	24	very	7,06
5	of	23,69	15	my	10,35	25	she	7,00
6	to	20,34	16	but	9,99	26	will	6,82
7	a	18,54	17	as	9,39	27	do	6,51
8	it	18,41	18	he	8,87	28	your	6,32
9	not	16,14	19	her	8,10	29	with	6,22
10	be	14,92	20	that	7,60	30	at	5,88

Source: Burrows, *Computation* 82

La table 8.5 indique la fréquence relative des trente vocables les plus fréquents dans l'idiolecte des personnages. Elle est composée de huit pronoms personnels, de six verbes auxiliaires, de cinq prépositions, de trois conjonctions, de deux adverbes, de l'article défini et de l'article indéfini et de quatre autres vocables (*to*, *that*, *for* et *all*). La plupart de ces vocables sont des mots-outils. Le vocable *I*, qui a une fréquence relative de 38,75 est le plus fréquent de tous.

Pour la plupart des critiques littéraires, ces mots-outils jouent un rôle stylistique négligeable, mais Burrows concentre son étude sur ces trente ou quarante vocables (qui sont pour la plupart des mots-outils) étant donné qu'ils constituent entre un tiers et une moitié du texte:

Since so few as thirty word-types make up about two fifths of all the word-tokens most writers employ [...] their contribution to the size of any text suggests that they may contribute to its 'shape' ("An Ocean" 311).

En appliquant plusieurs tests de statistique (le χ^2 et la corrélation de Pearson entre autres) Burrows démontre que la fréquence des types les plus nombreux permet de différencier entre le langage des personnages: "it is possible to differentiate sharply and appropriately between the idiolects of Jane Austen's characters and even to trace the ways in which an idiolect can develop during the course of a novel" (*Computation* 4). Burrows démontre aussi que la fréquence de ces trente types permet de distinguer entre les personnages:

Of all of Jane Austen's characters, it emerges that Elizabeth Bennett and Emma Woodhouse speak most like each other, whereas Mr Collins and Harriet Smith speak least like each other. There are close resemblances [...] between such authoritarians as Sir Bertram and Lady Catherine de Bourgh and [...] among such vulgarians as Mrs Bennet, her daughter Lydia, Lucy Steele and Elizabeth Thorpe ("Modal Verbs" 21).

Burrows réfute donc l'argument selon lequel les mots-outils sont hors du processus de la création littéraire et du contrôle de l'auteur et ne jouent qu'un rôle négligeable. En 1989, Burrows publie un article dans lequel il fait une analyse statistique du style de plusieurs romanciers tels que Henry Fielding (1707-54), Sarah Fielding (1710-68), Jane Austen (1775-1817) et Tobias Smollet (1721-71) ("An Ocean"). Il examine les types les plus fréquents dans la partie constituant la biographie des personnages, afin d'établir le genre de similarités qui existe entre eux (312).

En faisant des tests de corrélation entre les romans de Sarah Fielding et de Henry Fielding, Burrows observe qu'il existe plus de ressemblances entre les romans écrits par le même auteur qu'entre ceux écrits par différents auteurs. Lorsque Burrows compare les oeuvres de Henry Fielding entre elles, il obtient de très fortes corrélations qui varient entre 0,995 et 0,936.

En faisant les mêmes tests pour les romans de Sarah Fielding, Burrows obtient des corrélations qui varient entre 0,995 et 0,934. Mais lorsqu'il compare les romans de Sarah Fielding à ceux de son frère Henry, il obtient les corrélations moins fortes qui varient entre 0,969 et 0,873, quoiqu'elles demeurent très significatives. En ce qui concerne les romans des Fielding et ceux de Tobias Smollett, Burrows obtient des corrélations moins fortes. Lorsque Burrows compare les romans des seize auteurs entre eux "the results amount to a clustering by chronology and by gender" ("An Ocean" 317). Il est vrai que

les corrélations sont toutes très significatives, mais ce sont les romans d'un même auteur qui produisent les plus fortes corrélations.

Burrows a été le premier à employer une telle méthodologie. Elle est très reconnue aujourd'hui et les chercheurs l'adoptent de plus en plus. Je me servirai de cette méthodologie dans ce travail.

8.10. La Fréquence des synonymes

Durant la première phase de leur analyse des articles *Federalist*, Mosteller et Wallace se concentrent sur plusieurs aspects du style tels que la proportion des substantifs et des adjectifs, la longueur des vocables et la fréquence de l'article défini *the*. Aucune de ces méthodes ne leur permet de distinguer entre le style de Madison et celui de Hamilton.

À la deuxième phase de leur enquête, ils se concentrent sur les mots-outils, à savoir les vocables qui ne dépendent ni du sujet ni du contexte. Ils dressent une liste de vocables synonymes tels que *while/whilst*, *on/upon* et se concentrent sur la fréquence de ces vocables dans le texte. En examinant dix-huit articles, ils constatent que Hamilton se sert de l'adverbe *enough* dans quatorze de ses articles et que Madison ne l'utilise que dans un seul de ses articles. Par ailleurs, chez Hamilton la préposition *on* se rencontre à une fréquence relative de 3,28, tandis que la préposition *upon* figure à une fréquence relative de 3,35. Par contre, Madison utilise la préposition *on* à une fréquence relative de 7,83 et la préposition *upon* à une fréquence relative de 0,14. (Voir table 8.6).

Table 8.6

Fréquence relative des vocables par tranche de mille vocables

	On	Upon	While	Whilst
Hamilton	3,28	3,35	0,28	0,00
Madison	7,83	0,14	0,02	0,48

Source: Mosteller et Wallace, *Inference* 102

Ayant ainsi établi la préférence de chaque auteur pour tel ou tel vocable, Mosteller et Wallace font plusieurs tests statistiques et déclarent par la suite qu'il est fort probable que Madison ou Hamilton ait écrit tel ou tel essai. Les recherches entreprises par Mosteller et Wallace indiquent que l'emploi des synonymes tels que *while* ou *whilst* dans un texte servent à distinguer entre les auteurs. Les experts considèrent que cette recherche est l'une des plus importantes dans le domaine d'attribution d'auteur. Dans cette analyse du style de Gary/Ajar je me concentrerai, sur l'emploi des synonymes comme un indice de style.

8.11. Conclusions

Plusieurs méthodes d'investigation sont possibles dans le domaine d'attribution d'auteur aux oeuvres anonymes ou pseudonymes et existent depuis très longtemps. Toutes ces méthodes ne sont pas pour autant aussi efficaces les unes que les autres.

La longueur des phrases comme un indice du style n'a pas encore produit de résultats satisfaisants. Plusieurs chercheurs ont démontré qu'elle varie d'après les genres plutôt que d'après les auteurs. De plus cet aspect du style change souvent durant la vie d'un auteur. À l'époque moderne, ce sont les éditeurs et non pas les auteurs qui décident la longueur des phrases dans une oeuvre:

Statistics on sentence length [...] are best suited to providing confirmatory evidence which corroborates and substantiates an indication of authorship obtained by other means [...] the information provided is not sufficiently strong as to warrant the use of such measures as a stand-alone technique to discriminate between authors. (Smith, "Recent Experience" 78)

C'est pour cette raison que je n'utiliserai pas cette méthodologie dans ce travail. Je n'entreprendrai pas non plus une analyse de la longueur des vocables. Comme le démontrent Mosteller et Wallace la longueur des vocables est presque identique chez Hamilton et chez Madison, ce qui fait qu'il est impossible de différencier entre leur style.

Brainerd postule que le nombre de syllabes par vocable varie d'après les genres. Il reconnaît néanmoins que pour certains auteurs le nombre de syllabes demeure le même dans leurs écrits, quel que soit le genre traité ("On the Distinction" 270). En fait, d'après David Holmes: "the distribution of syllables per word [...] being an easily accessible index of its style, is one area that might prove profitable in authorship studies" (88).

Néanmoins, la méthodologie adoptée dans ce travail porte sur l'étude du vocabulaire, et ne se limite pas aux vocables qui dépendent du contexte.

Elle se base sur celle de Burrows (l'analyse des vocables les plus fréquents) et celle de Mosteller et Wallace (l'emploi des synonymes). Ces deux méthodologies me permettront de déterminer s'il existe une différence entre le style de Gary et celui d'Ajar.

Chapitre 9

Quelques Notions fondamentales sur la statistique

La littérature a toutes les caractéristiques de la science [...] comme la science, la littérature est méthodique: elle a ses programmes de recherche [...] ses règles d'investigation, parfois même ses prétentions expérimentales. (Barthes 11-12)

Ce chapitre fournit des détails sur la langue et la statistique. Il donne d'abord des définitions sur la terminologie employée dans ce travail. Ensuite, il explique quelques notions fondamentales de la statistique. Si les explications sont parfois détaillées, elles apparaissent néanmoins dans un langage que tout littéraire est en mesure de comprendre.

9.1. Analyse statistique de l'oeuvre littéraire

Mais, avant d'aborder une discussion de la statistique elle-même, il est utile de détailler les sept étapes qui s'imposent pour l'analyse statistique (voir table 9.1).

Table 9.1

Les Sept Étapes nécessaires pour l'analyse statistique du style

-
- (1) Définition du problème
 - (2) Choix du modèle
 - (3) Collecte des données
 - (4) Choix de l'algorithme
 - (5) Calculs
 - (6) Première Phase de l'interprétation (phase I)
 - (7) Deuxième Phase de l'interprétation (phase II)
-

Les étapes 1 et 2 appartiennent aux littéraires qui choisissent le problème duquel vont dépendre les conclusions. Par exemple, j'analyse quatre romans de Romain Gary/Émile Ajar pour déterminer lequel des deux points de vue concernant le style d'un auteur domine, celui qui postule que "Le style c'est l'homme même" ou l'autre qui maintient qu'un auteur peut changer son style quand il le veut. À l'étape deux, il est nécessaire de fournir des définitions sur la nature du langage, de définir quelques

notions de base (telle que la nature du mot), d'indiquer la longueur du texte à étudier et de signaler les unités en fonction desquels le texte se compte.

Le reste de l'analyse (étapes 3, 4, 5, 6 et 7) nécessite une connaissance de la statistique. C'est là que le chercheur fait le traitement des textes, choisit le logiciel, applique les formules et l'algorithme (l'ensemble des règles opératoires propres à un calcul) et fait les tests statistiques. Le chercheur détermine aussi à quel degré de certitude ou de probabilité un résultat statistique est significatif. Les étapes 6 et 7 sont celles de l'interprétation. Elle comporte une interprétation statistique des tests (étape 6). Enfin, la septième et dernière étape appartient au littéraire qui explique en quoi le résultat est significatif.

9.2. *Mot, vocable, lemme, vocabulaire, lexique, type, occurrence*

La tradition classe les mots selon les parties du discours fondées sur leur fonction syntaxique. Le dictionnaire *Robert* explique qu'un mot est "chacun des sons ou groupe de sons (de lettres ou groupes de lettres) correspondant à un sens isolable spontanément dans le langage". Cette définition s'applique au mot oral aussi bien qu'au mot écrit. Charles Muller explique que le mot est "une unité graphique, séparée des unités voisines par un blanc ou une ponctuation" (*Initiation* 145) et Benoît Mandelbrot affirme que le mot est "toute suite de lettres contenue entre deux signes 'intervalle' successifs, indépendamment de toute question de sens" (Galdi 49).

Ce qui ressort de toutes ces définitions, c'est que le mot est une unité de pensée qui peut être notée graphiquement entre les blancs et que cette unité est découpée dans le discours. Le mot inclut la forme conjuguée du verbe, tel que le participe passé *parties* ou la contraction *du*. L'on sait cependant que la séparation des mots dans le langage n'a pas toujours existé. L'ancien grec liait les mots ensemble et la langue latine était la première en Occident à utiliser les points pour séparer les mots; plus tard des blancs ont remplacé ces points.

On peut ainsi se demander si une étude des mots est légitime ou même désirable, si le langage ne constitue pas un tout, un ensemble, qu'il ne faut pas dissocier. La réponse à cette question est négative, car même lorsque le mot français est mal délimité phonétiquement et que son autonomie n'est marquée par aucune coupure ou pause lors de l'énonciation d'une phrase, son autonomie est évidente.

Il faut aussi distinguer entre les mots principaux (substantif, adjectif, verbe, adverbe) qui représentent des êtres ou des objets ayant un sens plein, et les mots accessoires ou mots-outils (tels que *le*, *de*, *ma*) qui sont des déterminants de relation. Même si les locuteurs ont souvent tendance à croire que les mots pleins sont plus significatifs que les mots-outils, John Burrows a démontré que les mots-outils en disent beaucoup sur les différences entre les énoncées (*Computation*).

Dans cette étude le mot *vocable* est utilisé pour les unités étudiées, tandis que le terme *mot* est employé dans son sens le plus large. Le terme *lemme* comprend

l'ensemble des formes flexionnelles faisant partie du vocable. Pour *vocabulaire* et *lexique*, encore une fois, la définition que propose Charles Muller est retenue. Le *vocabulaire* comprend des ensembles délimités: par exemple il peut comprendre le vocabulaire de Racine, le vocabulaire de la médecine, des mathématiques etc., ou il peut inclure tous les mots présents dans un texte. Quant au *lexique*, il comprend l'ensemble des ressources lexicales dont disposent les locuteurs ou les mots en jeu dans l'acte de discours. Ainsi ce travail se fait sur le *vocabulaire* et non sur le *lexique*. Pour la définition des termes *type* et *occurrence*, consulter le chapitre 7.

Après ces quelques notes explicatives sur les termes employés dans ce travail, on peut pouvoir maintenant aborder la question de la statistique elle-même.

9.3. La Statistique

La statistique employée au singulier avec l'article est "un art de penser" (*Initiation* 6). Elle demeure pour certains un objet de foi, comme lorsqu'il s'agit de Dieu ou des extra-terrestres, on y croit ou on n'y croit pas. Selon les pessimistes, la statistique est la façon moderne de mentir, c'est-à-dire qu'on fait dire aux statistiques ce que l'on veut. Pour les optimistes, au contraire, la statistique permet de résoudre des problèmes qui peuvent paraître épineux ou insolubles.

Dans les temps les plus reculés, les Grecs et les Romains utilisaient les statistiques pour se renseigner sur leurs biens. En Angleterre, au onzième siècle,

Guillaume le Conquérant (1027-1097) imposait de lourds impôts au peuple anglais. Afin de s'assurer que la Couronne obtienne ce qui lui revenait, Guillaume a fait compiler un relevé de statistique (le fameux *Domesday Book*) qui détaillait les biens de ses vassaux.

La statistique ne commence à prendre le statut d'une science en Angleterre, en Allemagne et en France qu'au dix-septième siècle. Pierre de Ferma (1601-1662), mathématicien français et Blaise Pascal (1623-1662), mathématicien, philosophe et écrivain français, fondent la théorie des probabilités. Jacques Bernoulli (1654-1705), mathématicien suisse, formule la loi des grands nombres; Thomas Bayes (1671-1746), théologien anglais, étudie la probabilité des causes, et Graunt (1620-1665), un autre Anglais, établit les taux de mortalité par âge et par sexe.

Au dix-huitième siècle Antoine Deparcieux (1703-1768), mathématicien français, calcule les chances de longévité pour chaque âge. Il indique que les femmes vivent plus longtemps que les hommes, que les personnes mariées vivent plus longtemps que les célibataires et ainsi de suite. Plus tard Pierre Laplace (1749-1827), mathématicien, physicien et astronome français démontre les utilisations pratiques des probabilités dans les sciences politiques.

C'est au milieu du dix-huitième siècle que le terme allemand *Statistik* (dont l'emploi est généralement attribué au Professeur Achenwall de l'Université de Göttingen) apparaît dans la langue. Ce terme est emprunté au mot italien *Statista*, qui

veut dire "homme d'état", puisqu'on considérait qu'un homme d'état devait posséder la connaissance des faits économiques et sociaux fondés sur des données numériques. Mais le mot "statistique" lui-même fait son apparition dans la langue française en 1835 (Legris v). La statistique a beaucoup évolué depuis le temps des Romains et des Grecs; elle est devenue

l'ensemble des procédés qui ont pour but l'étude scientifique des statistiques, c'est-à-dire des collections des chiffres, des tables, des graphiques décrivant et regroupant des observations faites sur des ensembles nombreux. (Legris, viii)

9.4. La Statistique dans le domaine littéraire et linguistique

De nos jours, la statistique est utile à tous; il n'existe aucune science (humaine, sociale, physique ou naturelle) qui ne fasse appel à ses services. Elle trouve une place importante dans d'innombrables versions de la stylistique moderne. Ce qui surprend néanmoins, c'est que de plus en plus les littéraires fassent eux aussi appel aux services de la statistique. En principe, les littéraires sont les esprits très rebelles au raisonnement mathématique, et un traitement statistique des textes littéraires soulève plus d'une objection de leur part. Plusieurs d'entre eux soulignent, à juste titre, que la valeur stylistique d'un texte littéraire se trouve dans l'agencement et l'harmonie des mots, dans

leur contexte, dans la syntaxe et non pas dans une liste alphabétique de vocables dépourvus de contexte.

Or, l'analyse statistique de l'oeuvre littéraire ne remplace nullement l'approche stylistique conventionnelle ni les investigations littéraires les plus traditionnelles (telle que l'analyse philosophique, biographique, sociologique, narratologique, linguistique ou stylistique). Loin de contredire ces autres approches, la statistique constitue un outil qui va de pair avec elles. Elle permet d'éliminer les barrières traditionnelles entre les disciplines et s'avère plutôt "un secours pour l'esprit littéraire." (Muller, *Initiation* 3).

Une approche statistique dans le domaine de la littérature peut ainsi se justifier par le besoin d'appliquer une méthodologie objective aux oeuvres qui depuis très longtemps n'ont été traitées que de manière impressionniste. De plus en plus, les littéraires, les philosophes, les linguistes, les stylisticiens, ne pouvant plus ignorer le concept de "corrélation", de "probabilité" ou de l'aspect statistique du langage, font eux aussi appel aux services de la statistique. Aussi faut-il se mettre en garde contre la méfiance instinctive des non-initiés à l'égard de la statistique.

Pour faciliter l'analyse statistique de l'oeuvre littéraire, l'ordinateur (qui est maintenant à la portée de tous) est un outil essentiel. Des logiciels de plus en plus conviviaux permettent de dénombrer et de catégoriser tout un ensemble de matériel avec rapidité et efficacité. Cette sorte de dépouillement est une aide précieuse à la critique littéraire.

9.5. La Nature statistique du langage

Avant d'aborder une analyse statistique de l'oeuvre littéraire, il est nécessaire de considérer un texte comme une population à l'intérieur de laquelle se trouvent des sous-ensembles, à savoir, des paragraphes, des phrases, des vocables. Gustave Herdan le précise en ces termes: " 'la langue' now has the essential characteristic of a statistical population [...] 'la parole' then appears as it ought to according to its meaning, as a term for statistical samples withdrawn from 'la langue' as the population" (27).

Par ailleurs, Muller affirme que les phénomènes humains (y compris le langage) lorsqu'ils ont lieu en grand nombre se conforment aux lois de la statistique (*Étude 9*). Muller explique d'ailleurs que, quoiqu'il existe des contraintes linguistiques qui interdisent d'écarter du discours certains éléments de relation ou certaines catégories de vocables, le langage ressemble à un jeu de hasard et "les mécanismes inconscients jouent un trop grand rôle dans l'élaboration du discours pour que celui-ci puisse être considéré comme un acte entièrement et constamment contrôlé par la volonté individuelle" (*Étude 10*).

Ceci implique que le style d'un auteur (y compris son vocabulaire) est soumis aux lois statistiques et que l'emploi que fait l'auteur du langage est semblable au tirage aléatoire, même si l'auteur a le sentiment d'être complètement libre dans le choix de son vocabulaire. Si le langage contient cet élément probabiliste, il ne faut pas ignorer le rôle important que jouent les préférences verbales et l'aptitude linguistique de l'auteur

dans la création de l'oeuvre littéraire. En fait, l'écrivain fait de la langue un emploi volontaire et conscient en lui imposant une intention esthétique. Il existe de grosses fluctuations dans la fréquence des mots dans le texte.

L'occurrence d'un vocable dans le texte n'indique pas quand il réapparaîtra; il peut réapparaître à la phrase suivante ou après les mille vocables qui suivent, car dans la réalité que constitue l'art de l'auteur, celui-ci peut répéter un mot plusieurs fois pour des effets de style. L'extrait suivant tiré de *Clair de Femme* peut servir d'exemple:

--*Elle* reviendra. Bien sûr, *elle* ne sera plus tout à fait la même. *Elle* aura un autre regard, un autre physique. *Elle* s'habillera autrement, quoi. Il est normal, naturel, qu'une femme change. Qu'*elle* change d'apparence, qu'*elle* ait des cheveux blancs, par exemple, une autre vie, d'autres malheurs. *Elle* reviendra. (58-9)

Dans cet extrait de cinquante-quatre vocables le type *elle* apparaît sept fois (129 fois par tranche de mille vocables). Et pourtant quand le triage a été fait, la fréquence de *elle* dans *Clair de femme* s'élève à 8,89 par tranche de mille vocables (voir appendice 10.1.1). Dans le passage cité, la répétition du vocable *elle* existe pour des effets de style, ce qui démontre que la fréquence du vocable *elle* ne peut ainsi être fixée qu'en moyenne. Lorsqu'on déclare qu'un auteur utilise *elle* à une fréquence de 8,89 par tranche de mille vocables, il ne s'agit que d'une tendance générale de la chose.

9.6 La Quantification de la langue

La langue peut se prêter à la quantification, c'est-à-dire que des concepts fondamentaux des mathématiques peuvent s'y retrouver. Quelle que soit la forme de la langue (littéraire, prosodique, poétique, utilitaire sous forme de documents publicitaires ou administratifs) elle peut fournir des données numériques. Le texte peut être divisé en paragraphes (ou en strophes), en phrases, en mots et en phonèmes qu'on peut compter. On peut aussi compter le nombre de verbes, de substantifs, d'articles, d'adjectifs, d'adverbes, de conjonctions et de prépositions aussi bien que le nombre de lettres.

Chacune de ces unités peut être caractérisée par l'étendue qu'elle a dans la chaîne du discours. Le texte peut alors être considéré comme une "population" et les fréquences comme des sous-ensembles de cette population.

Avant de procéder à une analyse statistique des oeuvres littéraires, on doit être en mesure de comprendre quelques notions mathématiques. Pour faciliter le raisonnement, la statistique utilise des symboles. (Ils sont indiqués dans le glossaire qui figure au début de cette étude à la page xxix). La statistique opère sur des données numériques qui sont soit *continues*, soit *discontinues*.

La taille et le poids des personnes sont des éléments qui permettent de les décrire quantitativement. Toute une gamme de valeurs est possible: par exemple, la taille d'une femme peut varier entre 165 et 166 cm et le poids peut varier par exemple entre 59 et 60 kilos. La taille et le poids de cette femme constituent alors des éléments

continus, c'est-à-dire que plusieurs valeurs existent entre 165 et 166 cm (165,1 cm; 165,256 cm, 165,3 cm et ainsi de suite.). Par contre, si cette femme est une mère de famille, le nombre de ses enfants peut varier entre 2, 3, 4 et 5 par exemple. Le nombre d'enfants constitue un élément *discontinu* et *discret* (aucune valeur intermédiaire ne pouvant exister). La mère de famille ne peut avoir 2,5 ou 4,7 enfants. Mais entre la taille de ses enfants, toute une série de valeurs intermédiaires est possible.

En ce qui concerne l'oeuvre littéraire, les critiques peuvent s'intéresser à la proportion de substantifs ou de verbes conjugués dans un texte, par exemple. Les éléments *discontinus* sont le nombre absolu des substantifs ou des verbes, telle que la fréquence absolue de 1.518 du vocable *de* dans *La Vie devant soi* (appendice 10.1.4). Sa fréquence relative de 23,32 constitue un élément *continu*.

La statistique comprend aussi les concepts suivants: la *moyenne*, la *variance*, l'*écart-type*, l'*écart réduit* et la *corrélation*.

9.7. La Moyenne

La moyenne est souvent représentée par le signe μ . La formule en est:

$$\mu = \frac{\sum (x_n)}{n}$$

Le signe Σ (sigma majuscule) indique la somme de tout ce qui suit, en ce cas-ci, la somme des valeurs; " x_n " signale une variable; "n" correspond au nombre de variables, et " μ " représente la moyenne des valeurs.

Si le chercheur veut calculer la longueur moyenne des vocables d'une oeuvre, il peut prendre en échantillon une phrase de dix mots et compter la longueur de chaque mot. Les dix mots peuvent être représentés par les symboles $x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9$ et x_{10} . Ensuite, il additionne ces valeurs et divise cette somme par dix ($x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9$ et $x_{10} \div 10$). Il obtient ainsi la longueur moyenne des mots.

Pour prendre un exemple plus concret, un critique peut s'intéresser à la longueur du mot typique dans les tragédies de Racine, parce qu'il postule que la longueur du vocable est un signe de richesse lexicale. En faisant une lecture attentive des tragédies, il arrive intuitivement à la conclusion que *Phèdre* et *Athalie* possèdent un lexique plus riche que *Mithridate* et *Britannicus*. S'il veut le contrôler de façon plus précise, il pourrait procéder comme Charles Muller qui a analysé la richesse lexicale dans les tragédies de Corneille (*Étude*).

Comme il s'agit de vers alexandrins chez Racine comme chez Corneille, chaque vers contient douze syllabes. Donc, s'il compte le nombre de vocables dans la pièce et qu'il le divise par le nombre de vers dans chaque pièce, il obtiendra le nombre moyen de mots par vers. D'autre part, s'il divise le nombre de syllabes du vers alexandrin (12) par le nombre moyen de mots par vers, il obtiendra la longueur moyenne du mot

exprimée en syllabes. Après quelques calculs, il peut déterminer si la longueur des vocables dans *Phèdre* et *Athalie* est effectivement supérieure à celle qu'il retrouve dans *Mithridate* et *Britannicus*.

Si le calcul de la moyenne fonctionne très bien pour un test de ce genre, il n'est pas toujours très efficace pour des cas plus compliqués, car la moyenne n'est qu'une caractérisation approximative de la réalité. Elle ne représente qu'une "sorte de centre de gravité de ses variations" (Muller, *Initiation* 62). Les valeurs les plus nombreuses se trouvent souvent groupées à cet endroit et se raréfient de plus en plus, à mesure que l'on s'éloigne de ce centre. Par ailleurs, deux groupes de variables peuvent avoir la même moyenne, mais des dispersions très différentes, comme le cas suivant le montre.

Le même critique est de l'avis que le nombre de vocables par phrase est une indication de richesse lexicale (c'est-à-dire que plus la phrase est longue, plus le style est riche). Il veut déterminer lequel de trois auteurs a le style le plus riche. En calculant la longueur moyenne des phrases pour ces auteurs, il constate qu'elle est la même pour tous les trois et s'élève à 10, qu'il a obtenue de la manière suivante.

Pour chaque auteur, ce critique a basé ses calculs sur trois de leurs romans. Dans le cas du premier auteur, le nombre de vocables par phrase dans son premier roman s'élève à 8 en moyenne, elle est de 13 en moyenne dans le deuxième roman, et de 9 dans le troisième roman, faisant un total de 30 et une moyenne de 10 pour les trois romans. Pour le deuxième auteur, le nombre de vocables par phrases dans son premier

roman est de 2 en moyenne, dans le deuxième roman ce nombre est de 12 en moyenne. Il est de 16 en moyenne dans le troisième roman, ce qui fait encore une fois un total de trente et une moyenne de 10. Pour le troisième auteur il a obtenu les résultats suivants: 15, 8 et 7 faisant encore une fois un total de 30 et une moyenne de 10. Cette moyenne commune de dix vocables par phrase ne caractérise pas ces trois auteurs du point de vue de l'homogénéité.

9.8. La Variance (v) et l'écart-type (σ)

Si le critique effectue le calcul de la variance (qui est un indice de dispersion autour de la moyenne) entre la longueur des phrases chez les trois auteurs, il obtiendra une meilleure interprétation de la chose.

Comme il vient d'être démontré dans le cas de ces trois auteurs, trois distributions ayant la même moyenne peuvent être l'une très étalée, l'autre très resserrée. Dans la première, les écarts (ou les intervalles entre chaque valeur et la moyenne) pris en valeur absolue seront grands, dans la seconde la moyenne de ces écarts sera faible. Pour trouver la variance, on calcule l'écart, c'est-à-dire la différence entre chaque valeur et la moyenne, aussi bien que le carré de chaque écart (étant donné que quelques valeurs peuvent être négatives). Par exemple si la moyenne est 50, et qu'il existe une donnée de 40 on aura là un écart de -10, mais son carré est 100.

Tous les écarts carrés sont ajoutés afin d'obtenir la somme des carrés. Cette nouvelle somme est divisée par le nombre de valeurs, permettant d'obtenir la variance. La variance est donc égale à la somme des carrés des écarts de la moyenne, somme qui est divisée par le nombre de mesures (n). Par une attitude conservatrice, le nombre de mesures est n moins 1 (n-1) dans des corpus de petite étendue. La somme des écarts est toujours zéro. On appelle la moyenne des écarts carrés la variance. La variance est indiquée selon la formule suivante:

$$v = \frac{\sum (x_n - \mu)^2}{(n-1)}$$

μ la moyenne

x une des données qui entrent dans le calcul de la moyenne

$x_n - \mu$ un écart entre une donnée et la moyenne

n le nombre de variables

Σ la somme de ce qui suit

Si on reprend le cas des trois écrivains, le calcul de la variance pour eux peut être fait selon la formule indiquée ci-dessus. En calculant la moyenne pour chacun, le critique avait déjà trouvé 10. Il soustrait alors la moyenne de chaque observation, tel qu'indiqué à l'étape (a) de la table 9.2 ci-dessous.

Table 9.2

Calcul de la variance pour les trois auteurs

auteur 1	(a)	8-10=-2	;	13-10=3	;	9-10=1	;
	(b)	(-2)(-2)	+	(3)(3)	+	(-1)(-1)	;
	=	4	+	9	+	1	= 14
auteur 2	(a)	2-10=-8	;	12-10=2	;	16-10=6	;
	(b)	(-8)(-8)	+	(2)(2)	+	(6)(6)	;
	=	64	+	4	+	36	= 104
auteur 3	(a)	15-10=5	;	8-10=-2	;	7-10=-3	;
	(b)	(5)(5)	+	(-2)(-2)	+	(-3)(-3)	;
	=	25	+	4	+	9	= 38
(c) auteur 1 (14/2) = 7 ; auteur 2 (104/2) = 51 ; auteur 3 (38/2) = 19							

La deuxième étape consiste à élever au carré chaque total obtenu à la phase (a) de l'opération afin d'en trouver la somme. Ceci est indiqué à la phase (b) de l'opération. On procède ainsi pour chaque auteur. Par la suite on applique la formule

$$v = \frac{\sum(x_n - \mu)^2}{(n-1)}$$

Comme on sait déjà que $n-1$ est égal à 2, on peut alors calculer la variance telle qu'elle est indiquée dans la partie (c).

La variance pour le premier auteur est 7, elle s'élève à 51 pour le deuxième et à 19 pour le troisième. La variance du deuxième auteur est plus grande que celles des deux autres, c'est-à-dire qu'il y a chez lui plus de dispersion autour de la moyenne. Le calcul de leur moyenne commune (10) caractérise moins ces trois auteurs du point de

vue de l'homogénéité que leurs écarts ou la variabilité entre la longueur des phrases. La moyenne de 10 ne tient pas compte de ces différences.

Il existe néanmoins quelques limitations au sujet de l'utilisation de la variance. Dans bien des cas, elle introduit une anomalie, comme dans le cas du poids d'une personne qui s'indique en kilogrammes carrés (kg^2) et la taille en centimètres carrés (cm^2). Par ailleurs, elle est rarement utilisée en statistique littéraire, surtout à cause de l'utilité de l'écart-type. L'écart-type, qui est une autre indice de dispersion autour de la moyenne est intéressante parce qu'on peut la situer sur une échelle de probabilités. Pour calculer l'écart-type, on cherche la racine carrée de la variance qui est indiqué comme suit : $\sigma = \sqrt{v}$.

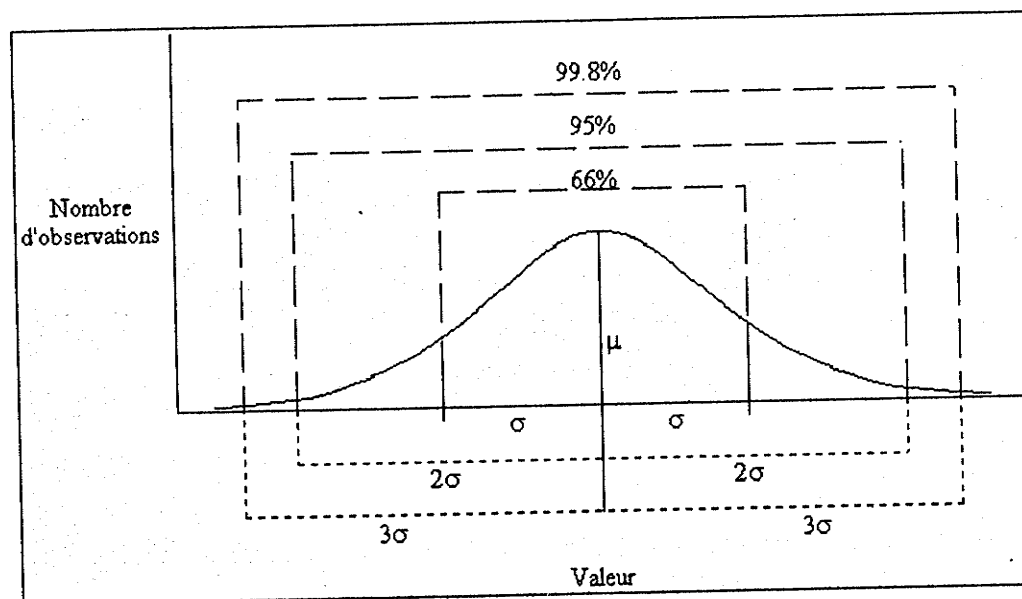
9.9. La Distribution gaussienne

Souvent les données se disposent selon la distribution gaussienne ou normale. Lorsque c'est le cas et qu'il existe un grand nombre d'observations, on trouve que 66% des valeurs sont comprises entre un écart-type de part et d'autre de la moyenne. On trouve aussi 95% des valeurs à deux écarts-types de part et d'autre de la moyenne et 99,9% des valeurs à trois écarts-types de part et d'autre de la moyenne. Ce que l'on vient de décrire peut être représentée par une courbe en cloche tracée sur les axes vertical et horizontal (Voir table 9.3).

Le symbole μ est utilisé pour la moyenne et le symbole σ pour l'écart-type. Pour toute valeur sur ces deux axes, la table de la distribution gaussienne (voir un manuel de statistique) donne l'aire comprise entre la moyenne et le point correspondant à cette valeur. Le centre de symétrie est μ , c'est ici que la courbe atteint son maximum. Il en découle que la médiane (la valeur qui occupe la position centrale) et la moyenne sont tous égaux à μ .

Table 9.3

La Courbe gaussienne



Comme il s'agit d'une courbe symétrique, on trouve 33% des valeurs entre la moyenne et un écart-type d'un côté ou de l'autre de la moyenne (entre μ et $-\sigma$, ou $\mu + \sigma$); entre la moyenne et deux écarts types (2σ) d'un côté ou de l'autre de la moyenne

on trouvera 47,7% des valeurs. D'un côté ou de l'autre de la moyenne, à trois écarts-types (3σ) se trouve 49,9% des valeurs. Par ailleurs, 13,59% des valeurs se situent entre un et deux écarts-types. On peut ainsi calculer la probabilité d'une valeur se situant entre l'intervalle 2 (2σ) et l'intervalle 3 (3σ) à 0,0215 (c'est-à-dire 0,4987-0,4772 ou 0,0215), ce qui indique qu'il existe 2,15% de chances d'obtenir une telle valeur.

L'exemple des notes obtenues lors d'un examen passé par 70 étudiants peut illustrer ce concept. La moyenne est 65%, l'écart type est de 13. Et pourtant on observe aussi des notes de 94% et même de 25%. On peut estimer la probabilité d'obtenir ces deux valeurs en faisant le calcul suivant:

$$Z = \frac{(x_i - \mu)}{\sigma}$$

Z l'écart réduit

x_i l'observation en question (94% ou 25%)

μ la moyenne (65%)

σ l'écart-type (13)

En faisant les calculs pour Z , on obtient 2,23 pour une note de 94%, et pour une note de 25% on obtient -3,077. Ceci veut dire qu'il y a une probabilité de 0,04868 (4,8%) d'obtenir une note de 94% et une probabilité de 0,04989 (4,9%) d'obtenir une note de 25%.

On doit ajouter cependant que, dans le cas de la courbe gaussienne, il s'agit surtout d'une interprétation probabiliste car cette courbe est la définition d'un objet idéal qui est rarement rencontrée dans le domaine de la stylométrie. Dugast la compare à la Vénus de Praxitèle qui est le résultat de plusieurs dessins faits par Praxitèle des jeunes filles d'Athènes. Elles étaient toutes très belles mais avaient quand même un défaut chacune. Praxitèle a su faire de toutes ces beautés une beauté unique, une beauté idéale sans défaut (35). La courbe gaussienne opère sur le même principe. Elle représente la courbe idéale qu'on ne rencontre presque jamais dans la réalité.

9.10. Le Test- t de Student-Fisher

La courbe gaussienne décrit les grandes populations et ne s'applique que lorsque l'échantillon est grand et dépasse soixante observations. Mais lorsque l'échantillon sur lequel on travaille est moins de 30, on doit utiliser le test- t de Student-Fisher. Encore une fois, il s'agit d'une courbe en cloche symétrique quoiqu'elle ne ressemble pas exactement à la courbe gaussienne. Le calcul de la moyenne de cette distribution s'effectue de la même manière que le calcul de la moyenne de grands échantillons, mais l'écart-type n'est pas semblable à celui d'une grande distribution.

On peut utiliser la formule de Student-Fisher pour une population de référence qui n'est pas normale. Pour interpréter les résultats obtenus, il faut aussi consulter la table du " t " de Student-Fisher dans un manuel de statistique. Une caractéristique

importante de cette table est qu'elle est basée sur le nombre de degrés de liberté (d.d.l.) qui est déterminé en fonction de la taille de l'échantillon (à savoir, le nombre de valeurs moins un). Pour un échantillon de 9 par exemple, le nombre des degrés de liberté est de 8.

La table-*t* de Student-Fisher est constituée de plusieurs colonnes: la colonne à gauche donne le nombre de degrés de liberté. Les autres colonnes indiquent le seuil de rejet ou le risque d'erreur à 0.05 (5%), 0.025 (2,5%), 0.01 (1%) et 0.005 (0,5%). D'après Muller (*Initiation* 43) un risque d'erreur de 0,05 (cinq pour cent) est généralement admis dans le domaine de la linguistique. Ceci veut dire qu'on a une chance sur vingt (ou cinq chances sur cent) de se tromper.

Si on veut isoler 5% des distributions ayant une taille de 9, en consultant la table-*t* de Student-Fisher, on cherche la colonne où le seuil de rejet est 0,05 (5%). Sous la colonne "ddl." on cherche 8. On trouve alors le chiffre 1,860 (à huit degrés de liberté).

Mais la limite de confiance choisie est souvent une question de jugement. Si on veut être plus prudent, on admet le risque d'erreur de 0,01, ce qui veut dire qu'il y a une chance sur cent de se tromper; cela produit des conclusions plus rares mais plus solides. Grâce au seuil défini, il est possible d'évaluer le degré d'assurance associé à une interprétation.

Le test *t* de Student-Fisher permet de tester une hypothèse lorsque l'échantillon est inférieur à soixante. Comme le corpus qui figure dans cette étude est constitué de

huit romans et du corpus d'Engwall, ce test est utilisé. Il permet d'estimer la moyenne de la population des romans écrits au vingtième siècle, étant donné qu'on ne possède des renseignements que sur ces huit romans et sur le corpus d'Engwall.

9.11. Le Test du khi-deux (χ^2)

La fréquence du type *de* varie entre 1.363 (fréquence relative de 33,03) et 2.290 (fréquence relative de 39,54) pour chaque texte sauf pour *La Vie devant soi* où elle est de 1.518 ayant une fréquence relative de 23,32 (voir appendices 10.1.1 à 10.1.9). Plusieurs autres types dans *La Vie devant soi* présentent des écarts semblables. La question est de savoir si les valeurs observées (1.518 pour le vocable *de* dans *La Vie devant soi*, par exemple) sont significativement différentes des autres valeurs pour ces mêmes types. Pour répondre à cette question, il suffit d'obtenir une valeur calculée pour *de* (valeur qui doit être déterminée à l'avance) et d'appliquer le test du χ^2 .

La valeur calculée peut être représentée soit par le corpus d'Engwall qui reflète un vocabulaire contemporain, soit par les deux romans de Gary (*Clair de Femme* et *Au-delà de cette limite votre ticket n'est plus valable*) qui servent d'échantillon du style de Gary. La formule utilisée pour le test est:

$$\chi^2 = \frac{\sum (O_i - C_i)^2}{C_i}$$

La lettre " C_i " représente une valeur calculée ou théorique (telle que 18.215 pour le vocable *de* dans le corpus d'Engwall) et " O_i " représente une valeur observée (telle que 1.518 pour le vocable *de* dans *La Vie devant soi*). Le χ^2 est la somme des écarts entre les valeurs observées et les valeurs calculées. Ces écarts sont ensuite élevés au carré $(O_i - C_i)^2$ et la nouvelle valeur est divisée par la valeur théorique $(O_i - C_i)^2 / C_i$, ce qui donne alors une valeur du χ^2 . Il faut savoir interpréter ces valeurs.

Pour cela il faut lire la table du χ^2 dans un manuel de statistique. Cette table est constituée de plusieurs colonnes. La première colonne de la table, à gauche, indique le nombre de degrés de liberté. Les colonnes suivantes donnent le seuil de rejet à 0,1 (10%), 0,05 (5%), 0,02 (2%), 0,01 (1%) et 0,001 (0,1%). Le nombre recherché se trouve à l'intersection de la ligne correspondant au nombre de degrés de liberté, et de la colonne correspondant au seuil de probabilité choisi.

Pour lire cette table le principe est le même que pour la lecture de la table *t* de Student-Fisher, c'est-à-dire qu'il faut respecter les contraintes imposées par les degrés de liberté (qui est $n-1$). Si la valeur du χ^2 calculée est inférieure à la valeur indiquée dans la table, on conclut que la différence entre la valeur observée et la valeur calculée est négligeable. Si elle est supérieure à la valeur indiquée dans la table, on accepte l'hypothèse inverse, c'est-à-dire que les résultats obtenus révèlent un phénomène significatif.

À un degré de liberté toute valeur entre 0 à 3,83 peut être considérée comme ayant peu d'importance. Toute valeur entre 3,84 et 6,62 est significative et a seulement 5% de probabilité de résulter du hasard. Toute valeur entre 6,63 et 10,82 n'a qu'une chance sur cent d'apparaître par pur hasard et toute valeur s'élevant au-dessus de 10,83 est très significative, ayant une chance sur mille d'apparaître par hasard.

Pour illustrer le calcul de ce test, on prendra comme exemple celui que propose Muller. Dans son analyse du style de Corneille, Muller calcule que le pourcentage de substantifs dans la pièce *L'Illusion Comique* s'élève à 18%. Il croit intuitivement que le rôle de Matamore qui est constitué de 2.438 vocables se caractérise par un grand nombre de substantifs. Si 18% du dialogue de Matamore est consacré aux substantifs, il devrait en contenir 438 (*Initiation* 120), mais Muller en rencontre 46 substantifs de plus (donnant un total de 484).

Afin de déterminer si cette différence est statistiquement significative, Muller fait les calculs du χ^2 pour le rôle de Matamore et obtient une valeur de 5,889. La table du χ^2 montre qu'à un degré de liberté, la valeur 5,41 est indiquée pour 0,02; et la valeur 6,63 est indiquée pour 0,01. Étant donné que le rôle de Matamore produit un χ^2 de 5,889, on peut s'assurer avec moins de deux chances sur cent de se tromper que le rôle de Matamore se caractérise par son grand nombre de substantifs (*Initiation* 120).

Les résultats n'expriment pas une chose absolue, mais ils expriment les degrés de relation entre la valeur calculée et la valeur observée. Ce test ne s'applique qu'à des valeurs absolues et non pas à des valeurs relatives, comme le pourcentage, par exemple.

Une manière d'effectuer le test du χ^2 est par le biais du test d'adéquation. Pour obtenir les valeurs hypothétiques, on prend le corpus d'Engwall comme source. Pour le vocable *de* dans *Clair de femme* par exemple, le nombre absolu du vocable *de* dans le corpus d'Engwall (118.215) est divisé par 500.000 (le nombre total de vocables dans le corpus d'Engwall) et ensuite multiplié par 33.755 (le nombre total de vocables dans *Clair de femme*). La valeur 1.229,69 est alors obtenue. En procédant de la sorte pour tous les autres vocables, on obtient les valeurs indiquées dans l'appendice 11.2. À partir de ces valeurs calculées, il est possible de comparer le vocabulaire des textes au vocabulaire du corpus d'Engwall, et d'effectuer le test du χ^2 . Pour le type *de* dans *Clair de femme* par exemple, on obtient un χ^2 de 6,11 (voir appendice 11.3).

La manière traditionnelle d'effectuer le test du χ^2 est par la méthode qu'on appelle la "table de contingence". Comme d'habitude, une valeur observée est comparée à une valeur hypothétique. Pour obtenir les valeurs théoriques, on additionne toutes les valeurs qui figurent dans les colonnes et dans les rangées. (Voir table 9.4).

Table 9.4

Valeurs observées et table de contingence

TYPE	<i>Clair</i>	<i>Tick</i>	<i>Câl</i>	<i>Vie</i>	
de	1.143	2.290	1.899	1.518	6.850
et	736	1.459	1.294	1.989	5.478
à	641	1.082	1.046	1.336	4.105
un	558	891	704	868	3.021
que	474	863	816	1.103	3.256
d'	391	749	593	420	2.153
	3.943	7.334	6.352	7.234	24.863

La première rangée indique la fréquence du vocable *de* dans les quatre romans de Gary; la première colonne indique les vocables dans *Clair de femme* et ainsi de suite. En ajoutant toutes les occurrences du vocable *de* dans la première rangée, la somme de 6.850 est obtenue, et, en additionnant toutes les occurrences dans la première colonne le total de 3.943 est obtenu. En procédant de la sorte pour toutes les autres rangées et les autres colonnes, et en ajoutant la somme des occurrences dans la dernière colonne (ou dans la dernière rangée), on obtient la somme de 24.863.

À partir de ces nouvelles valeurs, on peut obtenir les valeurs hypothétiques. Pour le vocable *de* dans *Clair de femme* par exemple, on multiplie la somme 3.943 par 6.850 qu'on divise ensuite par 24.863. Ceci donne alors 1.086,34. En procédant de la sorte pour toutes les autres occurrences, on obtient les valeurs indiquées dans la table 9.5.

Table 9.5**Valeurs calculées**

TYPE	<i>Clair</i>	<i>Tick</i>	<i>Câl</i>	<i>Vie</i>
de	1.086,34	2.020,59	1.750,04	1.993,04
et	868,75	1.615,88	1.399,52	1.593,85
à	651,01	1.210,88	1.048,75	1.194,37
un	479,10	891,12	771,81	878,97
que	516,37	960,44	831,84	947,35
d'	341,44	635,08	550,05	626,42

En appliquant la formule du χ^2 , on obtient les résultats qui figurent dans la table 9.6.

Table 9.6**Valeurs du χ^2**

TYPE	<i>Clair</i>	<i>Tick</i>	<i>Câl</i>	<i>Vie</i>
de	2.96	35.92	12.68	113.22
et	20.29	15.23	7.96	97.97
à	0.15	13.72	0.01	16.80
un	12.99	0.00	5.96	0.14
que	3.48	9.89	0.30	25.57
d'	7.19	20.43	3.35	68.02
χ^2	47.06	95.19	30.26	321.72

Il faut maintenant appliquer la correction de Bonferroni à ces tests, étant donné qu'il existe une multiplicité de tests. Pour cela, il suffit de diviser le seuil de rejet de 5%, par le nombre total de tests qui est au nombre de 28, ce qui donne un nouveau seuil de rejet de 0,17857. À 28 degrés de liberté, la valeur significative du χ^2 au seuil de rejet de

0,17857 est de 48,278. Ainsi toutes les valeurs au-dessus de 48,278 sont significatives; elles sont indiquées en chiffres gras dans la table 9.6.

9.12 Le Coefficient de corrélation de Pearson

Le concept de corrélation est très important dans l'étude statistique des textes littéraires, concept qui sera illustré par les exemples suivants. La taille et le poids d'une population d'étudiants constituée de neuf personnes du même sexe sont tels qu'indiqués dans la table 9.7.

Table 9.7

Rapport entre le poids et la taille (a)

Rang	Taille	Poids
1	5'5"	140 livres
2	5'6"	145 livres
3	5'7"	150 livres
4	5'8"	155 livres
5	5'9"	160 livres
6	5'10"	165 livres
7	5'11"	170 livres
8	6'0"	175 livres
9	6'1"	180 livres

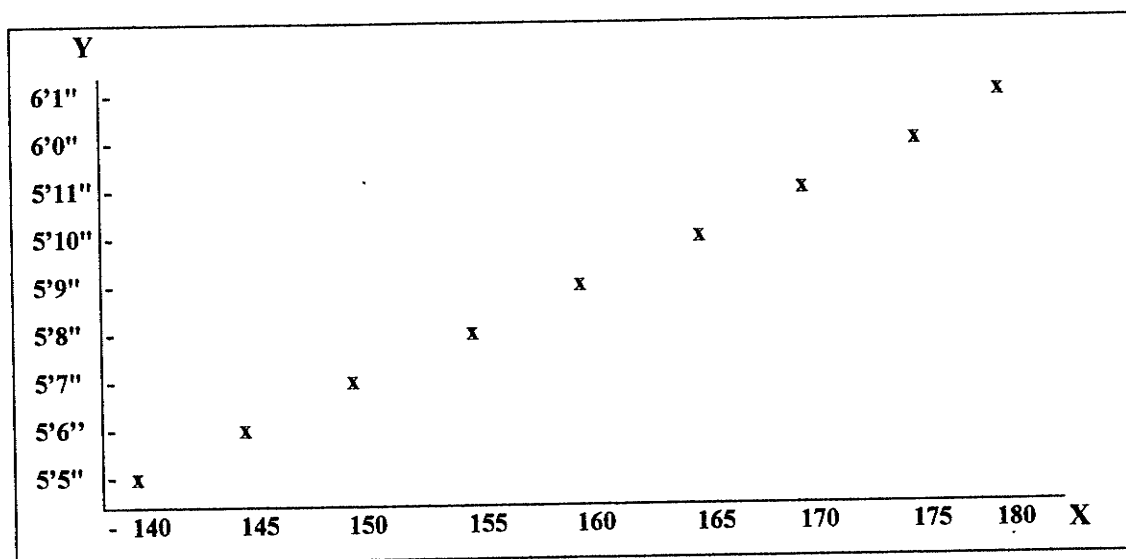
Une corrélation parfaitement positive s'observe entre le poids et la taille de ces neuf étudiants. À mesure que la taille augmente, le poids augmente aussi. Le moins grand est aussi celui qui pèse le moins. Mais lorsqu'on analyse ces correspondances

plus étroitement, l'on constate qu'à mesure que la taille augmente d'un pouce, le poids augmente de cinq livres; chaque pouce supplémentaire correspond à cinq livres.

Non seulement existe-t-il une parfaite corrélation entre le poids et la taille de cette population, mais encore cette corrélation est d'une nature spéciale, c'est-à-dire que la correspondance entre un pouce et cinq livres est uniforme à travers toutes les observations. Il va de soi qu'une telle observation ne peut être que théorique. Cette corrélation peut néanmoins être représentée par la table 9.8.

Table 9.8

Corrélation entre le poids et la taille (a)



Il existe une corrélation positive à cent pour cent entre les variables X et Y. Ainsi r est-il égal à +1. La réalité est cependant que les correspondances ne sont presque jamais parfaites: elles sont plutôt relatives. Le cas suivant en sert d'exemple.

Table 9.9

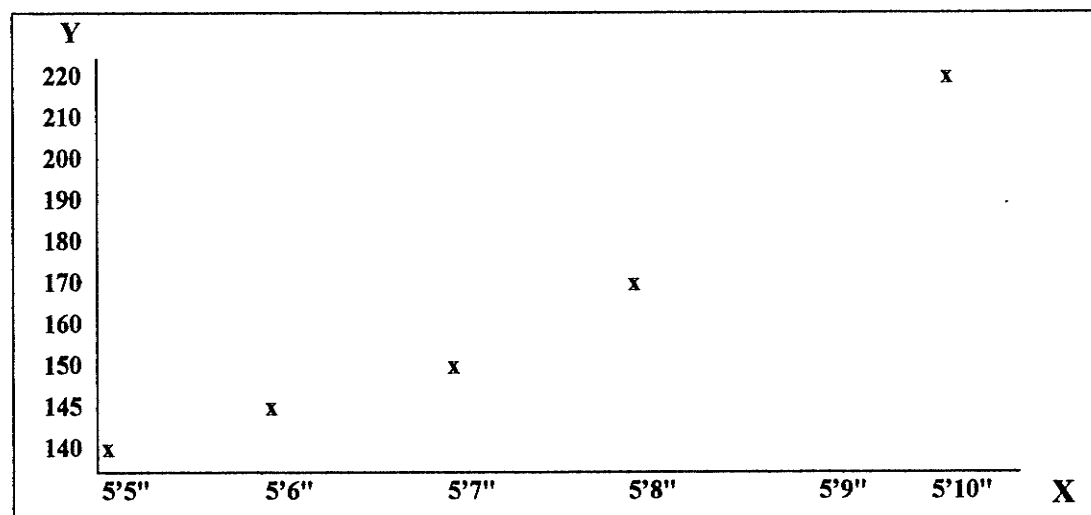
Rapport entre le poids et la taille (b)

Rang	Taille	Poids
1	5'5"	140 livres
2	5'6"	145 livres
3	5'7"	150 livres
4	5'8"	170 livres
5	5'10"	220 livres

La corrélation entre les observations n'est parfaite qu'entre les trois premières observations où chaque pouce supplémentaire correspond à cinq livres. De la troisième à la quatrième observation, chaque pouce supplémentaire correspond à vingt livres. De la quatrième à la cinquième observation deux pouces supplémentaires correspondent à cinquante livres. La table 9.10 présente une indication de cette correspondance.

Table 9.10

Corrélation entre le poids et la taille (b)



Il est également possible de vérifier la corrélation entre le poids de quelques personnes obèses et leur flexibilité. La flexibilité et le poids de cette population sont indiqués dans la table 9.11.

Table 9.11

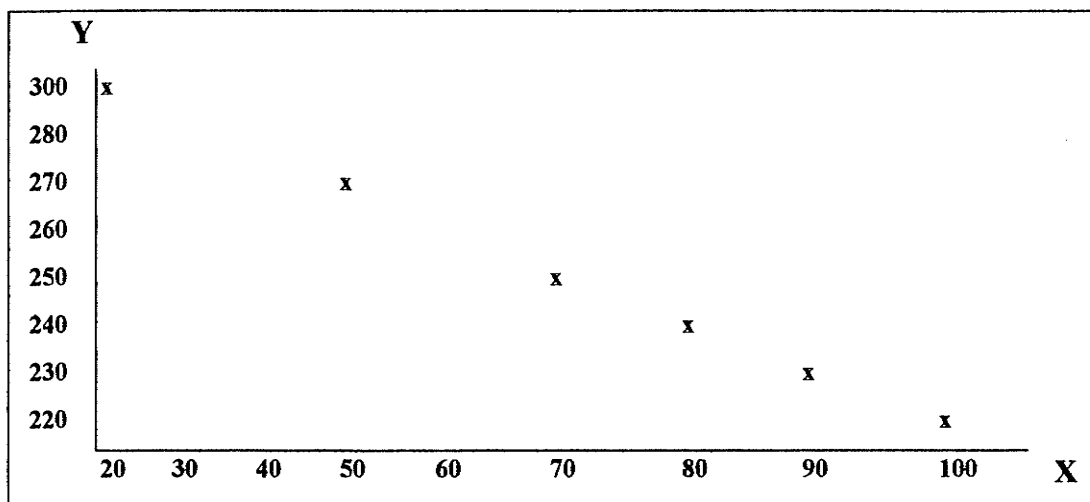
Rapport entre le poids et la flexibilité

Rang	Poids	Flexibilité
1	220 livres	100
2	230 livres	90
3	240 livres	80
4	250 livres	70
5	270 livres	50
6	300 livres	20

Il existe une parfaite corrélation inverse entre les deux variables, à savoir que moins on est obèse plus on est souple. Du premier au troisième rang chaque dix livres supplémentaires équivaut à une réduction de flexibilité de dix. Du quatrième au cinquième rang, il existe une augmentation de vingt livres et aussi une réduction de flexibilité. Cette corrélation peut être représentée en forme graphique de la façon suivante:

Table 9.12

Corrélation entre le poids et la flexibilité



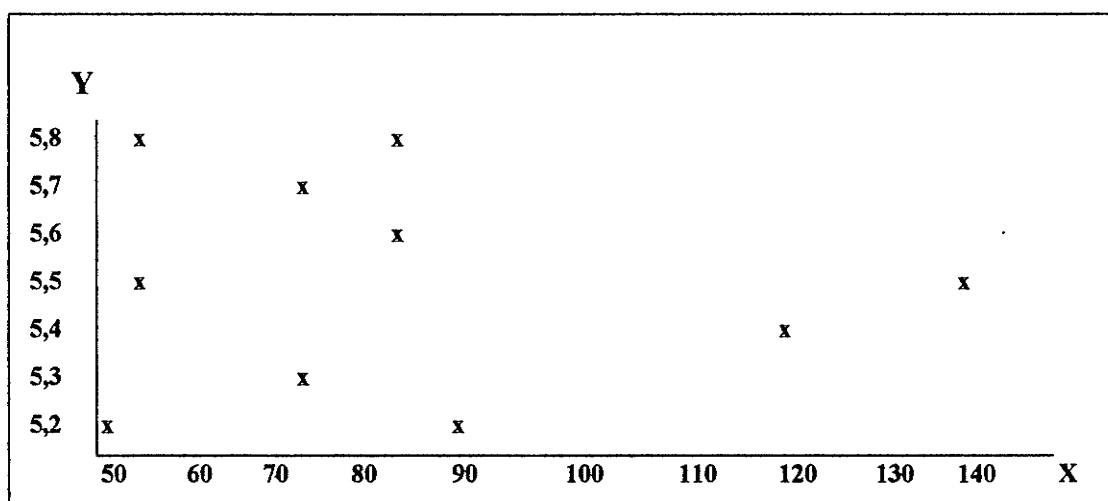
Cette fois-ci l'augmentation d'une variable entraîne la diminution de l'autre. À mesure que Y diminue X augmente ce qui veut dire que $r = -1$. Il y a donc une corrélation négative à cent pour cent entre ces deux variables.

Lorsqu'il n'existe aucune corrélation entre les deux variables X et Y , c'est-à-dire lorsqu'une variable ne diminue ni n'augmente en relation avec l'autre, la corrélation est de zéro. La corrélation entre la taille des gens du même sexe et du même âge et leur aptitude linguistique permet d'illustrer ce concept. En prenant une population de dix personnes et en leur faisant passer une épreuve dont la note maximale s'élève à 140, on obtient les résultats suivants:

Table 9.13**Rapport entre la taille et l'aptitude linguistique**

Rang	Taille	Aptitude linguistique
1	5'2"	90
2	5'2"	50
3	5'3"	75
4	5'4"	120
5	5'5"	140
6	5'5"	50
7	5'6"	85
8	5'7"	75
9	5'8"	55
10	5'8"	85

La table 9.14 donne une représentation graphique de ces notions. La corrélation est établie ici entre les classements selon la taille (l'axe vertical) et selon l'aptitude linguistique (l'axe horizontal).

Table 9.14**Corrélation entre la taille et l'aptitude linguistique**

Il est évident qu'aucune corrélation ne se révèle entre une valeur et une autre; X ne diminue ni n'augmente en fonction de Y .

Un mot de prudence s'impose: lorsqu'une corrélation a été calculée, il est important de savoir si elle représente une relation linéaire simple, c'est-à-dire si le résultat obtenu diffère significativement de zéro (la corrélation nulle). Lorsqu'il ne diffère pas de zéro, on ne peut pas dire que la relation linéaire est significative. La table des valeurs significatives de r permet de résoudre ce problème.

Cette table se trouve dans n'importe quel manuel de statistique. Elle est à double entrée. La première colonne N indique le nombre de valeurs analysées. En tête des colonnes suivantes se trouvent les seuils de probabilité (à 0,10; 0,05; 0,02 et 0,01). Cette table indique les valeurs qui devraient être dépassées par la corrélation observée pour que cette dernière diffère d'une manière significative de zéro.

Par exemple, dans la table 12.1 on observe une corrélation de 0,9538 entre *Clair de femme* à *Au-delà de cette limite votre ticket n'est plus valable*. Le nombre de vocables sur lequel on a travaillé est de soixante (c'est-à-dire que $n=60$). En consultant la table des valeurs significatives de r on constate que 60 ne figure pas dans la liste, ce qui veut dire qu'on doit choisir entre 52 et 62. On a choisi 62 étant donné que ce chiffre se rapproche le plus de 60. Comme presque toutes les corrélations se situent au-dessus de 0,32, on peut déclarer qu'elles sont significatives.

Une corrélation de 0,7100 semble à première vue forte et sérieuse. Elle est plus proche de la corrélation parfaite (+1), que de l'absence de corrélation (-1) ou de la corrélation nulle (0). Mais ceci n'a de sens qu'en fonction du nombre de valeurs avec lesquelles on a travaillé. Par exemple, si on avait obtenu un r égalant à 0,7100, et que le nombre avec lequel on travaille est de sept, la table montre que cette corrélation ne diffère pas significativement de zéro même au seuil de 0.05.

9.13. Le r^2 (le Coefficient de détermination) et les limites de confiance de " r "

Une corrélation de 0,9558 entre *L'immoraliste* et *La Porte étroite* lorsque $n=60$ (table 12.1) et 0,9767 (table 12.2) semble significative. Mais, pour obtenir une meilleure mesure de la similarité entre ces deux romans, on doit calculer les coefficients de détermination (r^2) à partir de ces corrélations.

Le coefficient de détermination est le coefficient de corrélation (r) élevé au carré. Ces nouvelles valeurs sont alors 0,9136 ($0,9558^2$) lorsque $n=60$, et 0,9539 ($0,9767^2$ lorsque $n=31$). Une corrélation de 0,9558 permet d'expliquer 91,36% de la variabilité linéaire (lorsque $n=60$) et une corrélation de 0,9767 permet d'expliquer 95,39% de la variabilité linéaire (lorsque $n=31$).

On doit aussi considérer les corrélations obtenues comme un échantillon de la population en général. Il faut alors savoir à quelle limite se trouve la valeur exacte du

coefficient de corrélation pour l'ensemble de la population d'où provient l'échantillon.

Pour cela, il faut calculer les limites de confiance de ces corrélations.

9.14. Conclusions en matière de corrélation

Lorsqu'il y a une corrélation entre deux variables telles que la taille et le poids, son coefficient varie entre -1 et +1. S'il existe une corrélation positive parfaite le coefficient est +1, s'il existe une corrélation négative parfaite, le coefficient est -1. Dans les cas concrets, presque aucune corrélation linéaire n'est parfaite. Lorsqu'il existe une corrélation de 0,8753, cela indique que 0,76615 ($0,8753^2$) de la variation linéaire est représentée par le coefficient et 23,38% ne l'est pas.

Le prochain chapitre applique les notions statistiques qui viennent d'être présentées afin de faire une analyse du style de Gary/Ajar.

Chapitre 10

Analyses Préliminaires

Rien n'est négligeable dans une oeuvre. C'est peut-être dans la zone souterraine de sa conscience linguistique, et dans les ramifications enfouies de ses habitudes, qu'un auteur cache les traits originaux de son style. (Brunet, 399)

Plusieurs experts de la littérature postulent que certains aspects du style sont invariables chez un auteur, ce qui implique que consciemment ou inconsciemment, tout auteur qui écrit sous un pseudonyme laisse des traces de sa personnalité dans ses oeuvres, traces qu'un chercheur assidu peut déceler. (Voir Brinegar, Mosteller et Wallace, Smith, Kjetsaa.)

Ce chapitre aussi bien que les trois chapitres qui suivent évaluent ce postulat en l'appliquant au style de Gary dans les quatre romans mentionnés dans l'introduction, et cela par le biais des tests statistiques suivants: le test-*t* de Student-Fischer, le test du χ^2 , et la corrélation de Pearson.

10.1. Méthodologie

Dans *Structure du langage poétique* Jean Cohen recommande que pour toute étude statistique littéraire le corpus se conforme à deux règles: être suffisamment étroit pour ne pas décourager la recherche et suffisamment large pour permettre l'induction (87). Plus le choix est large, plus le corpus a des chances d'être représentatif. Pour satisfaire à la première règle, cette étude comprend neuf sources de données: quatre romans de Romain Gary, quatre autres romans d'auteurs contemporains auxquels s'ajoute le corpus d'Engwall. Le chapitre 3 indique les raisons pour le choix de ces textes.

Les quatre autres romans qui font partie de l'étude sont *Le Noeud de Vipères* (1932), de Mauriac, *L'Étranger* (1942) de Camus, *L'Immoraliste* (1902) et *La Porte étroite* (1909) de Gide. Ils sont presque de la même longueur que les quatre romans de Gary et dans chacun, le narrateur est homodiégétique, ce qui revient à dire que l'emploi du pronom personnel *je* est privilégié.

Le nombre de textes étudiés est restreint, mais la totalité du corpus est suffisamment large pour permettre l'induction, le nombre total de vocables s'élevant à environ 800,000. Le corpus répond alors à la deuxième exigence de Cohen mentionnée ci-dessus. Je considère les neuf textes dans leur intégralité, ce qui veut dire que je ne travaille pas sur des échantillons et ne me fie pas au principe de sondage, car pour cela il aurait fallu savoir sur quelle partie du texte travailler. Le texte (et non pas des échantillons du texte) est considéré comme un ensemble, et les dialogues (y compris ceux qui sont indiqués entre guillemets ou par des tirets) ne sont pas isolés des parties narratives.

Quoiqu'il existe plusieurs façons d'aborder une analyse statistique du style, mon étude se concentre particulièrement sur la fréquence des vocables, étant donné que "le matériau premier de l'écrivain [...] (est) le vocabulaire, et son ouvrage à lui, l'organisation en discours de ce vocabulaire" (Dugast 103).

10.2. Traitement des données

Étant donné que le genre d'analyse que j'envisage nécessite un traitement des textes par ordinateur, les quatre romans de Romain Gary ont été traités par un lecteur optique et transcrits sur des disquettes en une forme lisible par machine. Ensuite, la version générée par ordinateur a été vérifiée, pour s'assurer qu'elle corresponde au texte imprimé, et au cas contraire les corrections qui s'imposent ont été faites. Par

exemple le vocable *Boeing* (*Ticket 7*), qui avait été accidentellement transcrit comme *Being*, a été corrigé en la forme requise.

Il n'était pas nécessaire de faire traiter les autres romans étudiés (*L'Etranger*, *L'Immoraliste*, *La Porte Étroite* et *Le Noeud de Vipères*), étant donné qu'ils sont disponibles en ligne à la banque des données du *Trésor de la langue française*, surtout le fonds ARTFL. Pour le corpus d'Engwall, *Fréquence et distribution du vocabulaire dans un choix de romans français* a été consulté.

Après l'obtention des données, la représentation graphique de quelques mots apparaissant dans les neuf textes a été modifiée. Par exemple, la forme des vocables *d'ailleurs*, *quelqu'une* a été transcrite de la manière suivante: *d_ailleurs*, *quelqu_une* pour faire de sorte que ces vocables n'apparaissent pas comme deux occurrences chacune, à savoir, *d'* et *ailleurs*; *quelqu'* et *une*. Les mots composés, tels que le bar *Clapsy's*, *peut-être*, *au-dessus*, *en-dessous* ont aussi été transformés de cette façon.

Après ce stade de préparation des textes, le logiciel *Oxford Concordance Package* (O.C.P.) a regroupé toutes les occurrences graphiquement identiques. Les types *que* sont regroupés ensemble, qu'il s'agisse du pronom relatif, du pronom interrogatif ou de la conjonction. Ce même principe s'applique aux types *qui*, *de*, *en*, *la*, *le*, et *les*. Par contre, les formes comme *de* et *d'* sont traitées séparément. Ceci s'applique aussi aux mots *le* et *l'*, *la* et *l'* et *ce* et *c'*, *ne* et *n'* etc. Les mots tels que *du* et *des*, *au* et *aux* ne sont pas décomposées en *de* et *le* (ou *les*), *à* et *le* (ou *les*).

Les homonymes (les unités lexicales distinctes dont la forme graphique ou sonore est accidentellement la même, mais qui n'ont pas de sens commun et n'appartiennent pas aux mêmes classes de vocables) ont aussi été regroupés ensemble. Les homonymes homographes, c'est-à-dire des formes qui s'écrivent de la même façon mais qui ne se prononcent pas de la même façon (par exemple "je vis" et "un vis") n'ont pas été séparés non plus. Dans la plupart des cas, les locutions figées sont traitées en fonction de leurs mots constitutifs: *afin de*, *tandis que* font deux vocables chacun.

À part ces quelques exceptions, chaque vocable est traité tel qu'il paraît dans le texte original. Il est vrai que Dugast suggère qu'on réduise les vocables à la forme que l'on retrouve dans les dictionnaires (57), mais cette méthode exige l'utilisation d'autres logiciels et plus de temps pour traiter les textes. Le logiciel *Oxford Concordance Package* a alors trié toutes les formes qui paraissent dans chaque roman, les a ensuite classées par ordre alphabétique et a calculé la fréquence absolue des types.

Un autre logiciel, le *Microsoft Excel*, a classé par ordre de fréquence décroissante tous les vocables contenus dans chaque roman. Il a compté la fréquence des types et les a convertis en fréquence relative par tranche de 1.000 vocables, ce qui est la norme pour de telles études (voir les appendices 10.1.1 à 10.1.9). L'étude

statistique découle ainsi d'une série de concordances faite par le logiciel *O.C.P.* qui a servi de logiciel primaire, et du logiciel *Microsoft Excel*.

10.3. Constitution du corpus

Le corpus d'Engwall contient 500.000 vocables tandis que le nombre total de vocables de chaque roman étudié varie entre 33.625 (*L'Étranger*) et 65.095 (*La Vie devant soi*). La moyenne des vocables pour ces huit romans est 48.162. L'écart-type est 12,129. La table 10.1 indique la moyenne des vocables et l'écart réduit pour les huit romans.

Table 10.1

Fréquences et Écarts réduits

Roman	Nombre de vocables	Écart réduit
<i>Clair</i>	33.755	-1,181 σ
<i>Ticket</i>	57.915	+0,799 σ
<i>Câlin</i>	53.264	+0,418 σ
<i>Vie</i>	65.095	+1,388 σ
<i>L'Étranger</i>	34.625	-1,110 σ
<i>Vipères</i>	59.426	+0,923 σ
<i>Porte</i>	39.953	-0,673 σ
<i>L'Immoraliste</i>	41.260	-0,566 σ
Total	385.293	
μ	48.162	
σ	12.192	

Les observations se situent entre -1,18 et 1,39 écarts-type de la moyenne. *L'Immoraliste*, *La Porte étroite*, *Au-delà de cette limite votre ticket n'est plus valable*, *Gros-Câlin* et *Le Noeud de Vipères* se situent entre -1 et +1 écart-type de la moyenne. *La Vie devant soi*, *Clair de femme* et *L'Étranger* se trouvent à plus d'un écart-type de la moyenne. Les romans sont raisonnablement homogènes du point de vue du nombre des vocables, justifiant ainsi l'emploi de l'analyse statistique.

10.4. Les soixante types les plus fréquents dans les huit romans et dans le corpus d'Engwall

Cette étude se base sur les soixante types les plus fréquents du corpus, puisqu'à l'exemple de Burrows, je considère que les mots les plus fréquents constituent un élément essentiel du style (*Computation* 4). Les appendices 10.1.1 à 10.1.9 indiquent (à titre de référence seulement) les soixante premiers types dans chaque texte, présentés en ordre de fréquence décroissante. La table 10.2 regroupe ces données et indique le rang de ces types. Souvent leur rang varie d'un roman à l'autre.

Table 10.2

Les Soixante Premiers Types dans chaque roman et dans le corpus d'Engwall

Rang	Clair	Tick	Câl	Vie	Étran	Vip	Imm	Porte	Eng
1	de	de	de	et	de	de	de	de	de
2	je	je	je	de	et	je	je	je	la
3	et	et	et	il	il	à	et	que	le
4	il	la	à	je	a	la	la	à	et
5	la	il	la	elle	je	que	à	la	il
6	un	à	il	que	le	le	que	le	à
7	à	un	le	pas	la	il	le	le	les
8	est	que	est	à	que	et	à	elle	je
9	le	l'	que	la	j'	l'	il	ne	l'
10	pas	le	l'	le	à	ne	me	l'	un
11	que	d'	un	est	ai	les	un	d'	que
12	vous	une	pas	a	l'	d'	ne	pas	d'
13	l'	pas	les	j'	un	un	d'	qu'	en
14	ne	les	c'	un	m'	pas	les	me	des
15	une	est	d'	qu'	qu'	ce	en	ce	pas
16	d'	ne	j'	les	était	me	j'	un	une
17	les	en	en	c'	les	qu'	qu'	j'	elle
18	c'	me	dans	était	d'	tu	plus	en	est
19	n'	des	une	qui	pas	qui	m'	m'	qu'
20	elle	j'	me	des	est	j'	pas	mon	ne
21	ce	dans	des	avait	en	en	elle	il	dans
22	a	ce	ne	l'	me	elle	ce	n'	qui
23	qui	qu'	qui	pour	avait	n'	une	est	n'
24	nous	c'	pour	ne	n'	une	est	nous	se
25	j'	vous	qu'	ai	une	des	n'	plus	ce
26	des	qui	vous	en	mais	m'	nous	les	pour
27	qu'	n'	on	madame	ne	dans	pour	tu	s'
28	on	pour	avec	ça	dit	était	était	ma	j'
29	y	a	a	n'	lui	pour	des	qui	vous
30	en	plus	n'	on	dans	est	mais	mais	du
31	dans	était	ce	Rosa	c'	mais	qui	dans	a
32	pour	tu	ai	mais	s'	plus	mon	une	c'
33	me	du	à	une	elle	au	se	pour	plus
34	mais	elle	mais	plus	ce	nous	dans	moi	sur
35	était	se	y	dans	des	mon	mais	si	me
36	du	y	du	d'	qui	ma	moi	s'	était
37	plus	m'	au	ce	pour	vous	du	du	au
38	se	au	tout	comme	du	cette	s'	sans	lui
39	tout	mon	était	me	nous	moi	Marceline	Alissa	mais
40	ai	mais	se	y	mon	avait	tout	tout	son
41	avait	avait	elle	vous	avais	du	c'	lui	nous

42	moi	avec	m'	s'	tout	s'	comme	ai	avait
43	au	on	moi	se	sur	se	vous	au	ils
44	s'	moi	mon	tout	avec	lui	au	était	tout
45	m'	ça	s'	avec	on	a	sur	se	avec
46	tu	nous	plus	lui	au	c'	on	son	on
47	ça	sur	même	quand	moi	comme	lui	c'	comme
48	si	s'	sur	m'	plus	sur	y	par	tu
49	sur	ai	comme	ils	son	sa	si	cette	sa
50	comme	tout	suis	même	peu	son	bien	a	ai
51	avec	même	bien	parce	si	si	par	te	m'
52	mon	son	avait	du	y	ai	avait	sans	si
53	encore	ma	là	monsieur	sa	ils	même	comme	y
54	suis	si	par	moi	alors	tout	son	t'	ses
55	lui	par	lui	suis	suis	avais	avais	bien	bien
56	femme	comme	fait	encore	se	avec	sans	avait	par
57	bien	bien	dit	bien	même	ces	encore	sur	moi
58	même	encore	très	son	cela	même	rien	même	mon
59	là	lui	nous	fait	fait	rien	peu	y	même
60	Lydia	Laura	ils	rien	étais	être	avec	avec	dit

Le type *de* occupe le premier rang dans sept de ces romans et dans le corpus d'Engwall; dans *La Vie devant soi* il se situe au deuxième rang. Le type *je* se trouve au deuxième rang dans six de ces romans, sauf dans *La Vie devant soi* où il se situe au quatrième rang, dans *L'Etranger* il se situe au cinquième rang. Dans le corpus d'Engwall *je* se situe au huitième rang, ce qui n'est pas étonnant étant donné que les vingt-cinq romans qui constituent ce corpus n'utilisent pas tous un narrateur homodiégétique. Les types les plus fréquents dans ces textes sont presque tous des mots-outils. La table 10.3 indique la proportion des soixante types les plus fréquents dans les neuf textes.

Cette étude se base sur les soixante types les plus fréquents dans les neuf textes ce qui représente le double du nombre de types que Burrows étudie (Voir chapitre 8).

Ces soixante types constituent entre 48,39 % et 56,58 % du corpus. Étant donné que l'étude statistique envisagée se rapporte à une si grande partie du texte, les résultats statistiques obtenus sont dignes de foi.

Table 10.3

Répartition des 60 premiers types dans chaque roman et dans le corpus d'Engwall

Titre	Nombre de vocables	Types 1-60	%	Liste combinée	%
<i>Clair</i>	33.755	17.164	50,84	17.079	50,59
<i>Ticket</i>	57.915	29.037	50,14	28.812	49,75
<i>Câlin</i>	53.264	28.068	52,69	27.894	52,37
<i>Vie</i>	65.095	36.833	56,58	35.627	54,73
<i>Étranger</i>	34.625	19.105	55,18	18.799	54,30
<i>Vipères</i>	59.426	29.501	49,64	28.674	48,25
<i>Immoraliste</i>	41.260	20.226	49,02	19.585	47,47
<i>Porte</i>	39.953	19.911	49,83	19.108	47,83
Engwall	500.000	241.499	48,39	239.487	47,90

10.5. Constitution d'une liste unique des soixante types les plus fréquents dans les neuf textes

Les appendices 10.1.1 à 10.1.9 démontrent que le rang des soixante types les plus fréquents varie d'un texte à l'autre; quelques types apparaissent dans un seul roman et non pas dans les autres, surtout en ce qui concerne les vocables qui dépendent du contexte tels que *madame* et *Rosa* (rangs 27 et 31) de *La Vie devant soi*. Malgré cela, il est possible d'obtenir une liste des 60 types les plus fréquents,

communs à tous ces neuf textes. Il suffit d'obtenir la moyenne de chaque type dans les neuf textes et de classer les moyennes par fréquence décroissante.

L'appendice 10.2 présente cette liste combinée commune à tous les neuf textes. La colonne indiquée sous le signe μ donne la moyenne des types classés en ordre décroissant. Les types tels que *Lydia (Clair)*, *Laura (Ticket)*, *Rosa (La Vie)*, *Marceline (Immoraliste)* et *Alissa (Porte)* ne figurent plus dans cette liste. Encore une fois, c'est le type *de* qui est souvent le plus fréquent, suivi de *je*, *et*, *la* et ainsi de suite.

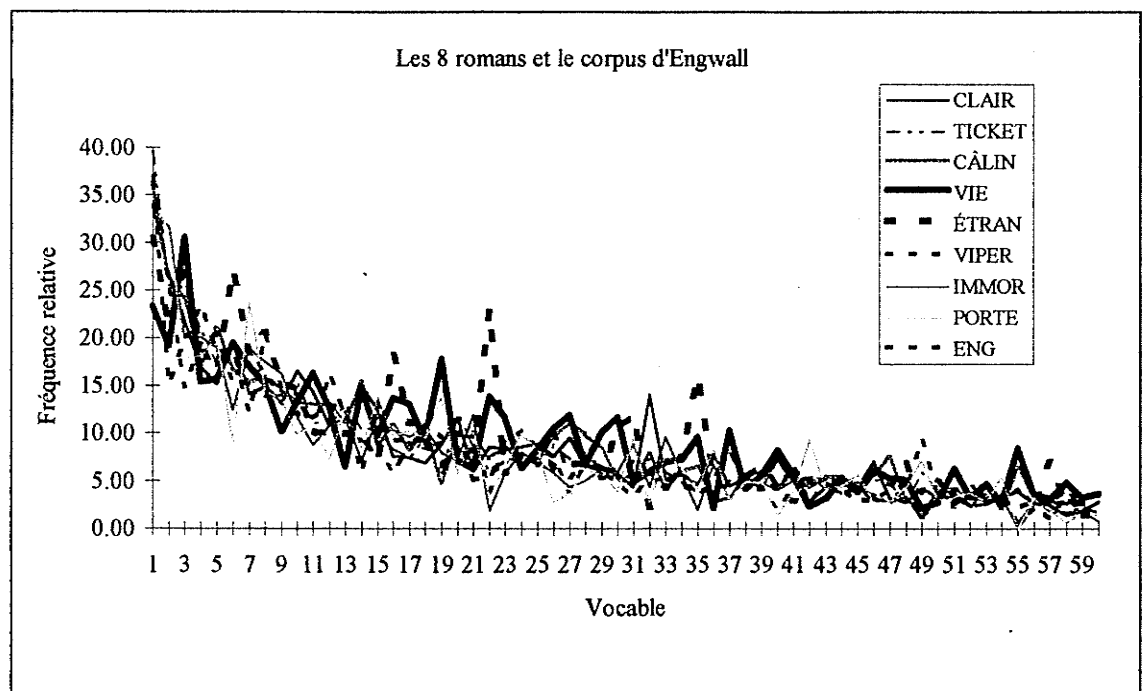
Lorsque les soixante premiers types sont regroupés, ils constituent entre 47,47% et 54,73% des textes pris individuellement. Étant donné que ces valeurs varient à peine de celles produites par les soixante premiers types de chaque texte (voir la table 10.3), il est légitime de travailler sur ces nouvelles valeurs. Les soixante types les plus fréquents dans ces textes (voir l'appendice 10.2) produisent les graphiques 10.1 à 10.4. Ces graphiques représentent visuellement les similarités et les différences entre les textes.

Le premier graphique (10.1) regroupe les soixante types les plus communs dans les neuf oeuvres (appendice 10.2) et montre les similarités et les différences entre elles. L'axe vertical représente la fréquence relative des types par tranches de 1.000 vocables et l'axe horizontal indique le rang des types: il commence par le type le plus fréquent à gauche (type 1) et se termine par le type le moins fréquent à droite (type 60). Il s'agit des mêmes types pour tous les textes.

Étant donné que le graphique 10.1 renferme un si grand nombre de courbes, il est difficile visuellement de faire une distinction entre elles. Malgré cela, des caractéristiques se dégagent; notamment que *La Vie devant soi* et *L'Étranger* se détachent le plus des autres textes. En se basant toujours sur les types les plus fréquents dans les neuf textes, il est légitime de regrouper ces romans afin d'obtenir une meilleure approximation de leur similarité et leur différence.

Graphique 10.1

Les Huit Romans et le corpus d'Engwall

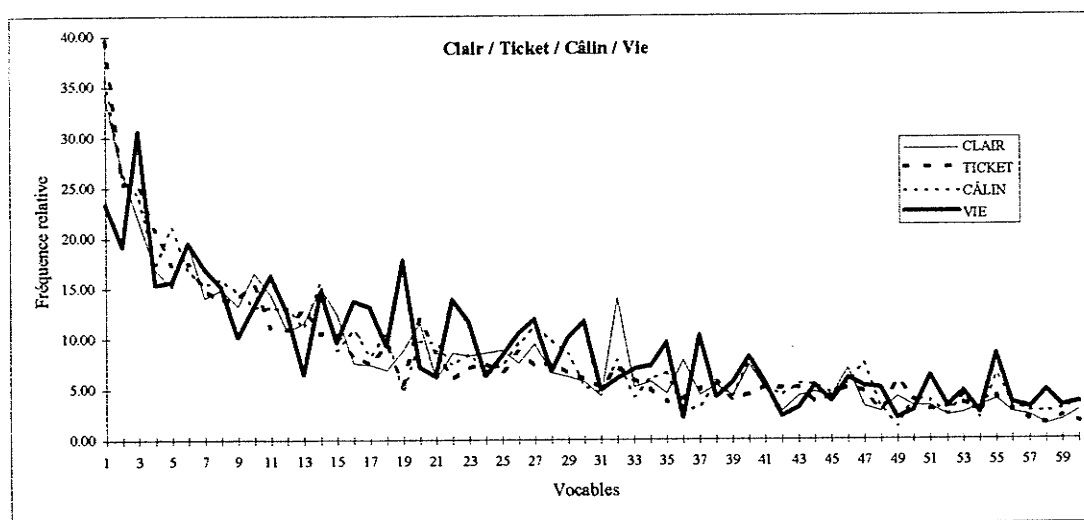


Le graphique 10.2 est obtenu lorsque les quatre romans de Gary/Ajar sont comparés entre eux. Ce graphique démontre que dans l'ensemble, les courbes pour *Au-delà de cette limite votre ticket n'est plus valable*, *Clair de femme* et *Gros-Câlin* se

correspondent entre eux. Par contre quelques types dans *La Vie devant soi* tels que *de* (rang 1), *et* (rang 3), *d'* (rang 13), *elle* (rang 19) *a* (rang 22) se détachent des autres courbes. Ce graphique semble démontrer qu'*Au-delà de cette limite votre ticket n'est plus valable*, *Clair de femme* (Gary) et *Gros-Câlin* (Ajar) se ressemblent tandis que *La Vie devant soi* s'écarte de ces trois romans.

Graphique 10.2

Clair, Ticket, Câlin et La Vie

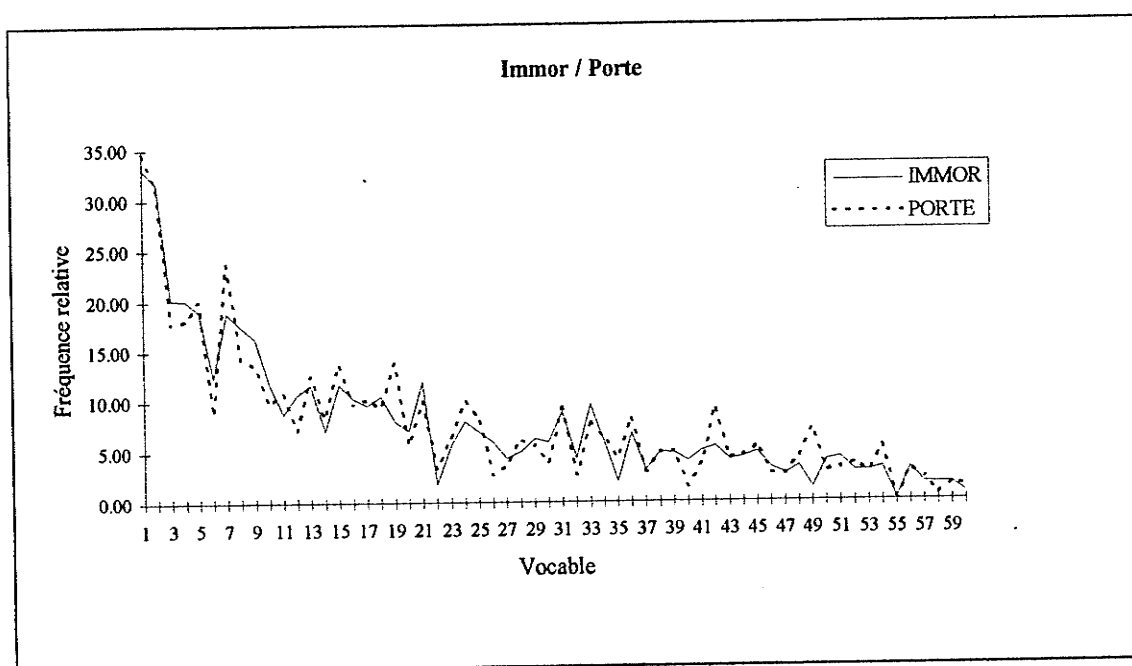


Pour obtenir une meilleure approximation de cet écart, il est nécessaire de comparer les romans rédigés par les mêmes auteurs entre eux. Le graphique 10.3 montre les similarités et les différences entre *L'Immoraliste* et *La Porte étroite* de Gide.

Ce graphique indique que quelques différences existent entre ces deux romans, surtout en ce qui concerne les types *il* (rang 6), *que* (rang 7), *les* (rang 12), *elle* (rang 19). Malgré cela, il existe plus de ressemblance que de différence entre ces romans. On pourrait ainsi conclure que les courbes se rapportant à des romans rédigés par le même auteur et publiés à quelques années d'intervalle ont tendance à se grouper. Pour tester cette hypothèse, il est nécessaire de comparer les deux romans signés Gary entre eux.

Graphique 10.3

L'Immoraliste et La Porte étroite

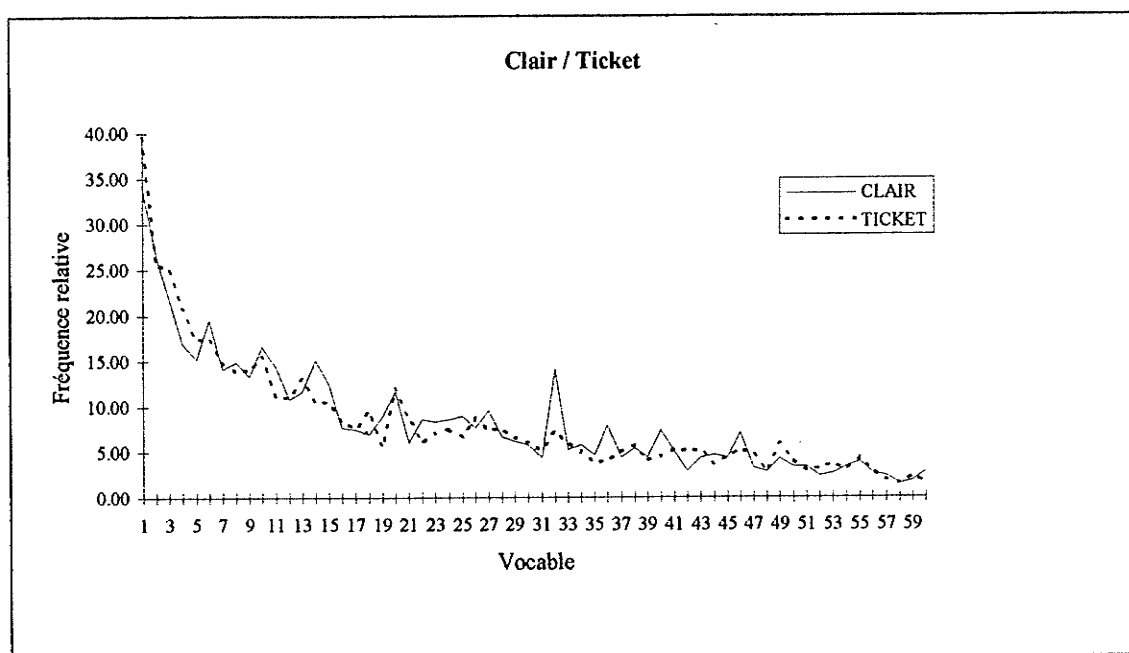


Le graphique 10.4 présente les similarités et les différences entre ces deux romans. Exception faite de quelques types tels que *de* (rang 1), *la* (4e rang), *est* (14e

rang) et *vous* (32e rang), il existe une forte ressemblance entre *Clair de femme* et *Au-delà de cette limite votre ticket n'est plus valable*. Ceci n'est pas surprenant, étant donné qu'ils ont été tous deux publiés sous le nom de Gary, et cela à deux années d'intervalle.

Graphique 10.4.

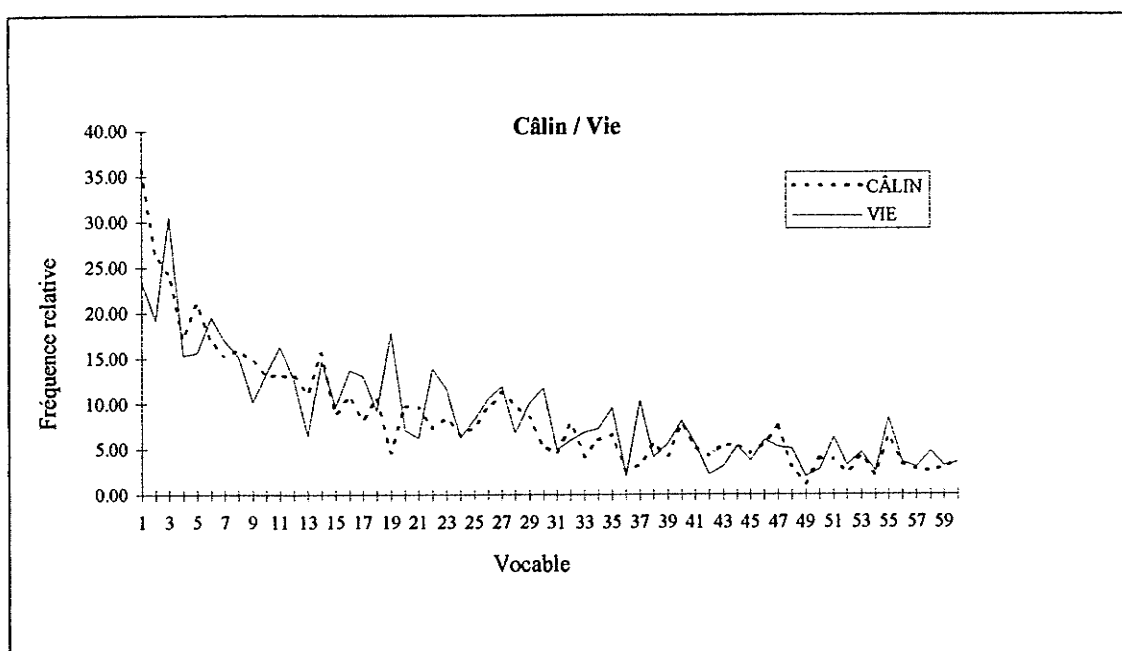
Clair et Ticket



Le graphique 10.5 présente les similarités et les différences entre les deux romans publiés sous le nom d'Ajar. Cette fois-ci, les similarités entre les courbes sont moins nombreuses, c'est-à-dire qu'il y a plus d'endroits où elles divergent d'une manière remarquable, surtout en ce qui concerne les types *de* (rang 1), *je* (rang 2) et (rang 3), *à* (rang 5), *l'* (rang 9), *d'* (rang 13), *elle* (rang 19), *a* (rang 22), *était* (rang 30).

Graphique 10.5

Gros-Câlin et La Vie devant soi



Les graphiques suggèrent que *Clair de femme*, *Au-delà de cette limite votre ticket n'est plus valable* et *Gros-Câlin* se ressemblent, tout comme *La Porte étroite* et *L'Immoraliste*. *La Vie devant soi* se détache non seulement des autres romans de Gary/Ajar, mais encore de *Gros-Câlin* également signé Ajar (voir graphique 10.2). Une investigation plus approfondie de ce phénomène semble ainsi nécessaire. Le test-*t* de Student -Fisher s'y avère alors très importante.

10.6. Le test *t* de Student-Fisher

Tel qu'indiqué au chapitre 9, le test-*t* de Student-Fisher (qui est une correction du test de Z, mais qui est utilisé lorsque le nombre d'observations est petit) permet

d'estimer la probabilité de rencontrer une observation à une certaine valeur. L'une des exigences du test est qu'il ne peut être effectué que sur des populations homogènes. Ceci a été facilité par le choix des romans eux-mêmes.

L'appendice 10.2 présente la liste des 60 types les plus fréquents dans les neuf textes par ordre décroissant de leur fréquence moyenne aussi bien que les écarts-type de chaque côté de la moyenne. Pour une distribution gaussienne toutes les observations qui se situent au delà de deux écarts-types de chaque côté de la moyenne n'ont qu'environ deux chances sur cent de s'y trouver si le hasard seul entre en jeu. Toute observation qui se situe au-delà de deux écarts-type est donc un événement rare. Ces observations sont indiquées en chiffres gras dans l'appendice 10.2. La table 10.4 regroupe ces données et indique le nombre d'observations se situant au delà de deux écarts-type de la moyenne dans chaque texte.

Table 10.4

Le Nombre d'observations se situant au delà de deux écarts-type de la moyenne

<i>Clair</i>	2	<i>Ticket</i>	0	<i>Câlin</i>	2	<i>Vie</i>	8
<i>Étranger</i>	4	<i>Vipères</i>	0	<i>Immoraliste</i>	1	<i>Porte</i>	2
Engwall	0						

Au-delà de cette limite votre ticket n'est plus valable, Le Noeud de Vipères, et le corpus d'Engwall ne présentent aucune observation se situant au delà de 2 écarts-

type de la moyenne. *L'Immoraliste* en a une; *Clair de femme*, *Gros-Câlin* et *La Porte étroite* en présentent deux chacun, *L'Étranger* en présente quatre et *La Vie devant soi* huit. Le test-*t* de Student-Fisher suggère que *La Vie devant soi* est le roman le plus extraordinaire du groupe.

10.7. Les Limites de confiance

Compte tenu des résultats obtenus, il est légitime de se demander si une telle situation aurait pu se produire si le hasard seul entre en jeu. C'est le calcul des limites de confiance qui permet de répondre à une telle question. Ces limites constituent une fourche autour d'une valeur obtenue et permettent d'évaluer la probabilité d'obtenir les mêmes résultats avec d'autres romans tels que *Les Cerfs-Volants* de Gary, *L'Angoisse du Roi Salomon* d'Ajar ou encore *La Nausée* de Sartre.

Un risque d'erreur de 5% a d'abord été choisi ce qui veut dire que toutes les valeurs obtenues n'ont que cinq chances sur cent d'être le résultat du hasard. (L'appendice 10.3 contient ces limites de confiance à 95%). Toutes les valeurs significatives sont indiquées en chiffres gras. La moyenne est encore une fois désignée par le signe μ et l'écart-type par le signe σ .

Pour choisir le parti pris du conservatisme, et pour que le risque d'erreur soit réduit considérablement, les limites de confiance à 99% ont aussi été calculées (Voir

l'appendice 10.4), ce qui veut dire qu'il n'existe qu'une chance sur cent de se tromper.

Cela permet alors d'obtenir des résultats plus solides quoique plus rares.

Étant donné que pour chaque oeuvre, cette liste est constituée de soixante types, aux limites de confiance de 95%, on accepte la probabilité que les résultats soient fautifs pour trois de ces soixante observations, sans qu'on sache lesquels. Aux limites de confiances de 99%, on accepte la probabilité qu'un résultat (0,6) soit fautif sur les soixante observations sans qu'on sache lequel. Néanmoins un grand nombre d'observations dépasse les limites établies (voir les appendices 10.3 et 10.4). Elles sont regroupées dans la table 10.5.

Table 10.5

Limites de confiance à 95% et à 99%

<i>Limites de confiance à 95%</i>			<i>Limites de confiance à 99%</i>		
Titre	Obs/60	%	Titre	Obs/60	%
<i>Clair</i>	22	36	<i>Clair</i>	11	18
<i>Ticket</i>	14	23	<i>Ticket</i>	04	07
<i>Câlin</i>	23	38	<i>Câlin</i>	15	25
<i>La Vie</i>	39	65	<i>La Vie</i>	31	52
<i>Étranger</i>	27	45	<i>Étranger</i>	16	27
<i>Vipères</i>	19	31	<i>Vipères</i>	10	17
<i>Porte</i>	32	53	<i>Porte</i>	16	27
<i>Immoraliste</i>	25	41	<i>Immoraliste</i>	09	15
<i>Engwall</i>	22	36	<i>Engwall</i>	14	23

La table 10.5 démontre que tous les romans accusent un grand nombre d'observations se situant hors des limites de confiance de 95%. Elles varient entre

23% des observations pour *Au-delà de cette limite votre ticket n'est plus valable* et 65% pour *La Vie devant soi*. Aux limites de confiance de 99%, encore une fois presque tous les romans contiennent des observations se situant hors de ces limites: elles varient entre 7% pour *Au-delà de cette limite votre ticket n'est plus valable* et 52% pour *La Vie devant soi*. Aux limites de confiance 95% et de 99% presque la moitié des types les plus fréquents dans *La Vie devant soi* se situent hors des bornes établies, tandis que cette valeur n'est pas si élevée pour les autres textes.

Les résultats obtenus dans ce chapitre sont prometteurs. Les graphiques, le test-t de Student-Fisher aussi bien que les limites de confiance démontrent que *La Vie devant soi* est considérablement différent des autres romans, y compris les autres romans de Gary et même de *Gros-Câlin* également signé Ajar. Ils nous convainquent qu'une étude plus approfondie et plus détaillée de ce phénomène s'impose. Les deux prochains chapitres sont consacrés à la corrélation de Pearson et au test de khi-deux appliqués aux données élaborées dans ce chapitre.

Chapitre 11

Le Test du χ^2

Si l'être constitue un tout organisé, on doit retrouver le symbole du moi dans chacune de ses manifestations [...] la même cause produit toujours les mêmes effets. (Morier 7)

Le principe du test du χ^2 est d'apprécier l'écart entre une valeur observée et une valeur théorique. Il existe plus d'une manière de calculer les valeurs théoriques; elles peuvent être obtenues par l'intermédiaire d'une table de contingence ou encore par le test d'adéquation.

Pour le premier test effectué dans ce chapitre (le test d'adéquation) le corpus d'Engwall fournit les valeurs théoriques, étant donné qu'elle a établi une liste de fréquences des types les plus communs dans le roman français des années soixante. Le but d'Engwall était d'orienter la recherche

vers un vocabulaire reflétant la France de nos jours et employé par les jeunes auteurs français [...] et à procurer un vocabulaire moderne traduisant les activités de la France contemporaine. (11)

D'abord, la fréquence des types les plus fréquents de chaque roman sera comparée à la fréquence de ces mêmes types dans le corpus d'Engwall et ces écarts seront mesurés. Ceci permettra d'établir les ressemblances et les différences entre le corpus d'Engwall et les huit romans. Ensuite, ce sont les deux romans signés Gary, *Clair de femme* et *Au-delà de cette limite votre ticket n'est plus valable*, qui fourniront les valeurs théoriques.

11.1. La Valeur théorique représentée par le corpus d'Engwall

L'appendice 11.1 présente la liste combinée des types dans les neuf textes et indique leur fréquence. Elle est composée de dix colonnes et de plusieurs rangées. La première rangée indique le nombre total des types pour chaque oeuvre (33.755 pour *Clair de femme* et ainsi de suite). Toutes les colonnes (sauf la première) donnent la fréquence de chaque type (1.143 pour le type *de* dans *Clair de femme* par exemple). La dernière rangée contient le total lorsque les soixante types dans chaque texte sont additionnés.

Afin de respecter l'une des exigences du test d'adéquation (que les unités de comparaison soient uniformes) les fréquences du corpus d'Engwall ont été converties

au nombre d'occurrences de chaque roman. Ces nouvelles valeurs sont indiquées dans l'appendice 11.2. Le test du χ^2 compare la fréquence des types dans chaque texte à ces nouvelles valeurs.

Les valeurs du χ^2 sont présentées dans l'appendice 11.3. (Dorénavant toute valeur significative à un degré de liberté sera indiquée en caractères gras dans les appendices). Plus de 50% des valeurs dans l'appendice 11.3 sont significatives. La valeur globale du χ^2 pour chaque oeuvre est indiquée dans la dernière rangée de l'appendice; ces valeurs sont reproduites dans la table 11.1.

Table 11.1

Valeurs du χ^2 pour les huit romans

<i>Clair</i>	2.162,16	<i>Tick</i>	2.107,69	<i>Cálin</i>	3.796,45	<i>Vie</i>	9.410,77
<i>Étran</i>	7.135,68	<i>Vip</i>	3.199,77	<i>Immor</i>	3.196,77	<i>Porte</i>	4.744,52
ddl pour chaque oeuvre: 59							

Cette fois-ci, le nombre de degrés de liberté pour chaque oeuvre est de 59, mais étant donné que 59 ddl ne figurent pas dans la liste des seuils de signification, 60 ddl ont été choisis pour lesquels les valeurs à dépasser sont 83,30 (au seuil de 0,025), 88,38 (au seuil de 0,01) et 91,95 (au seuil de 0,005). Dorénavant toutes les valeurs présentées à 59 ddl seront indiquées à 60 ddl.

Comme les valeurs du χ^2 dépassent ces seuils, la répartition des types les plus fréquents dans ces romans diffère de celle des types dans le corpus d'Engwall. Ceci résulte en partie du fait que le corpus d'Engwall ne se limite pas aux romans homodiégétiques. Quoique ces résultats soient tous significatifs, il existe de grandes différences entre les valeurs du χ^2 . Elles varient entre 2.107,69 et 9.410,77 et c'est *La Vie devant soi* qui présente la valeur du χ^2 la plus élevée. Ce premier test démontre que ce roman se distingue le plus du corpus d'Engwall.

Étant donné que ces valeurs sont toutes très fortes et que le corpus d'Engwall ne se limite pas aux narrateurs homodiégétiques, tout vocable susceptible d'indiquer la perspective de la narration tels que les marqueurs de personne grammaticale (verbes conjugués, pronoms personnels etc., dotés d'un astérisque dans l'appendice 11.3) ont été ôtés de la liste. Ceci permet de comparer la fréquence des types qui sont indépendants du contexte à la fréquence de ces mêmes types dans le corpus d'Engwall. Les valeurs du χ^2 indiquées dans l'appendice 11.4 sont alors obtenues et la majorité de ces valeurs dépassent 3,84. Les valeurs globales du χ^2 pour chaque roman sont regroupées dans la table 11.2.

Cette fois-ci le nombre de degrés de liberté pour chaque oeuvre est de 30. Les seuils de signification statistique sont 46,98 (au seuil de 0,025), 50,89 (au seuil de 0,01) et 53,67 (au seuil de 0,005). La table 11.2 démontre que les valeurs du χ^2

dépassent ces seuils pour tous les romans. Encore une fois c'est *La Vie devant soi* qui présente la valeur la plus élevée du χ^2 .

Table 11.2

Valeurs du χ^2 (sans les marqueurs de personne grammaticale)

<i>Clair</i>	685,58	<i>Tick</i>	567,74	<i>Câlin</i>	1.414,27	<i>Vie</i>	4.156,33
<i>Étran</i>	548,12	<i>Vip</i>	925,24	<i>Immor</i>	717,10	<i>Porte</i>	1.350,51

ddl pour chaque oeuvre: 30

Les tests du χ^2 démontrent sans aucune équivoque que lorsque les huit romans sont comparés au corpus d'Engwall, chacun s'est situé à l'écart de ce corpus. On observe en même temps que parmi tous ces romans (y compris les quatre romans de Gary), *La Vie devant soi* s'écarte le plus du corpus d'Engwall, même lorsque les marqueurs de personne grammaticale ne figurent plus dans ces tests.

11.2. La Valeur théorique représentée par *Clair* et *Ticket*

On pourrait postuler que les fortes valeurs de la statistique du χ^2 obtenues jusqu'ici résultent en partie du fait que le corpus d'Engwall (le groupe témoin) ne se limite pas au style d'un seul auteur, mais qu'il représente une sorte de dénominateur commun du style de vingt-cinq auteurs. Pour écarter l'éventualité que ceci aurait pu influencer les valeurs du χ^2 , les romans seront comparés à un autre groupe témoin,

Clair de femme et *Au-delà de cette limite votre ticket n'est plus valable*, étant donné qu'ils ont tous deux été écrits par Gary et publiés sous son propre nom.

Pour cette analyse, on calcule la moyenne de la fréquence des types dans *Clair de femme* et *Au-delà de cette limite votre ticket n'est plus valable* (voir la deuxième colonne de l'appendice 11.5). Ensuite, on convertit cette moyenne au nombre de types dans les autres romans (voir les six dernières colonnes de l'appendice 11.5). En comparant ces nouvelles valeurs calculées aux valeurs observées dans chaque roman (appendice 11.1) les χ^2 indiqués dans l'appendice 11.6 sont obtenues. La table 11.3. contient les valeurs globales du χ^2 indiquées au bas des colonnes.

Table 11.3

Valeurs du χ^2 (*Clair* et *Ticket* représentent la valeur théorique)

<i>Câlin</i>	1.436,93	<i>Vie</i>	7.215,57			
<i>Étran</i>	5.459,71	<i>Vip</i>	2.301,38	<i>Immor</i>	1.946,12	<i>Porte</i> 3.008,60
ddl 59						

Lorsque les deux romans de Gary fournissent les valeurs théoriques, on pourrait s'attendre à ce que *Gros-Câlin* et *La Vie devant soi* s'accordent le plus avec ces valeurs. En effet, *Gros-Câlin* s'en rapproche le plus (quoique la statistique du χ^2 soit bien au-delà du seuil de signification) mais *La Vie devant soi* est le plus éloigné. Effectivement, il y a plus de ressemblance entre *L'Étranger*, *L'Immoraliste*, *La Porte*

étroite, *Le Noeud de Vipères*, et les deux romans de Gary qu'entre *La Vie devant soi* et ces deux romans.

Encore une fois, afin d'éviter l'effet causée par la voix narrative, tout vocable susceptible d'indiquer la perspective de la narration est ôté de cette liste avant de faire le calcul du χ^2 . Les résultats obtenus sont indiqués au bas de l'appendice 11.6 dotés de deux astérisques. La table 11.4 présente la valeur du χ^2 pour chaque roman.

Table 11.4

Valeurs du χ^2 (sans les marqueurs de personne): *Clair* et *Ticket* représentent la valeur théorique

<i>Câlin</i>	662,22	<i>Vie</i>	2.289,48	<i>Étran</i>	546,47
<i>Immor</i>	1.018,10	<i>Vip</i>	821,35	<i>Porte</i>	1.258,67

ddl: 30

Le nombre de degrés de liberté pour chaque texte est de 30 et les seuils de signification à 30 ddl sont 46,98 (au seuil de 0,025), 50,89 (au seuil de 0,01) et 53,67 (au seuil de 0,005). Toutes les valeurs du χ^2 dépassent ces seuils. Cette fois-ci, c'est *L'Étranger* qui se rapproche le plus des deux romans de Gary, suivi de près par *Gros-Câlin*. Par contre, *La Vie devant soi* s'accorde le moins avec les deux romans de Gary.

Les tests d'adéquation du χ^2 effectués jusqu'ici démontrent que la répartition des vocables les plus fréquents dans les neuf textes est à peine semblable d'un texte à

l'autre, ce qui n'est pas surprenant. Par contre, certaines autres caractéristiques se dégagent. Malgré le fait qu'il existe une différence significative dans la fréquence des vocables dans ces neuf textes, des quatre romans de Gary/Ajar, *La Vie devant soi* se détache le plus de la valeur hypothétique, qu'elle soit représentée par le corpus d'Engwall ou par les deux romans de Gary. Il n'est nullement surprenant que ce roman ne ressemble pas au corpus d'Engwall, mais ce qui étonne, c'est qu'il soit si différent des deux autres romans signés Gary, surtout lorsque *Gros-Câlin* s'éloigne le moins de ces deux romans.

11.3. La table de contingence

La manière traditionnelle d'effectuer le test du χ^2 est par le biais de la méthode qui s'intitule "table de contingence" dont une description figure à la section 9.11 du chapitre 9. Encore une fois, pour ce test, une valeur observée est comparée à une valeur hypothétique, valeur qui est calculée en fonction des relations proportionnelles entre les valeurs observées elles-mêmes. Pour le premier test envisagé, les huit romans et le corpus d'Engwall sont comparés entre eux.

Les fréquences observées sont indiquées dans l'appendice 11.1 et les valeurs calculées se trouvent dans l'appendice 11.7. En appliquant la correction de Bonferroni, on obtient le seuil de rejet de 0,009 (à savoir 5/540) et un ddl: 472, à savoir (60-1)(9-1). Étant donné que 472 ddl ne figurent pas dans la liste des seuils de

signification 100 ddl ont été choisis pour lesquels les valeurs à dépasser à 0,001 est de 149,4. L'appendice 11.8 contient les valeurs du χ^2 . Toute valeur significative est indiquée en chiffre gras. Plusieurs types produisent un χ^2 significatif, le nombre de valeurs significatives varie entre 1 pour *Clair de femme* et 11 pour *La Vie devant soi*. C'est *La Vie devant soi* qui produit le plus grand nombre de valeurs significatives. Les valeurs globales du χ^2 pour chaque roman se trouvent dans la table 11.5.

Table 11.5

Valeurs du χ^2 pour la table de contingence.

<i>Clair</i>	1.258,67	<i>Ticket</i>	921,64	<i>Câlin</i>	1.817,71	<i>Vie</i>	5.157,73
<i>Étran</i>	3.980,68	<i>Vip</i>	1.941,05	<i>Immor</i>	1.891,56	<i>Porte</i>	3.089,81
<i>Eng</i>	3.230,05						

ddl pour chaque oeuvre 59

Les seuils de signification à 60 ddl sont 83,30 (au seuil de 0,025), 88,38 (au seuil de 0,01) et 91,95 (au seuil de 0,005). Toutes les valeurs du χ^2 dépassent ces seuils démontrant, encore une fois, qu'il existe une différence fondamentale entre les textes. De nouveau, la plus haute valeur du χ^2 est celle obtenue pour *La Vie devant soi*.

Afin d'éviter l'influence de la voix narrative, les verbes conjugués, les pronoms personnels et ainsi de suite sont exclus de la liste des types les plus fréquents

(Voir l'appendice 11.9). L'appendice 11.10 contient les résultats du test du χ^2 . Toute valeur individuelle significative est indiquée en chiffre gras. *Gros-Câlin* et *Le Noeud de Vipères* contiennent une valeur significative du χ^2 chacun, tandis que *La Porte étroite* en contient deux et *La Vie devant soi* cinq. Les valeurs du χ^2 pour chaque texte paraissent dans la table 11.6.

Table 11.6

Valeurs du χ^2 abstraction faite des vocables indiquant la perspective narrative

<i>Clair</i>	245,54	<i>Ticket</i>	355,62	<i>Câlin</i>	666,91	<i>Vie</i>	2.359,85
<i>Étran</i>	388,64	<i>Vip</i>	684,37	<i>Immor</i>	525,10	<i>Porte</i>	1.025,90
<i>Eng</i>	798,22						

ddl: 30 (pour chaque oeuvre)

Le nombre de degrés de liberté pour chaque oeuvre est de 30. Les seuils à dépasser sont les suivantes: 46,98 (au seuil de 0,025), 50,89 (au seuil de 0,01), 53,67 (au seuil de 0,005). Les valeurs du χ^2 pour chaque texte varient entre 245,54 (*Clair de femme*) et 2.359,85 (*La Vie devant soi*) et dépassent ces seuils. Encore une fois c'est *La Vie devant soi* qui s'éloigne le plus de la valeur théorique

Ces résultats indiquent, une fois de plus, que *La Vie devant soi* est le roman le plus remarquable des neuf textes étudiés.

11.4. Comparaison entre les huit romans (abstraction faite du corpus d'Engwall)

Étant donné que le corpus d'Engwall renferme 500.000 vocables et que les autres romans sont constitués de 35.000 à 65.000 vocables chacun, le corpus d'Engwall sera exclu des prochains tests statistiques, pour comparer les huit romans entre eux. L'appendice 11.11 indique les valeurs théoriques pour ces huit romans.

Le test du χ^2 produit les résultats qui figurent dans l'appendice 11.12. Les valeurs du χ^2 pour chaque roman sont regroupées dans la table 11.7. À 60 ddl les seuils à dépasser sont 83,30 (au seuil de 0,025), 88,38 (au seuil de 0,01), 91,95 (au seuil de 0,005). Les valeurs du χ^2 pour chaque texte varient entre 782,46 (*Au-delà de cette limite votre ticket n'est plus valable*) et 3.768,16 (*La Vie devant soi*), ce dernier texte ayant produit la plus haute valeur du χ^2 .

Table 11.7

Valeurs du χ^2 pour les huit romans

<i>Clair</i>	980,75	<i>Ticket</i>	782,46	<i>Câlin</i>	1.280,99	<i>Vie</i>	3.768,16
<i>Étran</i>	2.832,30	<i>Vip</i>	1.749,01	<i>Immor</i>	1.359,07	<i>Porte</i>	2.160,33

ddl pour chaque oeuvre: 59

Encore une fois, afin d'éviter l'effet de la voix narrative, les marqueurs de personne grammaticale sont ôtés de la liste. Les nouvelles valeurs calculées figurent

dans l'appendice 11.13 et les valeurs du χ^2 sont indiquées dans l'appendice 11.14. Les valeurs du χ^2 pour chaque roman se trouvent dans la table 11.8.

Les seuils à 30 ddl sont les suivants: 46,98 (au seuil de 0,025), 50,89 (au seuil de 0,01), 53,67 (au seuil de 0,005). Quoique toutes les valeurs du χ^2 dépassent ces seuils, la valeur la plus élevée est celle produite par *La Vie devant soi*.

Table 11.8

Valeurs du χ^2 pour les huit romans (sans les marqueurs de personne)

<i>Clair</i>	267,78	<i>Tick</i>	327,40	<i>Câlin</i>	528,48	<i>Vie</i>	1.675,63
<i>Étran</i>	327,19	<i>Vip</i>	693,85	<i>Immor</i>	459,99	<i>Porte</i>	809,99

ddl pour chaque oeuvre: 30

Tous les tests du χ^2 effectués jusqu'ici démontrent que le deuxième roman signé Ajar, *La Vie devant soi*, s'écarte le plus des autres textes, quelle que soit la source de la valeur hypothétique (le corpus d'Engwall, les deux romans signés Gary) et quelle que soit la méthode employée.

11.5. Les Quatre Romans de Gary/Ajar

Lorsqu'on compare les quatre romans de Gary/Ajar entre eux, encore une fois le test du χ^2 fournit les renseignements désirés. À partir des valeurs observées (appendice 11.1) et des valeurs calculées indiquées dans l'appendice 11.15, les valeurs

du χ^2 contenues dans l'appendice 11.16 sont obtenues. Quelques-unes de ces valeurs sont significatives: elles varient entre une valeur pour *Gros-Calin* et deux valeurs pour les trois autres romans. La table 11.9 regroupe les valeurs globales du χ^2 pour les quatre romans de Gary.

Les seuils à dépasser à 60 ddl sont 83,30 (au seuil de 0,025), 88,38 (au seuil de 0,01), 91,95 (au seuil de 0,005). Les χ^2 pour chaque roman dépassent ces seuils. *La Vie devant soi* présente la valeur la plus élevée du χ^2 , c'est-à-dire qu'il s'éloigne le plus de ces autres romans. De fait, la valeur du χ^2 obtenue pour *La Vie devant soi* est plus de deux fois plus élevée que pour les autres romans de Gary.

Table 11.9

Valeurs du χ^2 pour les quatre romans de Gary/Ajar

<i>Clair</i>	751,66	<i>Ticket</i>	1.111,37	<i>Câlin</i>	682,73	<i>Vie</i>	2.252,47
--------------	--------	---------------	----------	--------------	--------	------------	----------

ddl pour chaque oeuvre: 59

La prochaine étape de cette analyse consiste à supprimer les marqueurs de personne des fréquences analysées (afin d'écarter l'influence de la voix narrative). Les valeurs calculées pour ce nouveau test sont indiquées dans l'appendice 11.17. Les

valeurs du χ^2 indiquées dans l'appendice 11.18 sont obtenues. Les valeurs pour chaque roman sont indiquées au bas des colonnes et paraissent dans la table 11.10.

Les seuils de signification pour chaque roman à 30 degrés de liberté sont 46,98 (au seuil de 0,025), 50,89 (au seuil de 0,01) et 53,67 (au seuil de 0,005). Toutes les valeurs du χ^2 dépassent ces seuils, mais *La Vie devant soi* présente la valeur du χ^2 la plus élevée indiquant qu'elle se détache le plus des trois autres romans de Gary/Ajar.

Table 11.10

Valeurs du χ^2 pour les quatre romans de Gary/Ajar (abstraction faite des marqueurs de personne)

<i>Clair</i>	253,26	<i>Ticket</i>	419,29	<i>Câlin</i>	194,36	<i>Vie</i>	821,48
--------------	--------	---------------	--------	--------------	--------	------------	--------

ddl: 30

11.6. Conclusions

Un grand nombre de tests ont été effectués dans ce chapitre. Que les marqueurs de personnes soient inclus ou non dans ces tests, qu'ils soient effectués par l'intermédiaire du test d'adéquation ou par une table de contingence, ils révèlent tous certaines caractéristiques indiscutables. Le deuxième roman signé Ajar, *La Vie devant soi*, est toujours le plus éloigné des autres oeuvres étudiées, qu'il soit comparé au corpus d'Engwall, aux autres romans ou encore aux deux romans de Gary. Ces tests

confirment les résultats déjà obtenus et indiquent que, de tous les textes étudiés, *La Vie devant soi* est le plus remarquable.

Chapitre 12

Le Coefficient de corrélation de Pearson

Le style ne mérite ce nom que lorsqu'il porte la marque d'une personnalité (Morier 58).

Ce chapitre examine les similarités entre les romans étudiés par le biais de la corrélation de Pearson. Comme il n'est nullement mon intention de vérifier l'influence d'une variable sur une autre (les types les plus fréquents dans *Le Noeud de vipères* sur les types les plus fréquents dans *La Porte étroite*, par exemple), j'utiliserai les coefficients de corrélation de Pearson à des fins comparatives et descriptives. Aussi, l'exigence notée dans la section 13 du chapitre 9, que les données soient

normalement distribuées (comme dans une distribution gaussienne), n'est-elle pas importante ici.

12.1. Corrélation entre les oeuvres

Dans un premier temps, l'étude se base sur les soixante types les plus fréquents du corpus (voir appendice 10.2). Ensuite, elle se concentre sur une liste abrégée des trente et un types dotés d'un astérisque dans le même appendice afin d'écarter la perspective narrative. Ceci est dû au fait que le corpus d'Engwall ne recense pas seulement les textes à narrateurs homodiégétiques. La table 12.1 présente la corrélation entre les oeuvres lorsque le nombre de types est de soixante et la table 12.2 présente la corrélation entre ces mêmes oeuvres lorsque le nombre de types est de trente et un.

Dans ces deux tables, la première colonne à gauche, présente les neuf textes. Les autres colonnes indiquent le coefficient de corrélation lorsque chacun de ces textes est comparé aux huit autres. Les corrélations varient entre 0,9628 et 0,6810 (lorsque le nombre de types est de soixante) et entre 0,9817 et 0,7309 (lorsque le nombre de types est de 31).

Table 12.1

Coefficient de corrélation entre les huit romans et le corpus d'Engwall lorsque le nombre de types est de soixante

Analyse de corrélation (n = 60)								
CLAIR	TICKET	CÁLIN	VIE	ÉTRANGER	VIPÈRES	IMMOR	PORTE	ENG
	0,9538	0,9409	0,8182	0,7914	0,8889	0,8793	0,8341	0,8887
TICKET	CLAIR	CÁLIN	VIE	ÉTRANGER	VIPÈRES	IMMOR	PORTE	ENG
	0,9538	0,9628	0,7938	0,8135	0,9394	0,9282	0,8783	0,9352
CÁLIN	CLAIR	TICKET	VIE	ÉTRANGER	VIPÈRES	IMMOR	PORTE	ENG
	0,9409	0,9628	0,8334	0,8184	0,8909	0,9004	0,8351	0,9041
VIE	CLAIR	TICKET	CÁLIN	ÉTRANGER	VIPÈRES	IMMOR	PORTE	ENG
	0,8182	0,7938	0,8334	0,8272	0,7160	0,7258	0,6810	0,7711
ÉTRANGER	CLAIR	TICKET	CÁLIN	VIE	VIPÈRES	IMMOR	PORTE	ENG
	0,7914	0,8135	0,8184	0,8272	0,7668	0,7748	0,7251	0,7910
VIPÈRES	CLAIR	TICKET	CÁLIN	VIE	ÉTRANGER	IMMOR	PORTE	ENG
	0,8889	0,9394	0,8909	0,7160	0,7668	0,9495	0,9432	0,9125
IMMOR	CLAIR	TICKET	CÁLIN	VIE	ÉTRANGER	VIPÈRES	PORTE	ENG
	0,8793	0,9282	0,9004	0,7258	0,7748	0,9495	0,9558	0,8703
PORTE	CLAIR	TICKET	CÁLIN	VIE	ÉTRANGER	VIPÈRES	IMMOR	ENG
	0,8341	0,8783	0,8351	0,6810	0,7251	0,9432	0,9558	0,8014
ENG	CLAIR	TICKET	CÁLIN	VIE	ÉTRANGER	VIPÈRES	IMMOR	PORTE
	0,8887	0,9352	0,9041	0,7711	0,7910	0,9125	0,8703	0,8014

Table 12.2.

Coefficient de corrélation entre les huit romans et le corpus d'Engwall lorsque le nombre de types est de trente et un

Analyse de corrélation (n = 31)								
CLAIR	TICKET	CÂLIN	VIE	ÉTRANGER	VIPÈRES	IMMOR	PORTE	ENG
	0,9763	0,9504	0,8430	0,9283	0,9246	0,9192	0,8856	0,9485
TICKET	CLAIR	CÂLIN	VIE	ÉTRANGER	VIPÈRES	IMMOR	PORTE	ENG
	0,9763	0,9738	0,8281	0,9481	0,9329	0,9458	0,8971	0,9817
CÂLIN	CLAIR	TICKET	VIE	ÉTRANGER	VIPÈRES	IMMOR	PORTE	ENG
	0,9504	0,9738	0,8696	0,9420	0,9109	0,9183	0,8758	0,9645
VIE	CLAIR	TICKET	CÂLIN	ÉTRANGER	VIPÈRES	IMMOR	PORTE	ENG
	0,8430	0,8281	0,8696	0,8873	0,7309	0,7953	0,7429	0,7886
ÉTRANGER	CLAIR	TICKET	CÂLIN	VIE	VIPÈRES	IMMOR	PORTE	ENG
	0,9283	0,9481	0,9420	0,8873	0,8922	0,9429	0,9050	0,9304
VIPÈRES	CLAIR	TICKET	CÂLIN	VIE	ÉTRANGER	IMMOR	PORTE	ENG
	0,9246	0,9329	0,9109	0,7309	0,8922	0,9650	0,9696	0,9509
IMMOR	CLAIR	TICKET	CÂLIN	VIE	ÉTRANGER	VIPÈRES	PORTE	ENG
	0,9192	0,9458	0,9183	0,7953	0,9429	0,9650	0,9767	0,9472
PORTE	CLAIR	TICKET	CÂLIN	VIE	ÉTRANGER	VIPÈRES	IMMOR	ENG
	0,8856	0,8971	0,8758	0,7429	0,9050	0,9696	0,9767	0,8966
ENG	CLAIR	TICKET	CÂLIN	VIE	VIPÈRES	ÉTRANGER	IMMOR	PORTE
	0,9485	0,9817	0,9645	0,7886	0,9509	0,9304	0,9472	0,8966

Lorsqu'une corrélation a été calculée, comme dans les cas précédents, il est important de savoir si elle représente une relation réelle ou si elle est due au hasard. C'est la table des valeurs significatives de r qui permet de résoudre ce problème: elle

démontre que toutes les corrélations sont significatives même au seuil de 0,01 car elles se situent au-dessus de 0,32 (lorsque $n=60$) et de 0,45 (lorsque $n=31$).

Cette forte corrélation s'explique par l'identité des vocables à haute fréquence d'un texte à l'autre, quoique leur fréquence et leur rang changent (qu'il s'agisse des 60 types ou des 31 types). Burrows offre l'explication suivante pour ce phénomène linguistique: "the shape of the language makes it impossible, in any ordinary sort of English [...] for any word-types like "the" and "of" to slip towards the bottom of our list to be replaced at the top with words like "than" and "more" ("An Ocean" 313). Cette observation s'applique aussi à la langue française et explique, en partie, la forte corrélation observée entre les textes étudiés.

Les r^2 qui indiquent le pourcentage de similarité entre chaque paire de romans ont ensuite été calculés. Ces r^2 seront discutés au fur et à mesure dans ce chapitre. Les limites de confiance de r figurent dans les appendices 12.1 à 12.18. Dans ces appendices les r_1 et r_2 représentent une fourche des valeurs probables autour de la valeur de r . Si on répétait des tests de corrélation sur plusieurs autres romans semblables, l'existence de ces romans étant purement hypothétique, la valeur de ces corrélations varierait entre les deux bornes établies par r_1 et r_2 . Lorsque la valeur de r_2 est inférieure à celle indiquée au seuil de 0,32 ($n=60$) et au seuil de 0,45 ($n=31$), on ne peut pas se fier aux résultats. Toutes les valeurs de r_2 pour les textes dépassent ces

seuils, indiquant que les valeurs de r représentent une différence authentique entre les textes.

Ayant établi qu'il existe une différence entre les textes, il est maintenant question de déterminer la sorte de similarité ou de différence qui existe entre les romans rédigés par un seul auteur.

12.2. Les Textes de Gide

Cette analyse se concentre, premièrement, sur *L'Immoraliste* (1902) et *La Porte étroite* (1909) de Gide (rangées 7 et 8 des tables 12.1 à 12.2), ce qui fournit un indice de similarité entre les oeuvres rédigées par un même auteur à quelques années d'intervalle. Gide maintient dans son *Journal* que malgré les sept années qui séparent leur date de publication, ces deux romans ont coexisté dans son esprit (1135). Il est ainsi légitime de s'attendre à ce qu'il y ait une très forte corrélation entre ces deux romans.

Les plus fortes corrélations se produisent entre *L'Immoraliste* et *La Porte étroite*: elles s'élèvent à 0,9558 (lorsque $n=60$) et à 0,9767 (lorsque $n=31$). La table 12.3 montre les r^2 pour la corrélation entre les deux romans de Gide. Ces corrélations expliquent 91,36% de la variabilité commune entre les données pour *L'Immoraliste* et *La Porte étroite* (lorsque $n=60$) et 95,39% de cette variabilité (lorsque $n=31$).

Seulement entre 8,64% et 4,61% des données ne sont pas expliquées par les corrélations. Il s'agit, bien entendu, de très fortes corrélations.

Table 12.3

Le r^2 entre *L'Immoraliste* et *La Porte étroite*

Titre	n=60	n=31
<i>Porte</i>	91,36%	95,39%

Les limites de confiance de r entre *L'Immoraliste* et les autres oeuvres (voir appendices 12.13 et 12.14) et entre *La Porte étroite* et les autres oeuvres (voir appendices 12.15 et 12.16) corroborent ces résultats. Ces premiers tests mènent à croire que des romans rédigés par un même auteur à quelques années d'intervalle ont tendance à produire de très fortes corrélations.

12.3. Les Textes de Gary/Ajar

Les tables 12.1 et 12.2 présentent la corrélation entre les deux romans publiés sous le nom de Gary, d'une part, et les autres romans étudiés de l'autre. La corrélation la plus significative se trouve, comme l'on s'y attendrait, entre les deux romans signés Gary *Au-delà de cette limite votre ticket n'est plus valable* et *Clair de femme* (0,9763 lorsque $n=31$), aussi bien qu'entre *Au-delà de cette limite votre ticket n'est plus*

valable et *Gros-Câlin* (0,9628 lorsque $n=60$). La table 12.4 indique le r^2 entre *Clair de femme* et quelques autres romans.

Le r^2 pour la corrélation entre *Clair de femme* et *Au-delà de cette limite votre ticket n'est plus valable* explique entre 90,97% de la variabilité entre ces deux romans (lorsque $n=60$) et 95,32% de cette variabilité (lorsque $n=31$). Encore une fois, il s'agit de fortes corrélations. Le calcul des limites de confiance de r (voir appendices 12.1 et 12.2) confirme ces résultats.

Table 12.4

Le r^2 entre *Clair de femme* et quelques autres textes

Titre	$n=60$	$n=31$
<i>Ticket</i>	90,96%	95,32%
<i>Câlin</i>	88,52%	90,32%
<i>La Vie</i>	66,95%	71,07%
<i>Vipères</i>	79,02%	85,49%
<i>Engwall</i>	78,98%	89,96%

Il existe aussi de fortes corrélations entre *Clair de femme* et *Gros-Câlin* signé Ajar, le r^2 expliquant entre 88,52% de la variabilité commune entre ces deux romans (lorsque $n=60$) et 90,32% de cette variabilité (lorsque $n=31$). Les limites de confiance de r (voir les appendices 12.1 et 12.2) confirment la validité de ces corrélations. La corrélation entre *Clair de femme* et *La Vie devant soi* n'est pas aussi forte: elle s'élève à 0,8182 (lorsque $n=60$) et 0,8430 (lorsque $n=31$), le r^2 pour cette corrélation

n'expliquant qu'entre 66,95% de la variabilité commune entre les données (lorsque $n=60$) et 71,07% de la variabilité entre elles (lorsque $n=31$). Les limites de confiance de r entre *Clair de femme* et *La Vie devant soi* (voir appendices 12.1 et 12.2) indiquent que de telles corrélations varieraient à la longue entre 0,7127 et 0,8875 (lorsque $n=60$) et entre 0,6983 et 0,9215 (lorsque $n=31$).

Les appendices 12.1 et 12.2 démontrent que la corrélation entre des romans tels que *Clair de femme* et *Le Noeud de vipères* varierait à la longue entre 0,8204 et 0,9323 (lorsque $n=60$), et entre 0,9631 et 0,8490 (lorsque $n=31$). Ils indiquent aussi que la corrélation entre des romans tels *Clair de femme* et le corpus d'Engwall varierait entre 0,9321 et 0,8201 (lorsque $n=60$) et entre 0,9749 et 0,8955 (lorsque $n=31$). Ceci indique qu'il existe plus de ressemblance entre *Clair de femme* et les autres textes étudiés qu'entre ce même roman et *La Vie devant soi*. La table 12.5 montre le pourcentage de similarité entre *Au-delà de cette limite votre ticket n'est plus valable* et d'autres textes.

Table 12.5

Le r^2 entre *Au-delà de cette limite votre ticket n'est plus valable* et d'autres romans

Titre	$n=60$	$n=31$
<i>Clair</i>	90,96%	95,32%
<i>Câlin</i>	92,69%	94,83%
<i>La Vie</i>	60,74%	68,57%
<i>Vipères</i>	88,25%	87,02%
<i>Engwall</i>	87,45%	96,37%

La corrélation entre *Au-delà de cette limite votre ticket n'est plus valable* et *Gros-Câlin* est la plus forte de toutes, le r^2 expliquant 92,69% de la variabilité commune entre les données (lorsque $n=60$) et 94,83% de cette variabilité (lorsque $n=31$). Parmi les quatre romans de Gary, la corrélation la moins forte se rencontre entre *Au-delà de cette limite votre ticket n'est plus valable* et *La Vie devant soi*, le r^2 n'expliquant qu'entre 60,74% et 68,57% de la variabilité entre les données. En fait, il y a plus de similarité entre *Au-delà de cette limite votre ticket n'est plus valable* et *Le Noeud de vipères*, et entre ce roman et le corpus d'Engwall, qu'entre *Au-delà de cette limite votre ticket n'est plus valable* et *La Vie devant soi*.

Les tests de corrélation entrepris entre *Clair de femme* et *Au-delà de cette limite votre ticket n'est plus valable* démontrent presque la même tendance que celle observée entre les deux romans de Gide, à savoir qu'il existe une très forte corrélation entre les deux romans rédigés par le même auteur.

Table 12.6

Le r^2 entre *Gros-Câlin* et d'autres oeuvres

Titre	n=60	n=31
<i>La Vie</i>	69,45%	75,62%
<i>Étranger</i>	66,98%	88,74%
<i>Vipères</i>	79,38%	82,97%
<i>Porte</i>	69,74%	76,71%
<i>Engwall</i>	81,75%	93,02%

La table 12.6 démontre que la corrélation entre *Gros-Câlin* et *La Vie devant soi*, également signé Ajar, est moins significative que celle obtenue pour les deux romans signés Gary. Elle n'explique que 69,45% de la variabilité entre les données (lorsque $n=60$) et 75,62% de cette variabilité (lorsque $n=31$): il y a moins de similarité entre ces deux romans signés Ajar.

Il existe plus de similarité entre *Gros-Câlin* et certains textes étudiés (*L'Étranger*, *Le Noeud de vipères*, *La Porte étroite* et le corpus d'Engwall, par exemple) qu'entre ce roman et *La Vie devant soi*. Le calcul des limites de confiance de r entre *Gros-Câlin* et les autres oeuvres confirme la validité de ces résultats (voir appendices 12.5 et 12.6). La table 12.7 montre le r^2 entre *La Vie devant soi* et quelques autres textes.

Table 12.7

Le r^2 entre *La Vie devant soi* et quelques autres oeuvres

Texte	n=60	n=31
<i>Câlin</i>	69,45%	75,62%
<i>Étranger</i>	68,43%	78,74%
<i>Clair</i>	66,95%	71,07%
<i>Ticket</i>	60,74%	68,57%
<i>Engwall</i>	59,46%	62,19%

La corrélation entre *La Vie devant soi* et *Gros-Câlin* est la plus significative de toutes, expliquant 69,45% de la variabilité entre les données (lorsque $n=60$) et 75,62% de cette variabilité (lorsque $n=31$). Elle est suivie de la corrélation entre *La Vie devant soi* et *L'Étranger*, expliquant 68,43% de la variabilité entre les données (lorsque $n=60$) et 78,74% (lorsque $n=31$).

Les limites de confiance de r indiquées dans les appendices 12.7 et 12.8 confirment la validité de ces résultats et démontrent que si on répétait les tests de corrélation plusieurs fois entre des romans semblables aux deux romans signés Ajar, la corrélation varierait à la longue entre 0,8972 et 0,7354 (lorsque $n=60$) et entre 0,9353 et 0,7461 (lorsque $n=31$). On aurait pu s'attendre à ce qu'il y ait plus de correspondance entre les deux romans signés Ajar, mais les tests de corrélation indiquent le contraire. Il est vrai que *La Vie devant soi* s'accorde le plus avec *Gros-Câlin* (lorsque $n=60$), avec *L'Étranger* (lorsque $n=31$) aussi bien qu'avec les deux autres romans de Gary, mais des quatre romans de Gary étudiés, *La Vie devant soi* s'accorde le moins avec les autres romans de Gary.

12.4. Ce que révèlent les corrélations et les limites de confiance de r

Les tests de corrélation démontrent que, généralement, les plus grandes ressemblances se rencontrent entre deux romans rédigés par le même auteur, tels qu'entre *La Porte étroite* et *L'Immoraliste* de Gide et entre *Clair de femme* et *Au-delà*

de cette limite votre ticket n'est plus valable signés Gary, aussi bien qu'entre les deux romans signés Gary et le premier roman signé Ajar, *Gros-Câlin*. La forte corrélation entre les deux romans de Gide n'étonne nullement pour les raisons élaborées plus tôt. Quant aux deux romans signés Gary et *Gros-Câlin* les fortes corrélations entre eux ne sont pas étonnantes, non plus, car *Gros-Câlin* est la première tentative de Gary de se dissimuler sous la signature Ajar.

Une grande corrélation s'observe entre *La Vie devant soi* et *L'Étranger*. Ce phénomène mis à part, *La Vie devant soi* s'accorde le plus avec les deux romans signés Gary et avec *Gros-Câlin*. Néanmoins, la corrélation observée entre *La Vie devant soi* et les trois autres romans de Gary/Ajar n'est pas aussi forte qu'entre les autres romans rédigés par le même auteur.

Les tests entrepris dans ce chapitre démontrent que *La Vie devant soi* se détache le plus des trois autres romans de Gary; ils confirment les résultats des tests effectués dans les chapitres précédents.

Chapitre 13

La Fréquence des synonymes comme indice de style

Tout art est choix [...] et l'art du style est l'art de choisir entre les possibilités d'expression qui s'offrent dans chaque cas donné à l'usager de la langue" (Marouzeau 98).

Ce chapitre examine la répartition des vocables synonymes dans les textes, en partant du principe de Mosteller et Wallace que chaque auteur a une préférence pour l'un ou l'autre vocable synonyme, et que le choix que fait l'auteur entre ces deux vocables constitue un indice de son style. Dans l'usage courant un synonyme se définit de la façon suivante: lorsque deux vocables appartenant à une même classe grammaticale peuvent être remplacés l'un par l'autre sans que le sens général de la

phrase ne soit changé, ils sont synonymes. Cette définition reconnaît les équivalences sémantiques globales tout en conservant les nuances de sens.

L'étude se base sur la répartition des paires de synonymes les plus évidents dans les romans étudiés et dans le corpus d'Engwall. Elles sont les suivantes: *alors/donc*, *aucun(e)/nul(le)*, *ça/cela*, *commencer/débuter*, *commencement/début*, *croire/penser*, *demeurer/habiter*, *devoir/falloir*, *lorsque/quand*, *parce qu'(e)/puisque(e)*. La table 13.1 présente la fréquence de ces synonymes dans le corpus d'Engwall par ordre alphabétique.

Table 13.1

Engwall (500.000 vocables)

<i>A</i>	<i>Fr. Abs.</i>	<i>Fr. Rel.</i>	<i>B</i>	<i>Fr. Abs.</i>	<i>Fr. Rel.</i>
Alors	578	1,15	Donc	279	0,56
Aucun(e)	209	0,42	Nul(le)	46	0,09
Ça	1.084	2,17	Cela	584	1,17
Commencer	230	0,46	Débuter	2	0,01
Commencement	5	0,01	Début	63	0,13
Croire	595	1,19	Penser	577	1,15
Demeurer	47	0,09	Habiter	74	0,15
Devoir	782	1,56	Falloir	712	1,42
Lorsque	160	0,32	Quand	953	1,90
Parce qu'(que)	379	0,76	Puisqu'(que)	112	0,22

13.1. La Répartition des synonymes

Pour trouver la fréquence de ces synonymes dans le corpus d'Engwall, il a suffi de consulter le texte *Fréquence et distribution du vocabulaire dans un choix de*

romans français. Pour toutes les occurrences des verbes *débuter* et *commencer*, *devoir* et *falloir*, l'ensemble des formes flexionnelles que peuvent prendre ces verbes a été totalisée.

Pour trouver la fréquence de ces mêmes types dans les romans, il a fallu les répertorier dans la concordance, les compter et les additionner. Les appendices 13.1.1 à 13.1.8 montrent la fréquence de ces synonymes dans les huit romans et suivent le même système de classement.

13.2. Le Test du χ^2 pour les neuf textes

Le test du χ^2 permet de calculer si, du point de vue statistique, la distribution de ces synonymes diffère de façon significative dans les neuf textes. Dans ce chapitre le test du χ^2 est effectué par l'intermédiaire d'une table de contingence seulement. Ce test comporte certaines restrictions: notamment que les valeurs théoriques ne peuvent être inférieures à cinq, et qu'il ne peut être effectué que sur des valeurs absolues et non pas sur des proportions (telle que sur la fréquence relative).

L'appendice 13.2.1 présente les valeurs observées et la table de contingence, et l'appendice 13.2.2 présente les valeurs calculées pour ces synonymes. La table 13.2 regroupe les valeurs observées et les valeurs calculées (indiquées entre parenthèses) dans les neuf textes. Toute valeur calculée inférieure à cinq, par exemple *aucun(e)/nul(le)*, est exclue des calculs.

Table 13.2

Les Neuf Textes: valeurs observées et valeurs calculées

Lemmes	Clair	Tic	Câlin	Vie	Étran	Vip	Imm	Porte	Eng
Alors	44 (47,37)	66 (73,54)	64 (84,99)	78 (135,64)	107 (46,49)	27 (56,21)	52 (38,81)	36 (35,20)	578 (532,74)
Donc	18 (24,05)	26 (36,56)	50 (42,25)	20 (67,43)	10 (23,11)	17 (27,95)	67 (19,29)	36 (17,50)	279 (264,85)
Ça	132 (111,60)	249 (169,66)	340 (196,07)	549 (312,92)	24 (107,26)	45 (129,69)	00 (89,53)	4 (81,21)	1084 (1.229,06)
Cela	62 (47,32)	34 (71,93)	54 (83,13)	6 (132,67)	98 (45,48)	65 (54,99)	59 (37,96)	67 (34,43)	584 (521,10)
Croire	52 (52,24)	62 (79,41)	74 (91,77)	88 (146,47)	36 (50,20)	101 (60,70)	67 (41,91)	61 (38,01)	595 (575,28)
Penser	28 (48,01)	43 (72,98)	71 (84,34)	99 (134,60)	62 (46,14)	67 (55,79)	52 (38,51)	45 (34,93)	577 (528,69)
Devoir	62 (63,50)	118 (96,54)	65 (111,57)	86 (178,05)	46 (61,05)	82 (73,79)	78 (50,95)	62 (46,21)	782 (699,35)
Falloir	79 (64,61)	106 (98,22)	113 (113,51)	150 (181,51)	51 (62,09)	123 (75,08)	37 (51,83)	34 (47,01)	712 (711,51)
Lorsque	35 (23,73)	85 (36,07)	87 (41,69)	45 (66,53)	10 (22,80)	56 (27,57)	13 (19,04)	25 (17,27)	160 (261,31)
Quand	45 (79,73)	67 (121,22)	76 (140,08)	324 (223,57)	77 (76,63)	77 (92,66)	59 (63,97)	56 (58,02)	953 (878,11)
Parce que	57 (46,21)	73 (70,26)	86 (81,19)	280 (129,58)	63 (44,42)	43 (53,70)	9 (37,08)	15 (33,63)	379 (508,94)
Puisque	3 (7,63)	9 (11,60)	4 (13,41)	5 (21,40)	9 (7,34)	14 (8,87)	2 (6,12)	8 (5,55)	112 (84,06)

Les valeurs du χ^2 pour les neuf textes figurent dans l'appendice 13.2.3. En appliquant la correction de Bonferroni, on obtient le seuil de rejet de 0,046 (à savoir, 5/108); et un ddl de 88, à savoir (12-1)(9-1). Étant donné que 88 ddl ne figurent pas dans la liste des seuils de rejet, 90 ddl ont été choisis dont les valeurs à dépasser. Les valeurs significatives sont alors toutes celles qui sont supérieures à 124,3. Elles sont indiquées en chiffre gras dans l'appendice 13.2.3. Les valeurs globales du χ^2 pour chaque texte sont regroupées dans la table 13.3.

Table 13.3

Valeurs du χ^2 : les Neuf Textes

<i>Clair</i>	47,60	<i>Ticket</i>	173,76	<i>Câlin</i>	232,92	<i>Vie</i>	681,90
<i>Étran</i>	241,98	<i>Vipères</i>	174,16	<i>Immor</i>	288,33	<i>Porte</i>	164,50
Engwall	132,30						

Valeur globale du $\chi^2 = 2.137,45$

ddl: 88

Le nombre de degrés de liberté pour l'ensemble des textes est de 88, étant donné que 88 ddl ne figurent pas dans la liste des seuils de rejet, 90 ddl ont été choisis dont les valeurs à dépasser sont 118,136 (au seuil de 0,025) et 124,116 (au seuil de 0,01). Comme ces seuils sont largement dépassés, on peut s'assurer que du point de vue statistique, la répartition des synonymes est très différente dans les neuf textes. Ceci est semblable à la tendance générale observée dans les trois chapitres précédents.

Le nombre de degrés de liberté pour chaque roman est de 11; les seuils à dépasser sont 21,92 (au seuil de 0,025), 24,73 (au seuil de 0,01) et 26,76 (au seuil de 0,005). Comme la statistique du χ^2 obtenue pour *La Vie devant soi* (681,9) est la plus significative de toutes, elle démontre que ce roman contribue le plus à cette forte valeur du χ^2 .

13.3. Le Test du χ^2 pour les huit romans

On pourrait postuler que la très haute valeur du χ^2 obtenue pour les neuf textes résulte en partie du fait que ces textes sont hétérogènes, c'est-à-dire qu'ils sont composés de huit romans et du corpus d'Engwall (qui est constitué d'une sélection d'extraits des ouvrages de plusieurs auteurs). Pour écarter l'éventualité que le corpus d'Engwall aurait pu influencer cette valeur du χ^2 , il sera exclu des prochains tests.

Les valeurs calculées pour les huit romans sont obtenues à partir des valeurs observées (voir les appendices 13.3.1 et 13.3.2) et les valeurs du χ^2 sont indiquées en détail dans l'appendice 13.3.3. La dernière rangée de l'appendice indique la valeur globale du χ^2 pour chaque roman, valeurs qui sont présentées dans la table 13.4.

Table 13.4

Valeurs du χ^2 : les huit romans

<i>Clair</i>	31,52	<i>Ticket</i>	97,17	<i>Câlin</i>	121,47	<i>Vie</i>	443,58
<i>Étran</i>	304,40	<i>Vipères</i>	165,27	<i>Imm</i>	290,51	<i>Porte</i>	176,89

Valeur globale du χ^2 : 1630,82

ddl: 63

Le nombre de degrés de liberté pour l'ensemble des huit romans est de 63.

Étant donné que 63 ddl ne figurent pas dans la liste de seuils de rejet, 70 ddl ont été choisis dont les valeurs à dépasser sont 95,023 (au seuil de 0,025), 100,425 (au seuil

de 0,01) et 104,215 (au seuil de 0,005). La valeur globale du χ^2 pour les huit romans dans son ensemble (1630,82) dépasse de très loin ces seuils.

Lorsque le test du χ^2 est effectué sur les huit romans seulement, on observe des différences fondamentales dans la répartition des synonymes, ce qui démontre que la haute valeur du χ^2 obtenue pour le test précédent ne résulte pas du fait que le corpus d'Engwall est inclus dans ce test. Par ailleurs, sauf pour *Clair de femme*, toutes les valeurs individuelles du χ^2 sont significatives et dépassent de très loin le seuil de signification. C'est *La Vie devant soi*, qui contribue le plus à la valeur du χ^2 (443,58), indiquant que c'est le roman le plus remarquable de tous.

Les deux valeurs du χ^2 obtenues jusqu'ici démontrent beaucoup de variabilité d'un texte à l'autre. Il se pourrait que ceci résulte du fait que les textes étudiés sont rédigés par des auteurs différents. Il est alors légitime d'examiner la distribution des données en fonction des textes rédigés par un seul auteur afin d'établir le genre de variabilité qui existe entre eux.

13.4. Le Test du χ^2 pour *L'Immoraliste* et *La Porte étroite*

Les appendices 13.4.1 et 13.4.2 fournissent les valeurs observées et les valeurs calculées pour les deux romans de Gide et l'appendice 13.4.3 montre en détail les valeurs du χ^2 pour chaque synonyme. Comme aucune valeur ne dépasse le seuil de

signification, ce qui indique qu'il n'existe aucune différence statistique dans la fréquence des synonymes d'un roman à l'autre. Les valeurs globales du χ^2 pour chacun de ces deux romans sont regroupées dans la table 13.5.

Table 13.5

Valeurs du χ^2 : *Immoraliste* et *Porte*

<i>Immoraliste</i>	7,58	<i>Porte</i>	9,11
ddl:11			
Valeur globale du χ^2 : 16,69			

Le nombre de degrés de liberté pour les deux romans est de 11, les valeurs à dépasser sont 21,92 (au seuil de 0,025), 24,72 (au seuil de 0,01) et 26,76 (au seuil de 0,005). Comme la valeur globale du χ^2 pour ces deux romans (16,69) ne dépasse pas ces seuils, ceci indique qu'il n'existe aucune différence significative dans la répartition des synonymes dans ces deux romans, même au seuil de 0,025. Les résultats du χ^2 obtenus pour les deux romans de Gide suggèrent qu'un auteur peut avoir un style fort semblable dans deux textes différents:

13.5. Le Test du χ^2 pour *Clair et Ticket*, *Câlin* et *La Vie*

Étant donnés les résultats du χ^2 obtenus pour les deux romans de Gide, on pourrait s'attendre à ce qu'il n'existe presque aucune différence dans la distribution

des synonymes dans les deux romans signés Gary aussi bien que dans les deux romans signés Ajar. Pour vérifier cette hypothèse, les deux romans de Gary, *Clair de femme* et *Au-delà de cette limite votre ticket n'est plus valable*, sont comparés entre eux (Voir les appendices 13.5.1 et 13.5.2 pour les valeurs observées et les valeurs calculées). Les résultats du test du χ^2 sont indiqués dans l'appendice 13.5.3 et regroupés dans la table 13.6.

Comme la valeur globale du χ^2 (38,99) dépasse les valeurs significatives de 19,02 (au seuil de 0,025), de 21,67 (au seuil de 0,01) et de 23,59 (au seuil de 0,005), il existe une différence significative dans la fréquence des synonymes dans les deux romans signés Gary. Le style de ces deux romans est moins semblable que celui des deux romans de Gide.

Table 13.6

Valeurs du χ^2 : *Clair* et *Ticket*

<i>Clair</i>	23,62	<i>Ticket</i>	15,37
Valeur globale du χ^2 : 38,99			
ddl: 9			

Lorsque le même test est appliqué aux deux romans signés Ajar, *Gros-Câlin* et *La Vie devant soi*, on obtient les résultats du χ^2 indiqués dans l'appendice 13.6.3. Ils

sont regroupés dans la table 13.7 (voir les appendices 13.6.1 et 13.6.2 pour les valeurs observées et les valeurs calculées).

Table 13.7

Valeurs du χ^2 : *Câlin* et *La Vie*

<i>Câlin</i>	139,9	<i>Vie</i>	344,64
--------------	-------	------------	--------

ddl: 11

valeur globale du χ^2 : 484,54

La valeur globale du χ^2 de 484,54 dépasse de très loin les valeurs de 21,92 (au seuil de 0,025), de 24,73 (au seuil de 0,01) et de 26,76 (au seuil de 0,005). La statistique du χ^2 pour les deux romans signés Ajar est bien plus élevée que celle obtenue pour les deux romans signés Gary: elle est tellement forte qu'elle ne paraît pas dans les tables de valeurs significatives à 11 ddl.

Les résultats du χ^2 obtenus lorsque des romans rédigés par les mêmes auteurs sont comparés entre eux démontrent que, chez Gide, aucune différence statistiquement significative n'est perceptible dans la répartition des synonymes. Par contre, il existe de grandes variations dans les quatre romans rédigés par Gary; la plus grande valeur du χ^2 se trouve entre *Gros-Câlin* et *La Vie devant soi*.

13.6. Conclusions

La très forte valeur du χ^2 obtenue lorsque les neuf textes sont comparés entre eux indique que la répartition des synonymes dans l'ensemble des neuf textes est très différente. Lorsque le corpus d'Engwall est exclu de l'étude, la valeur du χ^2 est significative encore une fois, indiquant qu'il existe une différence fondamentale dans la répartition des synonymes dans ces huit romans. Dans les deux cas, c'est le deuxième roman signé Ajar, *La Vie devant soi* qui contribue le plus à cette haute valeur du χ^2 .

Lorsque la fréquence des synonymes dans *L'Immoraliste* (Gide) est comparée à la fréquence de ces mêmes synonymes dans *La Porte étroite* (Gide), aucune différence statistiquement significative n'est perceptible entre ces deux romans, ce qui démontre que le style d'un auteur peut demeurer le même dans différents textes. Lorsque les deux romans signés Ajar sont comparés entre eux, une très haute valeur du χ^2 est obtenue; elle est bien plus élevée que la valeur obtenue pour les deux romans signés Gary. Ceci indique qu'il existe moins de similarités entre *Gros-Câlin* et *La Vie devant soi* qu'entre *Clair de femme* et *Au-delà de cette limite votre ticket n'est plus valable*.

Tous les tests effectués dans ce chapitre démontrent que *La Vie devant soi* produit la plus haute valeur du χ^2 , indiquant qu'il diffère non seulement des trois

autres romans de Gary mais encore des autres romans étudiés. Les tests entrepris dans ce chapitre corroborent les résultats de ceux des trois chapitres précédents et démontrent que *La Vie devant soi* est le roman le plus extraordinaire des quatre romans de Gary étudiés. Ces résultats réfutent l'hypothèse selon laquelle le style d'un auteur constitue une manière singulière de s'exprimer lorsque le genre est le même (Ledger 235).

Le prochain chapitre offre une explication des phénomènes linguistiques et stylistiques révélés dans ce chapitre et dans les trois qui le précèdent.

Chapitre 14

Conclusion

La vérité est que j'ai été profondément atteint par la plus vieille tradition protéenne de l'homme: celle de la multiplicité [...] Je me suis toujours été un autre [sic] (Gary, *Vie et Mort* 18).

L'objet de ce travail était de déterminer si, comme le préconise Buffon "Le style c'est l'homme même" (Galdi 9), ce qui indique qu'un auteur ne peut le changer pas plus qu'il ne peut changer sa personnalité, ou si au contraire, tel que le préconise Spitzer (89), l'auteur peut changer certains aspects de son style quand il le veut. Les résultats de plusieurs tests statistiques entrepris pour résoudre ce problème, convergent pour mettre en lumière certaines caractéristiques indéniables.

14.1. Ce que révèlent les tests

La plupart des tests statistiques effectués (le test-*t* de Student-Fisher, le test du χ^2 et la corrélation de Pearson) démontrent que, généralement, les similarités les plus significatives se rencontrent entre les romans rédigés par le même auteur (telles qu'entre *La Porte étroite* et *L'Immoraliste* de Gide et entre *Clair de femme* et *Au-delà de cette limite votre ticket n'est plus valable* de Gary). Ces tests semblent confirmer le dicton de Buffon que "le style c'est l'homme même", c'est-à-dire qu'un auteur possède un style unique qui le distingue des autres.

Or, les similarités entre les deux romans signés Gary, aussi bien qu'entre ces deux romans et *Gros-Câlin*, n'étonnent nullement, car *Gros-Câlin*, paru en 1974, est le premier texte que Gary publie sous le pseudonyme d'Ajar. Le souci de distinguer son style de celui d'Ajar ne le préoccupe pas encore, et ne vient que plus tard.

D'autres caractéristiques se dégagent aussi. Le chapitre 10 démontre que le deuxième roman signé Ajar, *La Vie devant soi*, diffère considérablement des deux romans signés Gary aussi bien que de *Gros-Câlin*. De plus, les tests du χ^2 et de corrélation produisent des résultats forts similaires et démontrent, eux aussi, que *La Vie devant soi* s'éloigne le plus des autres romans de Gary. Ces tests contredisent ainsi l'hypothèse de Buffon et corroborent celle de Spitzer, et démontrent qu'un auteur peut adopter un autre style quand il le veut. En fait, Romain Gary a si bien réussi à changer de style dans *La Vie devant soi* qu'on peut être excusé de penser qu'il n'en est pas l'auteur.

Et pourtant, tout le monde sait que Romain Gary et Émile Ajar sont deux aspects d'un seul être (voir chapitres 2 et 4). Par ailleurs, le chapitre 5 démontre que souvent les mêmes thèmes, les mêmes personnages et les mêmes motifs se retrouvent dans *La Vie devant soi* aussi bien que dans les autres romans de Gary/Ajar.

Il n'est pas surprenant qu'un écrivain aussi habile que Romain Gary ait pu manipuler la fréquence avec laquelle les synonymes se produisent dans ses romans (étant donné leur faible fréquence). Il est plus difficile de comprendre comment Gary a réussi à contrôler la fréquence des vocables les plus nombreux dans ces mêmes romans.

On ne peut attribuer ce changement de style à la possibilité que le style de Gary a évolué, car seulement quatre ans séparent la publication de ses quatre romans. Par ailleurs, *La Vie devant soi* a paru en 1975, l'année durant laquelle Gary fait publier *Au-delà de cette limite votre ticket n'est plus valable*. De plus, *La Vie devant soi* est intercalé entre *Gros-Câlin* (1974) et *Clair de femme* (1977) et ne ressemble nullement à ces deux romans non plus, tandis que ces deux romans se ressemblent entre eux.

Tout ceci donne matière à réflexion. Dans la plupart des cas concernant l'attribution d'auteurs aux oeuvres anonymes ou pseudonymes décrits dans le chapitre 8, les chercheurs se concentrent sur certains aspects du style, mais dirigent leur attention sur des situations présumées ou hypothétiques. Leurs recherches désignent à un certain seuil de probabilité, Shakespeare, Marlowe, Cholokhov ou Krynkow comme

l'auteur présumé d'une oeuvre dont la paternité a été contestée. Mais les résultats statistiques obtenus pour *La Vie devant soi* indiquent que les méthodes traditionnelles d'identification de textes (telles que l'analyse des vocables les plus fréquents) ne produisent pas les résultats auxquels on pourrait s'attendre, ce qui met en doute le bien-fondé de ces méthodes.

14.2. Explications

Vraisemblablement plusieurs facteurs ont contribué au phénomène observé dans ce travail. Le chapitre 2 démontre à quel point les critiques ont cherché à dénigrer l'oeuvre de Gary dès la parution des *Racines du ciel* et comment ils ont accusé Gary de plagiat et de mal manier la langue française. L'on sait aussi combien Gary a voulu s'échapper des idées préconçues de ces critiques. On est tentée d'accepter l'explication de Catonné que Gary a publié des romans sous un pseudonyme parce qu'il voulait se venger de ces critiques (122).

Les chapitres 2 et 3 expliquent aussi comment quelques rares critiques ont cru reconnaître le style de Gary sous le pseudonyme d'Ajar après la parution de *Gros-Câlin*. Il est légitime de postuler que Romain Gary décide alors de faire taire les soupçons qu'il est Ajar en entreprenant la rédaction de *La Vie devant soi*. Il se détache de son style dans ce roman, en adopte un autre pour déjouer ces critiques. Ceci aide à comprendre, en partie, pourquoi, de son vivant, il n'a jamais avoué être Émile Ajar, et pourquoi, il l'a si catégoriquement répudié.

Vagn Outzen propose que Romain Gary voulait trouver une voix qui ne se perde pas dans la foule des écrivains et des moralisateurs de son époque et qu'il a dû hurler plus fort que les autres pour se faire entendre (324). Encore une fois, ceci explique comment il a pu se faire couronner deux fois du prix Goncourt, mais n'explique pas le changement de style observé dans *La Vie devant soi*. Paul Pavlowitch attribue la métamorphose de Gary en Ajar au fait que Gary était bâtard: "Il y a dans cette absence d'origine, la raison essentielle de sa trajectoire fantasque" (21). Une fois de plus, ceci explique le subterfuge Ajar, mais n'explique pas le changement de style observé dans *La Vie devant soi*.

Le chapitre deux décrit à quel point Gary se sentait tiraillé entre sa nature divisée de mi-juif, mi-slave, et comment il revendiquait sa bâtardise raciale d'où découle son alliage avec les maraudeurs des steppes et la souffrance des ghettos juifs d'Europe. Le chapitre 5 démontre que le thème de dédoublement et du *dibbouk* est l'un des plus importants dans l'oeuvre de Gary. Il est possible que Romain Gary ait réussi à créer son propre *dibbouk* et s'est métamorphosé en Ajar, ce qui expliquerait ce changement de style dans *La Vie devant soi*.

Le point de vue qu'on adopte pour expliquer l'une des mystifications qui a le plus réussi dans le monde littéraire (tant du point de vue psychologique que stylistique) importe peu. Ce qui est important c'est que cette étude a pu démontrer que la fréquence des vocables les plus nombreux aussi bien que la distribution des synonymes ne constituent pas toujours des indices du style ni de la personnalité d'un auteur.

Ces résultats mettent ainsi en doute le bien-fondé des méthodes traditionnelles d'identification d'auteurs, surtout celles des auteurs qui veulent cacher leur identité. Ils rendent contestable l'utilisation de ces méthodes dans un milieu medico-légal, telles qu'elles ont été mises en pratique par Morton à la Cour d'assises de Londres.

Appendices au chapitre 10

Appendice 10.1.1

Clair de femme

Rang	Type	Total	Fréquence relative	Rang	Type	Total	Fréquence relative
1	de	1143	33,86	31	dans	223	6,61
2	je	896	26,54	32	pour	209	6,19
3	et	736	21,80	33	me	203	6,01
4	il	658	19,49	34	mais	193	5,72
5	la	569	16,86	35	était	193	5,72
6	un	559	16,56	36	du	180	5,33
7	à	511	15,14	37	plus	175	5,18
8	est	508	15,05	38	se	168	4,98
9	le	502	14,87	39	tout	157	4,65
10	pas	483	14,31	40	ai	155	4,59
11	que	474	14,04	41	avait	148	4,38
12	vous	473	14,01	42	moi	146	4,33
13	l'	448	13,27	43	au	146	4,33
14	ne	419	12,41	44	s'	145	4,30
15	une	391	11,58	45	m'	145	4,30
16	d'	391	11,58	46	tu	141	4,18
17	les	364	10,78	47	ça	132	3,91
18	c'	321	9,51	48	si	112	3,32
19	n'	300	8,89	49	sur	112	3,32
20	elle	300	8,89	50	comme	112	3,32
21	ce	289	8,56	51	avec	108	3,20
22	a	288	8,53	52	mon	96	2,84
23	qui	281	8,32	53	encore	96	2,84
24	nous	265	7,85	54	suis	94	2,78
25	j'	256	7,58	55	lui	93	2,76
26	des	255	7,55	56	femme	90	2,67
27	qu'	251	7,44	57	bien	90	2,67
28	on	248	7,35	58	même	87	2,58
29	y	237	7,02	59	là	85	2,52
30	en	232	6,87	60	Lydia	82	2,43

Appendice 10.1.2

Au-delà de cette limite votre ticket n'est plus valable

Rang	Type	Total	Fréquence relative	Rang	Type	Total	Fréquence relative
1	de	2290	39,54	31	était	345	5,96
2	je	1477	25,50	32	tu	336	5,80
3	et	1459	25,19	33	du	324	5,59
4	la	1201	20,74	34	elle	321	5,54
5	il	1009	17,42	35	se	303	5,23
6	à	1007	17,39	36	y	300	5,18
7	un	891	15,38	37	m'	297	5,13
8	que	863	14,90	38	au	296	5,11
9	l'	811	14,00	39	mon	292	5,04
10	le	800	13,81	40	mais	289	4,99
11	d'	749	12,93	41	avait	287	4,96
12	une	694	11,98	42	avec	271	4,68
13	pas	647	11,17	43	on	257	4,44
14	les	635	10,96	44	moi	252	4,35
15	est	612	10,57	45	ça	248	4,28
16	ne	599	10,34	46	nous	232	4,01
17	en	546	9,43	47	sur	231	3,99
18	me	503	8,69	48	s'	230	3,97
19	des	501	8,65	49	ai	217	3,75
20	j'	484	8,36	50	tout	211	3,64
21	dans	437	7,55	51	même	210	3,63
22	ce	437	7,55	52	son	182	3,14
23	qu'	434	7,49	53	ma	174	3,00
24	c'	423	7,30	54	si	173	2,99
25	vous	418	7,22	55	par	170	2,94
26	qui	410	7,08	56	comme	169	2,92
27	n'	387	6,68	57	bien	163	2,81
28	pour	380	6,56	58	encore	158	2,73
29	a	351	6,06	59	lui	157	2,71
30	plus	346	5,97	60	Laura	141	2,43

Appendice 10.1.3

Gros-Câlin

Rang	Type	Total	Fréquence relative	Rang	Type	Total	Fréquence relative
1	de	1899	35,65	31	ce	356	6,68
2	je	1395	26,19	32	ai	348	6,53
3	et	1294	24,29	33	ça	339	6,36
4	à	1121	21,05	34	mais	321	6,03
5	la	914	17,16	35	y	304	5,71
6	il	899	16,88	36	du	298	5,59
7	le	849	15,94	37	au	291	5,46
8	est	827	15,53	38	tout	290	5,44
9	que	816	15,32	39	était	283	5,31
10	l'	782	14,68	40	se	280	5,26
11	un	704	13,22	41	elle	251	4,71
12	pas	693	13,01	42	m'	243	4,56
13	les	693	13,01	43	moi	242	4,54
14	c'	594	11,15	44	mon	230	4,32
15	d'	593	11,13	45	s'	227	4,26
16	j'	577	10,83	46	plus	227	4,26
17	en	566	10,63	47	même	224	4,21
18	dans	529	9,93	48	sur	210	3,94
19	une	522	9,80	49	comme	203	3,81
20	me	510	9,57	50	suis	188	3,53
21	des	508	9,54	51	bien	180	3,38
22	ne	475	8,92	52	avait	175	3,29
23	qui	451	8,47	53	là	170	3,19
24	pour	449	8,43	54	par	168	3,15
25	qu'	442	8,30	55	lui	166	3,12
26	vous	417	7,83	56	fait	160	3,00
27	on	412	7,74	57	dit	148	2,78
28	avec	398	7,47	58	très	144	2,70
29	a	396	7,43	59	nous	143	2,68
30	n'	395	7,42	60	ils	139	2,61

Appendice 10.1.4

La Vie devant soi

Rang	Type	Total	Fréquence relative	Rang	Type	Total	Fréquence relative
1	et	1989	30,56	31	Rosa	499	7,67
2	de	1518	23,32	32	mais	473	7,27
3	il	1268	19,48	33	une	467	7,17
4	je	1246	19,14	34	plus	452	6,94
5	elle	1157	17,77	35	dans	442	6,79
6	que	1103	16,94	36	d'	420	6,45
7	pas	1059	16,27	37	ce	411	6,31
8	à	1018	15,64	38	comme	408	6,27
9	la	998	15,33	39	me	407	6,25
10	le	989	15,19	40	y	395	6,07
11	est	958	14,72	41	vous	393	6,04
12	a	901	13,84	42	s'	365	5,61
13	j'	892	13,70	43	se	364	5,59
14	un	868	13,33	44	tout	349	5,36
15	qu'	852	13,09	45	avec	343	5,27
16	les	824	12,66	46	lui	328	5,04
17	c'	776	11,92	47	quand	324	4,98
18	était	763	11,72	48	m'	320	4,92
19	qui	758	11,64	49	ils	315	4,84
20	des	686	10,54	50	même	307	4,72
21	avait	670	10,29	51	parce	280	4,30
22	l'	662	10,17	52	du	271	4,16
23	pour	659	10,12	53	monsieur	257	3,95
24	ne	630	9,68	54	moi	245	3,76
25	ai	624	9,59	55	suis	235	3,61
26	en	604	9,28	56	encore	234	3,59
27	madame	576	8,85	57	bien	232	3,56
28	ça	549	8,43	58	son	214	3,29
29	n'	540	8,30	59	fait	210	3,23
30	on	532	8,17	60	rien	204	3,13

Appendice 10.1.5

L'Étranger

Rang	Type	Total	Fréquence relative	Rang	Type	Total	Fréquence relative
1	de	1054	30,44	31	c'	233	6,73
2	et	942	27,21	32	s'	228	6,59
3	il	935	27,01	33	elle	226	6,53
4	a	793	22,90	34	ce	226	6,53
5	je	739	21,34	35	des	223	6,44
6	le	711	20,54	36	qui	201	5,81
7	la	649	18,74	37	pour	186	5,37
8	que	639	18,46	38	du	183	5,29
9	j'	632	18,25	39	nous	180	5,20
10	à	572	16,52	40	mon	178	5,14
11	ai	552	15,94	41	avais	178	5,14
12	l'	536	15,48	42	tout	175	5,05
13	un	503	14,53	43	sur	168	4,85
14	m'	414	11,96	44	avec	165	4,77
15	qu'	385	11,12	45	on	163	4,71
16	était	366	10,57	46	au	159	4,59
17	les	353	10,20	47	moi	152	4,39
18	d'	344	9,94	48	plus	150	4,33
19	pas	343	9,91	49	son	138	3,99
20	est	338	9,76	50	peu	115	3,32
21	en	300	8,66	51	si	113	3,26
22	me	286	8,26	52	y	112	3,23
23	avait	281	8,12	53	sa	109	3,15
24	n'	278	8,03	54	alors	107	3,09
25	une	275	7,94	55	suis	106	3,06
26	mais	259	7,48	56	se	105	3,03
27	ne	243	7,02	57	même	98	2,83
28	dit	243	7,02	58	cela	98	2,83
29	lui	241	6,96	59	fait	96	2,77
30	dans	234	6,76	60	étais	94	2,71

Appendice 10.1.6

Le Noeud de Vipères

Rang	Type	Total	Fréquence relative	Rang	Type	Total	Fréquence relative
1	de	2192	36,89	31	mais	340	5,72
2	je	1509	25,39	32	plus	322	5,42
3	à	1212	20,40	33	au	310	5,22
4	la	1149	19,33	34	nous	301	5,07
5	que	1117	18,80	35	mon	298	5,01
6	le	967	16,27	36	ma	297	5,00
7	il	907	15,26	37	vous	287	4,83
8	et	885	14,89	38	cette	278	4,68
9	l'	799	13,45	39	moi	275	4,63
10	ne	790	13,29	40	avait	273	4,59
11	les	769	12,94	41	du	258	4,34
12	d'	759	12,77	42	s'	248	4,17
13	un	717	12,07	43	se	241	4,06
14	pas	690	11,61	44	lui	239	4,02
15	ce	610	10,26	45	a	232	3,90
16	me	595	10,01	46	c'	224	3,77
17	qu'	558	9,39	47	comme	222	3,74
18	tu	554	9,32	48	sur	219	3,69
19	qui	548	9,22	49	sa	219	3,69
20	j'	545	9,17	50	son	212	3,57
21	en	536	9,02	51	si	212	3,57
22	elle	526	8,85	52	ai	206	3,47
23	n'	510	8,58	53	ils	205	3,45
24	une	503	8,46	54	tout	200	3,37
25	des	443	7,45	55	avais	200	3,37
26	m'	423	7,12	56	avec	186	3,13
27	dans	401	6,75	57	ces	185	3,11
28	était	366	6,16	57	même	181	3,05
29	pour	361	6,07	59	rien	174	2,93
30	est	353	5,94	60	être	163	2,74

Appendice 10.1.7

L'Immoraliste

Rang	Type	Total	Fréquence relative	Rang	Type	Total	Fréquence relative
1	de	1363	33,03	31	qui	237	5,74
2	je	1305	31,63	32	mon	225	5,45
3	et	832	20,16	33	se	209	5,07
4	la	829	20,09	34	dans	206	4,99
5	à	783	18,98	35	ma	204	4,94
6	que	777	18,83	36	moi	203	4,92
7	le	722	17,50	37	du	202	4,90
8	l'	670	16,24	38	s'	199	4,82
9	il	513	12,43	39	Marceline	189	4,58
10	me	494	11,97	40	tout	181	4,39
11	un	486	11,78	41	c'	178	4,31
12	ne	482	11,68	42	comme	177	4,29
13	d'	479	11,61	43	vous	173	4,19
14	les	443	10,74	44	au	173	4,19
15	en	435	10,54	45	sur	167	4,05
16	j'	425	10,30	46	on	167	4,05
17	qu'	397	9,62	47	lui	142	3,44
18	plus	397	9,62	48	y	138	3,34
19	m'	368	8,92	49	si	136	3,30
20	pas	362	8,77	50	bien	134	3,25
21	elle	330	8,00	51	par	131	3,17
22	ce	330	8,00	52	avait	129	3,13
23	une	295	7,15	53	même	125	3,03
24	est	295	7,15	54	son	124	3,01
25	n'	287	6,96	55	avais	123	2,98
26	nous	281	6,81	56	sans	122	2,96
27	pour	260	6,30	57	encore	121	2,93
28	était	246	5,96	58	rien	114	2,76
29	des	245	5,94	59	peu	112	2,71
30	mais	242	5,87	60	avec	112	2,71

Appendice 10.1.8

La Porte étroite

Rang	Type	Total	Fréquence relative	Rang	Type	Total	Fréquence relative
1	de	1382	34,59	31	dans	247	6,18
2	je	1249	31,26	32	une	228	5,71
3	que	946	23,68	33	pour	223	5,58
4	à	796	19,92	34	moi	223	5,58
5	la	723	18,10	35	si	211	5,28
6	et	709	17,75	36	s'	197	4,93
7	le	579	14,49	37	du	193	4,83
8	elle	551	13,79	38	sans	191	4,78
9	ne	547	13,69	39	Alissa	190	4,76
10	l'	541	13,54	40	tout	179	4,48
11	d'	501	12,54	41	lui	176	4,41
12	pas	433	10,84	42	ai	173	4,33
13	qu'	409	10,24	43	au	168	4,20
14	me	401	10,04	44	était	156	3,90
15	ce	398	9,96	45	se	152	3,80
16	un	394	9,86	46	son	149	3,73
17	j'	390	9,76	47	c'	147	3,68
18	en	379	9,49	48	par	142	3,55
19	m'	374	9,36	49	cette	135	3,38
20	mon	371	9,29	50	a	134	3,35
21	il	354	8,86	51	te	131	3,28
22	n'	335	8,38	52	sa	128	3,20
23	est	335	8,38	53	comme	128	3,20
24	nous	332	8,31	54	t'	125	3,13
25	plus	310	7,76	55	bien	124	3,10
26	les	288	7,21	56	avait	121	3,03
27	tu	287	7,18	57	sur	118	2,95
28	ma	266	6,66	58	même	114	2,85
29	qui	256	6,41	59	y	110	2,75
30	mais	252	6,31	60	avec	110	2,75

Appendice 10.1.9

Le Corpus d'Engwall

Rang	Type	Total	Fréquence relative	Rang	Type	Total	Fréquence relative
1	de	18215	36,43	31	a	2740	5,48
2	la	11677	23,35	32	c'	2728	5,46
3	le	10139	20,28	33	plus	2613	5,23
4	et	9916	19,83	34	sur	2584	5,17
5	il	9317	18,63	35	me	2551	5,10
6	à	9299	18,60	36	était	2432	4,86
7	les	8086	16,17	37	au	2380	4,76
8	je	7579	15,16	38	lui	2359	4,72
9	l'	7483	14,97	39	mais	2353	4,71
10	un	6965	13,93	40	son	2282	4,56
11	que	6104	12,21	41	nous	2241	4,48
12	d'	5888	11,78	42	avait	2201	4,40
13	en	5372	10,74	43	ils	2110	4,22
14	des	5114	10,23	44	tout	2065	4,13
15	pas	5027	10,05	45	avec	2018	4,04
16	une	5005	10,01	46	on	2007	4,01
17	elle	4798	9,60	47	comme	1957	3,91
18	est	4632	9,26	48	tu	1867	3,73
19	qu'	4436	8,87	49	sa	1856	3,71
20	ne	4312	8,62	50	ai	1777	3,56
21	dans	4012	8,02	51	m'	1657	3,31
22	qui	4011	8,02	52	si	1546	3,09
23	n'	3335	6,67	53	y	1534	3,07
24	se	3267	6,53	54	ses	1480	2,96
25	ce	3236	6,47	55	bien	1459	2,92
26	pour	3014	6,03	56	par	1459	2,92
27	s'	2941	5,88	57	moi	1448	2,90
28	j'	2934	5,87	58	mon	1359	2,72
29	vous	2913	5,83	59	même	1345	2,69
30	du	2781	5,56	60	dit	1282	2,56

Les abréviations suivantes sont utilisées dans les tables qui suivent:

<i>Clair</i>	<i>Clair de femme;</i>
<i>Tick</i>	<i>Au-delà de cette limite votre ticket n'est plus valable;</i>
<i>Câl</i>	<i>Gros-Câlin;</i>
<i>Vie</i>	<i>La Vie devant soi;</i>
<i>Étran</i>	<i>L'Étranger;</i>
<i>Vip</i>	<i>Le Noeud de vipères;</i>
<i>Imm</i>	<i>L'Immoraliste;</i>
<i>Porte</i>	<i>La Porte étroite;</i>
<i>Eng</i>	<i>Engwall.</i>

Appendice 10.2

Les Types les plus fréquents dans les huit romans et dans le corpus d'Engwall

Rang	Type	<i>Clair</i>	<i>Tick</i>	<i>Cal</i>	<i>Vie</i>	<i>Étran</i>	<i>Vip</i>	<i>Imm</i>	<i>Porte</i>	<i>Eng</i>	μ	σ	-2 σ	+2 σ
1	de*	33,86	39,54	35,65	23,32	30,44	36,89	33,03	34,54	36,43	33,74	4,68	24,39	43,1
2	je	26,54	25,50	26,19	19,14	21,34	25,39	31,63	31,26	15,16	24,68	5,37	13,94	35,43
3	et*	21,80	25,19	24,29	30,56	27,21	14,89	20,16	17,75	19,83	22,41	4,89	12,63	32,18
4	la	16,86	20,74	17,16	15,33	18,74	19,33	20,09	18,10	23,35	18,86	2,38	14,09	23,62
5	à*	15,14	17,35	21,05	15,64	16,52	20,40	18,98	19,92	18,60	18,18	2,13	13,93	22,43
6	il	19,49	17,42	16,88	19,48	27,01	15,26	12,43	8,86	18,63	17,27	5,06	7,15	27,40
7	que*	14,04	14,90	15,32	16,94	18,46	18,80	18,83	23,63	12,21	17,01	3,38	10,25	23,78
8	le	14,87	13,81	15,94	15,19	20,54	16,27	17,50	14,49	20,28	16,54	2,44	11,66	21,42
9	l'	13,27	14,00	14,68	10,17	15,48	13,45	16,24	13,54	14,97	13,98	1,74	10,49	17,46
10	un*	16,54	15,38	13,22	13,33	14,53	12,07	11,78	9,86	13,93	13,40	2,01	9,39	17,42
11	pas*	14,31	11,17	13,01	16,27	9,91	11,61	8,77	10,84	10,05	11,77	2,37	7,03	16,51
12	les	10,78	10,96	13,01	12,66	10,20	12,94	10,74	7,21	16,17	11,63	2,47	6,69	16,57
13	d'*	11,58	12,93	11,13	6,45	9,94	12,77	11,61	12,54	11,78	11,19	2,00	7,19	15,2
14	est	15,05	10,57	15,53	14,72	9,76	5,94	7,15	8,38	9,26	10,71	3,57	3,57	17,85
15	ne*	12,41	10,34	8,92	9,68	7,02	13,29	11,68	13,69	8,62	10,63	2,28	6,06	15,19
16	j'	7,58	8,36	10,83	13,70	18,25	9,17	10,30	9,76	5,87	10,42	3,66	3,10	17,75
17	qu'*	7,44	7,49	8,30	13,09	11,12	9,39	9,62	10,24	8,87	9,51	1,81	5,89	13,13
18	en*	6,87	9,43	10,63	9,28	8,66	9,02	10,54	9,49	10,74	9,41	1,21	6,98	11,83
19	elle	8,89	5,54	4,71	17,77	6,53	8,85	8,00	13,79	9,60	9,30	4,14	1,02	17,57
20	une*	11,58	11,98	9,80	7,17	7,94	8,46	7,15	5,71	10,01	8,87	2,12	4,62	13,11
21	me	6,01	8,69	9,57	6,25	8,26	10,01	11,97	10,04	5,10	8,43	2,26	3,92	12,94
22	a	8,55	6,06	7,43	13,84	22,90	3,90	1,82	3,35	5,48	8,15	6,55	-4,95	21,25
23	qui*	8,32	7,08	8,47	11,64	5,81	9,22	5,74	6,41	8,02	7,86	1,88	4,10	11,61
24	ce*	8,56	7,55	6,68	6,31	6,53	10,26	8,00	9,96	6,47	7,81	1,51	4,79	10,83
25	n'*	8,89	6,68	7,42	8,30	8,03	8,58	6,96	8,36	6,67	7,77	0,85	6,07	9,47
26	des*	7,55	8,65	9,54	10,54	6,44	7,45	5,94	2,70	10,23	7,67	2,46	2,75	12,60
27	c'*	9,51	7,30	11,15	11,92	6,73	3,77	4,31	3,68	5,46	7,09	3,14	0,81	13,37

28 dans*	6,61	7,55	9,93	6,76	6,76	6,75	4,99	6,18	8,02	7,06	1,37	4,33	9,80
29 pour*	6,19	6,56	8,43	10,12	5,37	6,07	6,30	5,58	6,03	6,74	1,54	3,66	9,82
30 était	5,72	5,96	5,31	11,72	10,57	6,16	5,96	3,90	4,86	6,68	2,64	1,41	11,96
31 m'	4,30	5,13	4,56	4,92	11,96	7,12	8,92	9,36	3,31	6,62	2,90	0,82	12,42
32 vous	14,01	7,22	7,83	6,04	2,31	4,83	4,29	2,65	5,83	6,11	3,50	-0,89	13,11
33 plus*	5,18	5,97	4,26	6,94	4,33	5,42	9,62	7,76	5,23	6,08	1,75	2,58	9,58
34 mais*	5,72	4,99	6,03	7,27	7,48	5,72	5,87	6,31	4,71	6,01	0,92	4,17	7,85
35 ai	4,59	3,75	6,53	9,59	15,94	3,47	1,99	4,33	3,56	5,97	4,33	-2,70	14,64
36 nous	7,85	4,01	2,68	2,10	5,20	5,07	6,81	8,31	4,48	5,17	2,16	0,86	9,48
37 avait	4,38	4,96	3,29	10,29	8,12	4,59	3,13	3,03	4,40	5,13	2,47	0,19	10,07
38 du*	5,33	5,59	5,59	4,16	5,29	4,34	4,90	4,83	5,56	5,07	0,54	3,98	6,15
39 s'	4,30	3,97	4,26	5,61	6,59	4,17	4,82	4,93	5,88	4,95	0,90	3,15	6,75
40 on	7,35	4,44	7,74	8,17	4,71	2,07	4,05	1,53	4,01	4,90	2,39	0,11	9,69
41 se	4,98	5,23	5,26	5,59	3,03	4,06	5,07	3,80	6,53	4,84	1,05	2,74	6,94
42 mon	2,84	5,04	4,32	2,26	5,14	5,01	5,45	9,29	2,72	4,67	2,11	0,46	8,89
43 au*	4,33	5,11	5,46	3,13	4,59	5,22	4,19	4,20	4,76	4,55	0,70	3,15	5,96
44 tout*	4,65	3,64	5,44	5,36	5,05	3,37	4,39	4,48	4,13	4,50	0,72	3,07	5,94
45 moi	4,33	4,35	4,54	3,76	4,39	4,63	4,92	5,58	2,90	4,38	0,74	2,89	5,86
46 y*	7,02	5,18	5,71	6,07	3,23	2,73	3,34	2,75	3,07	4,34	1,65	1,05	7,64
47 avec*	3,20	4,68	7,47	5,27	4,77	3,13	2,71	2,75	4,04	4,22	1,54	1,15	7,30
48 lui	2,76	2,71	3,12	5,04	6,96	4,02	3,44	4,41	4,72	4,13	1,35	1,42	6,84
49 tu	4,18	5,80	1,26	2,06	0,98	9,32	1,36	7,18	3,73	3,99	2,94	-1,89	9,86
50 sur*	3,32	3,99	3,94	2,77	4,85	3,69	4,05	2,95	5,17	3,86	0,80	2,26	5,45
51 comme*	3,32	2,92	3,81	6,27	2,48	3,74	4,29	3,20	3,91	3,77	1,09	1,60	5,95
52 son	2,34	3,14	2,38	3,29	3,99	3,57	3,01	3,73	4,56	3,33	0,72	1,89	4,78
53 même*	2,58	3,63	4,21	4,72	2,83	3,05	3,03	2,85	2,69	3,29	0,74	1,80	4,77
54 si*	3,32	2,99	2,14	2,61	3,26	3,57	3,30	5,28	3,09	3,28	0,86	1,56	5,01
55 ça*	3,91	4,28	6,36	8,43	0,69	0,76	0,19	0,10	2,17	2,99	2,97	-2,95	8,93
56 bien*	2,67	2,81	3,38	3,56	2,43	2,29	3,25	3,10	2,92	2,93	0,43	2,08	3,79
57 dit	2,37	1,90	2,78	3,04	7,02	1,09	1,79	2,15	2,56	2,74	1,71	-2,37	6,15
58 ils	1,42	1,48	2,61	4,84	2,51	3,45	1,72	0,55	4,22	2,53	1,41	-0,29	5,36
59 fait	1,81	2,28	3,00	3,23	2,77	1,30	1,75	1,70	1,85	2,19	0,67	0,85	3,52
60 suis	2,78	1,59	3,53	3,61	3,06	1,35	0,80	1,33	1,55	2,18	1,06	0,05	4,31

Appendice 10.3

Les Neuf Textes: limites de confiance à 95%

Type	Clair	Tick	Câl	Vie	Étran	Vip	Imm	Porte	Eng	μ	σ	+95%	-95%
de	33,86	39,50	35,65	23,30	30,44	36,89	33,03	34,54	36,43	33,74	4,68	37,40	30,15
je	26,54	25,50	26,19	19,10	21,34	25,39	31,60	31,30	15,16	24,68	5,37	28,80	20,55
et	21,80	25,19	24,29	30,6	27,20	14,90	20,20	17,75	19,83	22,41	4,89	26,20	18,65
la	16,90	20,74	17,16	15,33	18,74	19,33	20,09	18,10	23,35	18,86	2,38	20,70	17,03
à	15,10	17,35	21,10	15,60	16,50	20,40	18,98	19,90	18,60	18,18	2,13	19,80	16,54
il	19,49	17,42	16,88	19,48	27,00	15,26	12,40	8,86	18,63	17,27	5,06	21,20	13,38
que	14,00	14,90	15,32	16,94	18,46	18,80	18,83	23,60	12,21	17,01	3,38	19,60	14,41
le	14,87	13,80	15,94	15,19	20,50	16,27	17,50	14,50	20,28	16,54	2,44	18,40	14,67
l'	13,27	14,00	14,68	10,17	15,50	13,45	16,20	13,54	14,97	13,98	1,74	15,30	12,64
un	16,50	15,40	13,22	13,33	14,53	12,07	11,80	9,86	13,93	13,40	2,01	14,90	11,86
pas	14,30	11,17	13,01	16,30	9,91	11,61	8,77	10,84	10,05	11,77	2,37	13,60	9,95
les	10,78	10,96	13,01	12,66	10,20	12,94	10,74	7,21	16,17	11,63	2,47	13,50	9,73
d'	11,58	12,90	11,13	6,45	9,94	12,80	11,61	12,54	11,78	11,19	2,00	12,70	9,65
est	15,10	10,57	15,50	14,70	9,76	5,94	7,15	8,38	9,26	10,71	3,57	13,50	7,96
ne	12,40	10,34	8,92	9,68	7,02	13,30	11,68	13,70	8,62	10,63	2,28	12,40	8,87
j'	7,58	8,36	10,83	13,70	18,30	9,17	10,30	9,76	5,87	10,42	3,66	13,20	7,61
qu'	7,44	7,49	8,30	13,10	11,10	9,39	9,62	10,24	8,87	9,51	1,81	11,00	8,12
en	6,87	9,43	10,60	9,28	8,66	9,02	10,50	9,49	10,74	9,41	1,21	10,30	8,48
elle	8,89	5,54	4,71	17,80	6,53	8,85	8,00	13,80	9,60	9,30	4,14	12,50	6,12
une	11,60	12,00	9,80	7,17	7,94	8,46	7,15	5,71	10,01	8,87	2,12	10,50	7,24
me	6,01	8,69	9,57	6,25	8,26	10,01	12,00	10,04	5,10	8,43	2,26	10,20	6,70
a	8,55	6,06	7,43	13,80	22,90	3,90	1,82	3,35	5,48	8,15	6,55	13,20	3,11
qui	8,32	7,08	8,47	11,60	5,81	9,22	5,74	6,41	8,02	7,86	1,88	9,30	6,41
ce	8,56	7,55	6,68	6,31	6,53	10,30	8,00	9,96	6,47	7,81	1,51	8,97	6,65
n'	8,89	6,68	7,42	8,30	8,03	8,58	6,96	8,36	6,67	7,77	0,85	8,42	7,11
des	7,55	8,65	9,54	10,50	6,44	7,45	5,94	2,70	10,23	7,67	2,46	9,56	5,78
c'	9,51	7,30	11,2	11,90	6,73	3,77	4,31	3,68	5,46	7,09	3,14	9,51	4,68
dans	6,61	7,55	9,93	6,76	6,76	6,75	4,99	6,18	8,02	7,06	1,37	8,11	6,01
pour	6,19	6,56	8,43	10,10	5,37	6,07	6,30	5,58	6,03	6,74	1,54	7,92	5,55
était	5,72	5,96	5,31	11,70	10,60	6,16	5,96	3,90	4,86	6,68	2,64	8,71	4,66
m'	4,30	5,13	4,56	4,92	12,00	7,12	8,92	9,36	3,31	6,62	2,90	8,85	4,39
vous	14,00	7,22	7,83	6,04	2,31	4,83	4,29	2,65	5,83	6,11	3,50	8,80	3,42
plus	5,18	5,97	4,26	6,94	4,33	5,42	9,62	7,76	5,23	6,08	1,75	7,42	4,74
mais	5,72	4,99	6,03	7,27	7,48	5,72	5,87	6,31	4,71	6,01	0,92	6,72	5,30
ai	4,59	3,75	6,53	9,59	15,90	3,47	1,99	4,33	3,56	5,97	4,33	9,30	2,64
nous	7,85	4,01	2,68	2,10	5,20	5,07	6,81	8,31	4,48	5,17	2,16	6,83	3,51
avait	4,38	4,96	3,29	10,30	8,12	4,59	3,13	3,03	4,40	5,13	2,47	7,03	3,23

du	5,33	5,59	5,59	4,16	5,29	4,34	4,90	4,83	5,56	5,07	0,54	5,48	4,65
s'	4,30	3,97	4,26	5,61	6,59	4,17	4,82	4,93	5,88	4,95	0,90	5,64	4,26
on	7,35	4,44	7,74	8,17	4,71	2,07	4,05	1,53	4,01	4,90	2,40	6,74	3,06
se	4,98	5,23	5,26	5,59	3,03	4,06	5,07	3,80	6,53	4,84	1,05	5,64	4,03
mon	2,84	5,04	4,32	2,26	5,14	5,01	5,45	9,29	2,72	4,67	2,11	6,29	3,06
au	4,33	5,11	5,46	3,13	4,59	5,22	4,19	4,20	4,76	4,55	0,70	5,09	4,01
tout	4,65	3,64	5,44	5,36	5,05	3,37	4,39	4,48	4,13	4,50	0,72	5,05	3,95
moi	4,33	4,35	4,54	3,76	4,39	4,63	4,92	5,58	2,90	4,38	0,74	4,95	3,81
y	7,02	5,18	5,71	6,07	3,23	2,73	3,34	2,75	3,07	4,34	1,65	5,61	3,08
avec	3,20	4,68	7,47	5,27	4,77	3,13	2,71	2,75	4,04	4,22	1,54	5,41	3,04
lui	2,76	2,71	3,12	5,04	6,96	4,02	3,44	4,41	4,72	4,13	1,35	5,17	3,09
tu	4,18	5,80	1,26	2,06	0,98	9,32	1,36	7,18	3,73	3,99	2,94	6,24	1,73
sur	3,32	3,99	3,94	2,77	4,85	3,69	4,05	2,95	5,17	3,86	0,80	4,47	3,25
comme	3,32	2,92	3,81	6,27	2,48	3,74	4,29	3,20	3,91	3,77	1,09	4,61	2,94
son	2,34	3,14	2,38	3,29	3,99	3,57	3,01	3,73	4,56	3,33	0,72	3,89	2,78
même	2,58	3,63	4,21	4,72	2,83	3,05	3,03	2,85	2,69	3,29	0,74	3,86	2,72
si	3,32	2,99	2,14	2,61	3,26	3,57	3,30	5,28	3,09	3,28	0,86	3,95	2,62
ça	3,91	4,28	6,36	8,43	0,69	0,76	0,19	0,10	2,17	2,99	2,97	5,27	0,70
bien	2,67	2,81	3,38	3,56	2,43	2,29	3,25	3,10	2,92	2,93	0,43	3,26	2,60
dit	2,37	1,90	2,78	3,04	7,02	1,09	1,79	2,15	2,56	2,74	1,71	4,06	1,43
ils	1,42	1,48	2,61	4,84	2,51	3,45	1,72	0,55	4,22	2,53	1,41	3,62	1,45
fait	1,81	2,28	3,00	3,23	2,77	1,30	1,75	1,70	1,85	2,19	0,67	2,70	1,67
suis	2,78	1,59	3,53	3,61	3,06	1,35	0,80	1,33	1,55	2,18	1,06	3,00	1,36

Appendice 10.4

Les Neuf Textes: limites de confiance à 99%

Type	Clair	Tick	Câl	Vie	Étran	Vip	Imm	Porte	Eng	μ	σ	+99%	-99%
de	33,86	39,50	35,65	23,30	30,44	36,89	33,03	34,54	36,43	33,74	4,68	38,97	28,51
je	26,54	25,50	26,19	19,14	21,34	25,39	31,60	31,26	15,16	24,68	5,37	30,69	18,67
et	21,80	25,19	24,29	30,60	27,21	14,90	20,16	17,75	19,83	22,41	4,89	27,88	16,94
la	16,86	20,74	17,16	15,30	18,74	19,33	20,09	18,10	23,35	18,86	2,38	21,52	16,19
à	15,10	17,35	21,10	15,60	16,52	20,40	18,98	19,92	18,60	18,18	2,13	20,56	15,80
il	19,49	17,42	16,88	19,48	27,00	15,26	12,43	8,86	18,63	17,27	5,06	22,93	11,61
que	14,04	14,90	15,32	16,94	18,46	18,80	18,83	23,60	12,21	17,01	3,38	20,79	13,23
le	14,87	13,81	15,94	15,19	20,50	16,27	17,50	14,49	20,28	16,54	2,44	19,27	13,81
l'	13,27	14,00	14,68	10,20	15,48	13,45	16,20	13,54	14,97	13,98	1,74	15,93	12,03
un	16,50	15,38	13,22	13,33	14,53	12,07	11,78	9,86	13,93	13,40	2,01	15,65	11,15
pas	14,31	11,17	13,01	16,30	9,91	11,61	8,77	10,84	10,05	11,77	2,37	14,42	9,12
les	10,78	10,96	13,01	12,66	10,20	12,94	10,74	7,21	16,17	11,63	2,47	14,39	8,87
d'	11,58	12,93	11,13	6,45	9,94	12,77	11,61	12,54	11,78	11,19	2,00	13,43	8,95
est	15,1	10,57	15,50	14,70	9,76	5,94	7,15	8,38	9,26	10,71	3,57	14,70	6,72
ne	12,41	10,34	8,92	9,68	7,02	13,30	11,68	13,70	8,62	10,63	2,28	13,18	8,08
j'	7,58	8,36	10,83	13,70	18,30	9,17	10,30	9,76	5,87	10,42	3,66	14,51	6,33
qu'	7,44	7,49	8,30	13,10	11,12	9,39	9,62	10,24	8,87	9,51	1,81	11,53	7,49
en	6,87	9,43	10,63	9,28	8,66	9,02	10,54	9,49	10,74	9,41	1,21	10,76	8,06
elle	8,89	5,54	4,71	17,80	6,53	8,85	8,00	13,79	9,60	9,30	4,14	13,93	4,67
une	11,6	12,00	9,80	7,17	7,94	8,46	7,15	5,71	10,01	8,87	2,12	11,24	6,50
me	6,01	8,69	9,57	6,25	8,26	10,01	12,00	10,04	5,10	8,43	2,26	10,96	5,90
a	8,55	6,06	7,43	13,84	22,90	3,90	1,82	3,35	5,48	8,15	6,55	15,48	0,82
qui	8,32	7,08	8,47	11,60	5,81	9,22	5,74	6,41	8,02	7,86	1,88	9,96	5,76
ce	8,56	7,55	6,68	6,31	6,53	10,30	8,00	9,96	6,47	7,81	1,51	9,50	6,12
n'	8,89	6,68	7,42	8,30	8,03	8,58	6,96	8,36	6,67	7,77	0,85	8,72	6,82
des	7,55	8,65	9,54	10,50	6,44	7,45	5,94	2,70	10,23	7,67	2,46	10,42	4,92
c'	9,51	7,30	11,20	11,90	6,73	3,77	4,31	3,68	5,46	7,09	3,14	10,60	3,58
dans	6,61	7,55	9,93	6,76	6,76	6,75	4,99	6,18	8,02	7,06	1,37	8,59	5,53
pour	6,19	6,56	8,43	10,10	5,37	6,07	6,30	5,58	6,03	6,74	1,54	8,46	5,02
était	5,72	5,96	5,31	11,70	10,60	6,16	5,96	3,90	4,86	6,68	2,64	9,63	3,73
m'	4,30	5,13	4,56	4,92	12,00	7,12	8,92	9,36	3,31	6,62	2,90	9,86	3,38
vous	14,01	7,22	7,83	6,04	2,31	4,83	4,29	2,65	5,83	6,11	3,50	10,02	2,20
plus	5,18	5,97	4,26	6,94	4,33	5,42	9,62	7,76	5,23	6,08	1,75	8,04	4,12
mais	5,72	4,99	6,03	7,27	7,48	5,72	5,87	6,31	4,71	6,01	0,92	7,04	4,98
ai	4,59	3,75	6,53	9,59	15,90	3,47	1,99	4,33	3,56	5,97	4,33	10,81	1,13
nous	7,85	4,01	2,68	2,10	5,20	5,07	6,81	8,31	4,48	5,17	2,16	7,59	2,75
avait	4,38	4,96	3,29	10,30	8,12	4,59	3,13	3,03	4,40	5,13	2,47	7,89	2,37

du	5,33	5,59	5,59	4,16	5,29	4,34	4,90	4,83	5,56	5,07	0,54	5,67	4,47
s'	4,30	3,97	4,26	5,61	6,59	4,17	4,82	4,93	5,88	4,95	0,90	5,96	3,94
on	7,35	4,44	7,74	8,17	4,71	2,07	4,05	1,53	4,01	4,90	2,40	7,58	2,22
se	4,98	5,23	5,26	5,59	3,03	4,06	5,07	3,80	6,53	4,84	1,05	6,01	3,67
mon	2,84	5,04	4,32	2,26	5,14	5,01	5,45	9,29	2,72	4,67	2,11	7,03	2,31
au	4,33	5,11	5,46	3,13	4,59	5,22	4,19	4,20	4,76	4,55	0,70	5,33	3,77
tout	4,65	3,64	5,44	5,36	5,05	3,37	4,39	4,48	4,13	4,50	0,72	5,31	3,69
moi	4,33	4,35	4,54	3,76	4,39	4,63	4,92	5,58	2,90	4,38	0,74	5,21	3,55
y	7,02	5,18	5,71	6,07	3,23	2,73	3,34	2,75	3,07	4,34	1,65	6,19	2,49
avec	3,20	4,68	7,47	5,27	4,77	3,13	2,71	2,75	4,04	4,22	1,54	5,94	2,50
lui	2,76	2,71	3,12	5,04	6,96	4,02	3,44	4,41	4,72	4,13	1,35	5,64	2,62
tu	4,18	5,80	1,26	2,06	0,98	9,32	1,36	7,18	3,73	3,99	2,94	7,28	0,70
sur	3,32	3,99	3,94	2,77	4,85	3,69	4,05	2,95	5,17	3,86	0,80	4,75	2,97
comme	3,32	2,92	3,81	6,27	2,48	3,74	4,29	3,20	3,91	3,77	1,09	4,99	2,55
son	2,34	3,14	2,38	3,29	3,99	3,57	3,01	3,73	4,56	3,33	0,72	4,14	2,52
même	2,58	3,63	4,21	4,72	2,83	3,05	3,03	2,85	2,69	3,29	0,74	4,12	2,46
si	3,32	2,99	2,14	2,61	3,26	3,57	3,30	5,28	3,09	3,28	0,86	4,24	2,32
ça	3,91	4,28	6,36	8,43	0,69	0,76	0,19	0,10	2,17	2,99	2,97	6,31	-0,33
bien	2,67	2,81	3,38	3,56	2,43	2,29	3,25	3,10	2,92	2,93	0,43	3,41	2,45
dit	2,37	1,90	2,78	3,04	7,02	1,09	1,79	2,15	2,56	2,74	1,71	4,65	0,84
ils	1,42	1,48	2,61	4,84	2,51	3,45	1,72	0,55	4,22	2,53	1,41	4,11	0,95
fait	1,81	2,28	3,00	3,23	2,77	1,30	1,75	1,70	1,85	2,19	0,67	2,94	1,44
suis	2,78	1,59	3,53	3,61	3,06	1,35	0,80	1,33	1,55	2,18	1,06	3,37	0,99

Appendices au chapitre 11

Les abréviations suivantes sont utilisées dans les tables:

<i>Clair</i>	<i>Clair de femme;</i>
<i>Tick</i>	<i>Au-delà de cette limite votre ticket n'est plus valable;</i>
<i>Câl</i>	<i>Gros-Câlin;</i>
<i>Vie</i>	<i>La Vie devant soi;</i>
<i>Étang</i>	<i>L'Étranger;</i>
<i>Vip</i>	<i>Le Noeud de vipères;</i>
<i>Imm</i>	<i>L'Immoraliste;</i>
<i>Porte</i>	<i>La Porte étroite;</i>
<i>Eng</i>	<i>Engwall.</i>

Appendice 11.1

Les Neuf Textes: fréquences observées

Texte Vocables	Clair 33.755	Tick 57.915	Câl 53.264	Vie 65.095	Étran 34.625	Vip 59.426	Imm 41.260	Porte 39.953	Eng 500.000
de	1.143	2.290	1.899	1.518	1.054	2.192	1.363	1.382	18.215
je	896	1.477	1.395	1.246	739	1.509	1.305	1.249	7.579
et	736	1.459	1.294	1.989	942	885	832	709	9.916
la	569	1.201	914	998	649	1.149	829	723	11.677
à	511	1.007	1.121	1.018	572	1.212	783	796	9.299
il	658	1.009	899	1.268	935	907	513	354	9.317
que	474	863	816	1.103	639	1.117	777	944	6.104
le	502	800	849	989	711	967	722	579	10.139
l'	448	811	782	662	536	799	670	541	7.483
un	559	891	704	868	503	717	486	394	6.965
pas	483	647	693	1.059	343	690	362	433	5.027
les	364	635	693	824	353	769	443	288	8.086
d'	391	749	593	420	344	759	479	501	5.888
est	508	612	827	958	338	353	295	335	4.632
ne	419	599	475	630	243	790	482	547	4.312
j'	256	484	577	892	632	545	425	390	2.934
qu'	251	434	442	852	385	558	397	409	4.436
en	232	546	566	604	300	536	435	379	5.372
elle	300	321	251	1.157	226	526	330	551	4.798
une	391	694	522	467	275	503	295	228	5.005
me	203	503	510	407	286	595	494	401	2.551
a	288	351	396	901	793	232	75	134	2.740
qui	281	410	451	758	201	548	237	256	4.011
ce	289	437	356	411	226	610	330	398	3.236
n'	300	387	395	540	278	510	287	334	3.335
des	255	501	508	686	223	443	245	108	5.114
c'	321	423	594	776	233	224	178	147	2.728
dans	223	437	529	442	234	401	206	247	4.012
pour	209	380	449	659	186	361	260	223	3.014
était	193	345	283	763	366	366	246	156	2.432
m'	145	297	243	320	414	423	368	374	1.657
vous	473	418	417	393	80	287	177	106	2.913
plus	175	346	227	452	150	322	397	310	2.613
mais	193	289	321	473	259	340	242	252	2.353
ai	155	217	348	624	552	206	82	173	1.778
nous	265	232	143	137	180	301	281	332	2.241
avait	148	287	175	670	281	273	129	121	2.201
du	180	324	298	271	183	258	202	193	2.781
s'	145	230	227	365	228	248	199	197	2.941
on	248	257	412	532	163	123	167	61	2.007
se	168	303	280	364	105	241	209	152	3.267
mon	96	292	230	147	178	298	225	371	1.359

au	146	296	291	204	159	310	173	168	2.380
tout	157	211	290	349	175	200	181	179	2.065
moi	146	252	242	245	152	275	203	223	1.448
y	237	300	304	395	112	162	138	110	1.534
avec	108	271	398	343	165	186	112	110	2.018
lui	93	157	166	328	241	239	142	176	2.359
tu	141	336	67	134	34	554	56	287	1.867
sur	112	231	210	180	168	219	167	118	2.584
comme	112	169	203	408	86	222	177	128	1.957
son	79	182	127	214	138	212	124	149	2.282
même	87	210	224	307	98	181	125	114	1.345
si	112	173	114	170	113	212	136	211	1.545
ça	132	248	339	549	24	45	8	4	1.085
bien	90	163	180	232	84	136	134	124	1.460
dit	80	110	148	198	243	65	74	86	1.280
ils	48	86	139	315	87	205	71	22	2.110
fait	61	132	160	210	96	77	72	68	925
suis	94	92	188	235	106	80	33	53	775
	17.079	20.256	27.894	35.629	18.799	28.674	19.585	19.108	239.487

Appendice 11.2

Les Huit Romans: Valeurs hypothétiques

Texte	Clair	Tick	Câl	Vie	Étran	Vip	Imm	Porte
Vocables	33.755	57.915	53.264	65.095	34.625	59.426	41.260	39.953
de	1.229,69	2.109,84	1.940,41	2.371,41	1.261,32	2.164,89	1.503,10	1.455,49
je	511,66	877,88	807,38	986,71	524,82	900,78	625,42	605,61
et	669,43	1.148,57	1.056,33	1.290,96	686,64	1.178,54	818,27	792,35
la	788,31	1.352,55	1.243,93	1.520,23	808,59	1.387,83	963,59	933,06
à	627,78	1.077,10	990,60	1.210,64	643,92	1.105,20	767,35	743,05
il	628,99	1.079,19	992,52	1.212,98	645,16	1.107,34	768,84	744,48
que	412,08	707,03	650,25	794,68	422,68	725,47	503,70	487,75
le	684,48	1.174,40	1.080,09	1.320,00	702,09	1.205,04	836,67	810,17
l'	505,18	866,76	797,15	974,21	518,17	889,37	617,50	597,94
un	470,21	806,76	741,97	906,77	482,30	827,80	574,75	556,55
pas	339,37	582,28	535,52	654,47	348,10	597,47	414,83	401,69
les	545,89	936,60	861,39	1052,72	559,92	961,04	667,26	646,12
d'	397,50	682,01	627,24	766,56	407,72	699,80	485,88	470,49
est	312,71	536,52	493,44	603,04	320,75	550,52	382,23	370,12
ne	291,10	499,46	459,35	561,38	298,59	512,49	355,83	344,55
j'	198,07	339,85	312,55	381,98	203,17	348,71	242,11	234,44
qu'	299,47	513,82	472,56	577,52	307,18	527,23	366,06	354,46
en	362,66	622,24	572,27	699,38	371,99	638,47	443,30	429,26
elle	323,91	555,75	511,12	624,65	332,24	570,25	395,93	383,39
une	337,89	579,73	533,17	651,60	346,58	594,85	413,01	399,93
me	172,22	295,48	271,75	332,11	176,65	303,19	210,51	203,84

a	184,98	317,37	291,89	356,72	189,73	325,65	226,10	218,94
qui	270,78	464,59	427,28	522,19	277,75	476,72	330,99	320,50
ce	218,46	374,83	344,72	421,29	224,08	384,61	267,03	258,58
n'	225,15	386,29	355,27	434,18	230,94	396,37	275,20	266,49
des	345,25	592,35	544,78	665,79	354,12	607,81	422,01	408,64
c'	184,17	315,98	290,61	355,16	188,90	324,23	225,11	217,98
dans	270,85	464,71	427,39	522,32	277,81	476,83	331,07	320,58
pour	203,48	349,11	321,08	392,39	208,71	358,22	248,72	240,84
était	164,18	281,70	259,08	316,62	168,41	289,05	200,69	194,33
m'	111,86	191,93	176,52	215,72	114,74	196,94	136,74	132,40
vous	196,66	337,41	310,32	379,24	201,71	346,22	240,38	232,77
plus	176,40	302,66	278,36	340,19	180,94	310,56	215,62	208,79
mais	158,85	272,55	250,66	306,34	162,94	279,66	194,17	188,02
ai	120,03	205,95	189,41	231,48	123,12	211,32	146,72	142,07
nous	151,29	259,58	238,73	291,76	155,18	266,35	184,93	179,07
avait	148,59	254,94	234,47	286,55	152,41	261,59	181,63	175,87
du	187,75	322,12	296,25	362,06	192,57	330,53	229,49	222,22
s'	198,55	340,66	313,30	382,89	203,65	349,54	242,69	235,00
on	135,49	232,47	213,80	261,29	138,98	238,54	165,62	160,37
se	220,56	378,42	348,03	425,33	226,23	388,29	269,59	261,05
mon	91,75	157,41	144,77	176,93	94,11	161,52	112,14	108,59
au	160,67	275,68	253,54	309,85	164,81	282,87	196,40	190,18
tout	139,41	239,19	219,98	268,84	142,99	245,43	170,40	165,01
moi	97,75	167,72	154,25	188,52	100,27	172,10	119,49	115,70
y	103,56	177,68	163,41	199,71	106,22	182,32	126,59	122,58
avec	136,24	233,74	214,97	262,72	139,74	239,84	166,53	161,25
lui	159,26	273,24	251,30	307,12	163,35	280,37	194,66	188,50
tu	126,04	216,25	198,89	243,06	129,28	221,90	154,06	149,18
sur	174,45	299,30	275,27	336,41	178,93	307,11	213,23	206,48
comme	132,12	226,68	208,48	254,78	135,51	232,59	161,49	156,38
son	154,06	264,32	243,10	297,09	158,02	271,22	188,31	182,35
même	90,80	155,79	143,28	175,11	93,14	159,86	110,99	107,47
si	104,30	178,96	164,59	201,14	106,99	183,63	127,49	123,45
ça	73,25	125,68	115,58	141,26	75,13	128,95	89,53	86,70
bien	98,56	169,11	155,53	190,08	101,10	173,52	120,48	116,66
dit	86,41	148,26	136,36	166,64	88,63	152,13	105,63	102,28
ils	142,45	244,40	224,77	274,70	146,11	250,78	174,12	168,60
fait	62,45	107,14	98,54	120,43	64,05	109,94	76,33	73,91
suis	52,32	89,77	82,56	100,90	53,67	92,11	63,95	61,93

Appendice 11.3

Les Huit Romans: χ^2 (Engwall fournit la valeur calculée)

Type	Clair	Tick	Câl	Vie	Étran	Vip	Imm	Porte
de	6,11	15,38	0,89	307,12	34,08	0,34	13,06	3,71
je*	288,71	408,88	427,68	68,14	87,41	410,68	738,43	683,53
et	6,62	83,90	53,47	377,43	94,96	73,11	0,23	8,77
la*	61,01	16,94	87,50	179,40	31,50	41,10	18,80	47,29
à	21,72	4,83	17,16	30,65	8,03	10,32	0,32	3,77
il*	1,34	4,56	8,81	2,50	130,21	36,25	85,13	204,81
que	9,30	34,41	42,25	119,62	110,71	211,30	148,29	426,79
le*	48,65	119,36	49,44	83,00	0,11	47,02	15,72	65,96
l'*	6,47	3,59	0,29	100,06	0,61	9,18	4,46	5,42
un	16,39	8,80	1,94	1,66	0,89	14,83	13,70	47,47
pas	60,79	7,19	46,31	250,05	0,07	14,33	6,73	2,44
les*	60,60	97,12	32,92	49,69	76,47	38,37	75,37	198,49
d'	0,11	6,58	1,87	156,68	9,96	5,01	0,10	1,98
est*	121,97	10,62	225,49	208,94	0,93	70,87	19,91	3,33
ne	56,19	19,84	0,53	8,39	10,35	150,27	44,74	118,95
j'*	16,94	61,15	223,74	680,99	905,15	110,49	138,15	103,21
qu'	7,85	12,40	1,98	130,45	19,72	1,80	2,62	8,39
en	47,08	9,34	0,07	13,01	13,93	16,45	0,16	5,88
elle*	1,77	99,16	132,38	453,68	33,97	3,43	10,98	73,28
une	8,35	22,52	0,23	52,30	14,78	14,18	33,72	73,91
me*	5,50	145,74	208,87	16,89	67,70	280,85	381,78	190,70
a*	58,50	3,56	37,14	830,45	1.918,11	26,93	100,98	32,95
qui	0,39	6,42	1,32	106,48	21,21	10,66	26,69	12,98
ce	22,78	10,31	0,37	0,25	0,02	132,09	14,85	75,18
n'	24,89	0,00	4,44	25,79	9,59	32,57	0,51	17,10
des	23,59	14,09	2,48	0,61	48,55	44,69	74,24	221,18
c'	101,66	36,24	316,74	498,67	10,29	30,98	9,86	23,11
dans	8,45	1,65	24,16	12,97	6,91	12,06	47,25	16,89
pour	0,15	2,73	50,97	181,14	2,47	0,02	0,51	1,32
était*	5,06	14,22	2,21	629,31	231,84	20,49	10,23	7,56
m'*	9,82	57,52	25,04	50,40	780,51	259,49	391,14	440,84
vous*	388,32	19,25	36,68	0,50	73,44	10,13	16,71	69,04
plus	0,01	6,20	9,48	36,75	5,29	0,42	152,57	49,06
mais	7,34	0,99	19,74	90,67	56,64	13,02	11,78	21,77
ai*	10,19	0,59	132,79	665,61	1.493,99	0,13	28,55	6,73
nous*	85,46	2,93	38,39	82,09	3,97	4,51	49,91	130,61
avait*	0,00	4,03	15,08	513,13	108,49	0,50	15,25	17,12
du	0,32	0,01	0,01	22,90	0,48	15,91	3,29	3,84
s'*	14,44	35,94	23,77	0,84	2,91	29,50	7,87	6,15
on*	93,42	2,59	183,73	280,47	4,15	55,96	0,01	61,57
se*	12,52	15,03	13,30	8,84	64,96	55,87	13,62	45,56
mon*	0,20	115,07	50,17	5,06	74,79	115,32	113,57	634,10
au	1,34	1,50	5,54	36,16	0,20	2,60	2,79	2,59
tout	2,22	3,32	22,29	23,90	7,16	8,41	0,66	1,19
moi*	23,81	42,35	49,92	16,92	26,69	61,53	58,37	99,50

y	171,94	84,20	120,95	190,96	0,31	2,26	1,03	1,29
avec	5,85	5,94	155,83	24,53	4,57	12,09	17,85	16,29
lui*	27,56	49,45	28,95	1,42	36,91	6,10	14,25	0,83
tu*	1,78	66,31	87,46	48,94	70,22	497,04	62,42	127,31
sur	22,35	15,59	15,48	72,72	0,67	25,28	10,02	37,91
comme	3,06	14,68	0,14	92,14	18,09	0,48	1,49	5,15
son*	36,57	25,64	55,44	23,24	2,54	12,93	21,96	6,10
même	0,15	19,02	45,75	99,72	0,25	2,86	1,77	0,38
si	0,58	0,19	15,56	4,85	0,32	4,43	0,59	62,01
ça	47,10	118,82	430,93	1.175,54	34,95	54,44	74,54	78,89
bien	0,72	0,24	3,86	9,13	2,85	8,08	1,54	0,44
dit*	0,48	9,85	1,01	5,86	269,03	50,16	9,56	2,62
ils*	62,71	103,03	32,72	5,93	23,99	8,35	61,11	127,52
fait*	0,03	5,79	38,08	67,01	15,84	9,72	0,22	0,49
suis*	32,95	0,06	134,72	178,22	50,93	1,53	14,97	1,25
χ^2	2.162,16	2.107,69	3.796,45	9.410,77	7.135,68	3.199,77	3.196,91	4.744,52

Les valeurs significatives du χ^2 à un degré de liberté sont les suivantes:

au seuil de 0,05: 3,84;

au seuil de 0,01: 6,63;

au seuil de 0,001: 10,82.

Appendice 11.4

Les Huit Romans: χ^2 (abstraction faite des marqueurs de personne)

Texte	Clair	Tick	Câl	Vie	Étran	Vip	Imm	Porte
de	6,11	15,38	0,88	307,12	34,08	0,34	13,06	3,71
et	6,62	83,9	53,47	377,43	94,96	73,11	0,23	8,77
à	21,72	4,83	17,16	30,65	8,03	10,32	0,32	3,77
que	9,3	34,41	42,25	119,62	110,71	211,3	148,29	426,79
un	16,39	8,80	1,94	1,66	0,89	14,83	13,7	47,47
pas	60,79	7,19	46,31	250,05	0,07	14,33	6,73	2,44
d'	0,11	6,58	1,87	156,68	9,96	5,01	0,10	1,98
ne	56,19	19,84	0,53	8,39	10,35	150,27	44,74	118,95
qu'	7,85	12,4	1,98	130,45	19,72	1,80	2,62	8,39
en	47,08	9,34	0,07	13,01	13,93	16,45	0,16	5,88
une	8,35	22,52	0,23	52,30	14,78	14,18	33,72	73,91
qui	0,39	6,42	1,32	106,48	21,21	10,66	26,69	12,98
ce	22,78	10,31	0,37	0,25	0,02	132,09	14,85	75,18
n'	24,89	0,00	4,44	25,79	9,59	32,57	0,51	17,10
des	23,59	14,09	2,48	0,61	48,55	44,69	74,24	221,18
c'	101,66	36,24	316,74	498,67	10,29	30,98	9,86	23,11

dans	8,45	1,65	24,16	12,97	6,91	12,06	47,25	16,89
pour	0,15	2,73	50,97	181,14	2,47	0,02	0,51	1,32
du	0,32	0,01	0,01	22,9	0,48	15,91	3,29	3,84
plus	0,01	6,2	9,48	36,75	5,29	0,42	152,57	49,06
mais	7,34	0,99	19,74	90,67	56,64	13,02	11,78	21,77
au	1,34	1,50	5,54	36,16	0,20	2,60	2,79	2,59
tout	2,22	3,32	22,29	23,90	7,16	8,41	0,66	1,19
y	171,94	84,2	120,95	190,96	0,31	2,26	1,03	1,29
avec	5,85	5,94	155,83	24,53	4,57	12,09	17,85	16,29
sur	22,35	15,59	15,48	72,72	0,67	25,28	10,02	37,91
comme	3,06	14,68	0,14	92,14	18,09	0,48	1,49	5,15
même	0,16	18,86	45,48	99,35	0,25	2,8	1,77	0,40
si	0,56	0,21	15,60	4,86	0,33	4,34	0,56	61,93
ça	47,28	119,4	432,67	1.178,81	34,74	54,55	74,17	78,8
bien	0,73	0,21	3,89	9,31	2,87	8,07	1,54	0,47
χ^2	685,58	567,74	1.414,27	4.156,33	548,12	925,24	717,10	1.350,51

Appendice 11.5

Les Six Romans: Valeurs calculées (*Clair* et *Ticket* représentent la valeur hypothétique)

TYPE	<i>Clair/Tick</i>	<i>Câl</i>	<i>Vie</i>	<i>Étran</i>	<i>Vip</i>	<i>Imm</i>	<i>Porte</i>
de	1.716,50	1.994,71	2.437,78	1.296,62	2.225,48	1.545,17	1.496,22
je	1.186,50	1.378,81	1.685,07	896,26	1.538,32	1.068,07	1.034,24
et	1.097,50	1.275,38	1.558,67	829,03	1.422,93	987,95	956,66
la	885,00	1.028,44	1.256,88	668,51	1.147,42	796,66	771,43
à	758,00	880,86	1.076,51	572,58	982,76	682,34	660,73
il	833,50	968,59	1.183,74	629,61	1.080,65	750,30	726,54
que	668,50	776,85	949,41	504,97	866,72	601,77	582,71
le	651,00	756,51	924,55	491,75	844,03	586,02	567,46
l'	629,50	731,53	894,02	475,51	816,16	566,67	548,72
un	724,50	841,93	1028,94	547,28	939,33	652,18	631,53
pas	565,00	656,58	802,41	426,79	732,53	508,60	492,49
les	499,50	580,46	709,39	377,31	647,61	449,64	435,40
d'	570,00	662,39	809,52	430,57	739,02	513,11	496,85
est	560,00	650,77	795,31	423,01	726,05	504,10	488,14
ne	509,00	591,50	722,88	384,49	659,93	458,19	443,68
j'	370,00	429,97	525,48	279,49	479,71	333,07	322,52
qu'	342,50	398,01	486,42	258,72	444,06	308,31	298,55
en	389,00	452,05	552,46	293,84	504,35	350,17	339,08
elle	310,50	360,83	440,97	234,55	402,57	279,51	270,65
une	542,50	630,43	770,46	409,80	703,36	488,35	472,88
me	353,00	410,21	501,33	266,65	457,67	317,77	307,70

a	320,00	371,87	454,46	241,72	414,89	288,06	278,93
qui	345,50	401,50	490,68	260,98	447,95	311,01	301,16
ce	363,00	421,84	515,53	274,20	470,64	326,77	316,42
n'	343,50	399,17	487,84	259,47	445,35	309,21	299,42
des	378,00	439,27	536,84	285,53	490,08	340,27	329,49
c'	372,00	432,29	528,32	281,00	482,31	334,87	324,26
dans	330,00	383,49	468,67	249,28	427,85	297,06	287,65
pour	294,50	342,23	418,25	222,46	381,83	265,10	256,71
était	269,00	312,60	382,03	203,20	348,76	242,15	234,48
m'	221,00	256,82	313,86	166,94	286,53	198,94	192,64
vous	445,50	517,71	632,70	336,52	577,60	401,03	388,33
plus	260,50	302,72	369,96	196,78	337,74	234,50	227,07
mais	241,00	280,06	342,27	182,05	312,46	216,94	210,07
ai	186,00	216,15	264,16	140,50	241,15	167,43	162,13
nous	248,50	288,78	352,92	187,71	322,19	223,70	216,61
avait	217,50	252,75	308,89	164,30	281,99	195,79	189,59
du	252,00	292,84	357,89	190,36	326,72	226,85	219,66
s'	187,50	217,89	266,29	141,63	243,10	168,78	163,44
on	252,50	293,43	358,60	190,73	327,37	227,30	220,10
se	235,50	273,67	334,46	177,89	305,33	211,99	205,28
mon	194,00	225,44	275,52	146,54	251,52	174,64	169,10
au	221,00	256,82	313,86	166,94	286,53	198,94	192,64
tout	184,00	213,82	261,32	138,99	238,56	165,63	160,39
moi	199,00	231,25	282,62	150,32	258,01	179,14	173,46
y	268,50	312,02	381,32	202,82	348,12	241,70	234,04
avec	189,50	220,21	269,13	143,15	245,69	170,59	165,18
lui	125,00	145,26	177,53	94,42	162,07	112,52	108,96
tu	238,50	277,16	338,72	180,16	309,22	214,69	207,89
sur	171,50	199,30	243,56	129,55	222,35	154,38	149,49
comme	140,50	163,27	199,54	106,13	182,16	126,48	122,47
son	130,50	151,65	185,34	98,58	169,20	117,47	113,75
même	148,50	172,57	210,90	112,17	192,53	133,68	129,44
si	142,50	165,60	202,38	107,64	184,75	128,28	124,21
ça	190,00	220,80	269,84	143,52	246,34	171,04	165,62
bien	126,50	147,00	179,66	95,56	164,01	113,87	110,27
dit	95,00	110,40	134,92	71,76	123,17	85,52	82,81
ils	67,00	77,86	95,15	50,61	86,87	60,31	58,40
fait	96,50	112,14	137,05	72,89	125,11	86,87	84,12
suis	93,00	108,07	132,08	70,25	120,58	83,72	81,07

Appendice 11.6

Les Six romans: χ^2 (*Clair* et *Ticket* représentent la valeur théorique)

Type	<i>Câl</i>	<i>Vie</i>	<i>Étran</i>	<i>Vip</i>	<i>Imm</i>	<i>Porte</i>
de*	4,59	347,03	45,40	0,50	21,48	8,72
je	0,19	114,41	27,59	0,56	52,56	44,60
et*	0,27	118,81	15,39	203,36	24,62	64,11
la	12,73	53,32	0,57	0,00	1,31	3,04
à*	65,47	3,18	0,00	53,47	14,85	27,70
il	5,00	6,00	148,13	27,90	75,05	191,02
que*	1,97	24,85	35,57	72,27	51,02	224,00
le	11,31	4,49	97,75	17,91	31,55	0,23
l'	3,48	60,21	7,69	0,36	18,84	0,11
un*	22,60	25,17	3,58	52,62	42,35	89,34
pas*	2,02	82,05	16,45	2,47	42,26	7,19
les	21,82	18,52	1,57	22,75	0,10	49,90
d'*	7,27	187,42	17,41	0,54	2,27	0,03
est	47,73	33,28	17,09	191,68	86,74	48,04
ne*	22,95	11,93	52,07	25,64	1,24	24,06
j'	50,28	255,66	444,60	8,89	25,37	14,12
qu'*	4,86	274,76	61,64	29,24	25,51	40,86
en*	28,72	4,81	0,13	1,99	20,55	4,70
elle	33,43	1162,64	0,31	37,84	9,12	290,39
une*	18,65	119,52	44,34	57,08	76,55	126,81
me	24,27	17,75	1,40	41,21	97,74	28,29
a	1,57	438,74	1257,25	80,62	157,59	75,31
qui	6,10	145,63	13,79	22,35	17,61	6,77
ce*	10,27	21,20	8,47	41,27	0,03	21,04
n'*	0,04	5,58	1,32	9,38	1,60	3,99
des*	10,75	41,45	13,70	4,52	26,67	148,89
c'*	60,49	116,12	8,20	138,34	73,49	96,90
dans*	55,21	1,75	0,94	1,69	27,91	5,74
pour*	33,31	138,58	5,98	1,14	0,10	4,43
du*	0,09	21,10	0,28	14,46	2,72	3,24
était	2,80	379,90	130,44	0,85	0,06	26,27
m'	0,74	0,12	365,63	65,00	143,67	170,74
vous	19,59	90,81	195,54	146,21	125,15	205,26
plus*	18,94	18,19	11,12	0,73	112,61	30,29
mais*	5,98	49,93	32,53	2,43	2,89	8,37
ai	80,43	490,19	1205,19	5,12	43,59	0,73
nous	73,59	132,10	0,32	1,39	14,68	61,47
avait	23,92	422,14	82,90	0,29	22,78	24,81
s'	0,38	36,59	52,66	0,10	5,41	6,89
on	47,92	83,85	4,03	127,58	16,00	115,00
se	0,15	2,61	29,87	13,55	0,04	13,83
mon	0,09	59,95	6,75	8,59	14,52	241,05
au*	4,55	38,46	0,38	1,92	3,38	3,15

tout*	27,14	29,42	9,33	6,23	1,43	2,16
moi	0,50	5,01	0,02	1,12	3,18	14,15
y*	0,21	0,49	40,67	99,50	44,49	65,74
avec*	143,53	20,28	3,34	14,50	20,12	18,43
lui	2,96	127,55	227,54	36,52	7,72	41,25
tu	159,35	123,73	118,58	193,77	117,30	30,10
sur*	0,57	16,59	11,41	0,05	1,03	6,63
comme*	9,67	217,78	3,82	8,71	20,18	0,25
son	4,01	4,43	15,77	10,83	0,36	10,92
même*	15,33	43,79	1,79	0,69	0,56	1,84
si*	16,08	5,18	0,27	4,02	0,47	60,64
ça*	63,28	288,81	99,54	164,56	155,41	157,71
bien*	7,41	15,25	1,40	4,78	3,56	1,71
dit	12,81	29,49	408,61	27,47	1,55	0,12
ils	48,01	507,94	26,16	160,65	1,89	22,69
fait	20,43	38,83	7,32	18,50	2,54	3,09
suis	59,11	80,20	18,19	13,65	30,73	9,72
χ^2	1.436,93	7.215,57	5.459,71	2.301,38	1.946,12	3.008,60
χ^2 **	662,22	2.289,48	546,47	1.018,10	821,35	1.258,67

Appendice 11.7

Les Neuf Textes: Valeurs calculées

Type	Clair	Tick	Câl	Vie	Étran	Vip	Imm	Porte	Eng
de	1219,14	2056,66	1991,13	2543,13	1341,91	2046,81	1398,02	1363,97	17095,11
je	682,86	1151,98	1115,28	1424,46	751,63	1146,46	783,06	763,99	9575,32
et	736,53	1242,51	1202,92	1536,40	810,70	1236,56	844,60	824,03	10327,80
la	734,45	1239,01	1199,53	1532,08	808,42	1233,08	842,22	821,71	10298,72
à	640,54	1080,59	1046,16	1336,18	705,05	1075,41	734,53	716,64	8981,92
il	622,60	1050,32	1016,86	1298,76	685,31	1045,29	713,96	696,57	8730,36
que	503,93	850,13	823,04	1051,21	554,68	846,06	577,88	563,80	7066,30
le	638,23	1076,68	1042,38	1331,35	702,50	1071,52	731,88	714,05	8949,44
l'	499,81	843,17	816,31	1042,61	550,15	839,13	573,15	559,19	7008,51
un	474,45	800,39	774,89	989,71	522,23	796,56	544,07	530,82	6652,91
pas	382,24	644,83	624,28	797,35	420,73	641,74	438,32	427,65	5359,87
les	488,94	824,83	798,55	1019,93	538,18	820,88	560,68	547,02	6856,03
d'	397,43	670,46	649,10	829,05	437,46	667,25	455,75	444,65	5572,90
est	347,73	586,62	567,93	725,37	382,75	583,81	398,75	389,04	4876,01
ne	333,56	562,71	544,78	695,81	367,15	560,02	382,50	373,19	4677,29
j'	280,09	472,51	457,46	584,28	308,30	470,25	321,19	313,37	3927,56
qu'	320,49	540,66	523,43	668,54	352,76	538,07	367,51	358,56	4493,99
en	352,13	594,04	575,11	734,55	387,59	591,19	403,80	393,96	4937,66
elle	332,11	560,26	542,41	692,78	365,55	557,58	380,84	371,56	4656,92
une	328,97	554,96	537,28	686,23	362,10	552,31	377,24	368,05	4612,89

me	233,57	394,04	381,48	487,24	257,10	392,15	267,85	261,32	3275,26
a	232,04	391,45	378,98	484,05	255,41	389,58	266,09	259,61	3253,79
qui	280,80	473,71	458,61	585,75	309,08	471,44	322,00	314,16	3937,47
ce	247,04	416,75	403,47	515,33	271,92	414,76	283,29	276,39	3464,07
n'	249,91	421,59	408,15	521,31	275,07	419,57	286,57	279,59	3504,25
des	317,31	535,29	518,24	661,91	349,26	532,73	363,87	355,01	4449,40
c'	220,78	372,45	360,58	460,54	243,01	370,66	253,17	247,01	3095,81
dans	264,16	445,63	431,43	551,03	290,76	443,49	302,92	295,54	3704,07
pour	225,37	380,20	368,08	470,13	248,07	378,38	258,44	252,14	3160,21
était	202,17	341,06	330,19	421,73	222,53	339,42	231,83	226,19	2834,89
m'	166,49	280,86	271,91	347,29	183,25	279,51	190,91	186,26	2334,52
vous	206,65	348,61	337,50	431,06	227,46	346,94	236,97	231,19	2897,64
plus	195,97	330,59	320,06	408,79	215,70	329,01	224,72	219,25	2747,92
mais	185,37	312,71	302,75	386,68	204,04	311,22	212,57	207,39	2599,29
ai	162,32	273,84	265,11	338,61	178,67	272,53	186,14	181,61	2276,17
nous	161,42	272,32	263,64	336,73	177,68	271,01	185,11	180,60	2263,51
avait	168,21	283,77	274,73	350,89	185,15	282,41	192,90	188,20	2358,74
du	184,11	310,59	300,70	384,06	202,65	309,11	211,13	205,98	2581,68
s'	187,65	316,55	306,47	391,43	206,54	315,04	215,18	209,94	2631,22
on	155,85	262,91	254,54	325,10	171,54	261,65	178,71	174,36	2185,34
se	199,78	337,02	326,28	416,73	219,89	335,40	229,09	223,51	2801,31
mon	125,46	211,65	204,91	261,72	138,10	210,64	143,87	140,37	1759,28
au	162,01	273,31	264,60	337,96	178,33	272,00	185,78	181,26	2271,76
tout	149,45	252,12	244,08	311,75	164,50	250,91	171,38	167,20	2095,62
moi	125,07	210,99	204,27	260,90	137,67	209,98	143,42	139,93	1753,78
y	129,23	218,01	211,07	269,58	142,25	216,97	148,19	144,58	1812,13
avec	145,68	245,76	237,93	303,89	160,35	244,58	167,06	162,99	2042,77
lui	153,14	258,34	250,11	319,45	168,56	257,11	175,61	171,33	2147,36
tu	136,45	230,20	222,86	284,65	150,20	229,09	156,48	152,67	1913,41
sur	156,59	264,17	255,75	326,66	172,36	262,91	179,57	175,20	2195,80
comme	135,91	229,27	221,97	283,50	149,59	228,17	155,85	152,05	1905,71
son	137,67	232,25	224,85	287,19	151,54	231,14	157,87	154,03	1930,48
même	105,68	178,27	172,59	220,44	116,32	177,42	121,18	118,23	1481,81
si	109,38	184,52	178,64	228,16	120,39	183,63	125,43	122,37	1533,73
ça	95,52	161,14	156,01	199,26	105,14	160,37	109,54	106,87	1339,42
bien	102,18	172,37	166,88	213,14	112,47	171,54	117,17	114,32	1432,75
dit	89,64	151,23	146,41	187,00	98,67	150,51	102,80	100,29	1257,03
ils	121,01	204,14	197,64	252,43	133,20	203,17	138,77	135,39	1696,85
fait	70,72	119,30	115,50	147,52	77,84	118,73	81,10	79,12	991,65
suis	65,02	109,69	106,19	135,63	71,57	109,16	74,56	72,74	911,71

Appendice 11.8

Les Neuf Textes: χ^2

Type	Clair	Tick	Câl	Vie	Étran	Vip	Imm	Porte	Eng
de	4,75	26,46	4,28	413,23	61,77	10,30	0,88	0,24	73,36
je	66,52	91,70	70,16	22,36	0,21	114,64	347,89	307,91	416,20
et	0,00	37,72	6,90	133,33	21,26	99,95	0,19	16,06	16,42
la	37,27	1,16	67,96	186,18	31,44	5,73	0,21	11,86	184,46
à	26,20	5,29	5,35	75,77	25,11	17,35	3,20	8,79	11,19
il	2,01	1,63	13,66	0,73	90,98	18,30	56,56	168,47	39,42
que	1,78	0,19	0,06	2,55	12,82	86,77	68,61	256,39	131,05
le	29,08	71,10	35,87	88,03	0,10	10,20	0,13	25,54	158,12
l'	5,37	1,23	1,44	138,94	0,36	1,92	16,37	0,59	32,12
un	14,71	10,26	6,49	14,97	0,71	7,95	6,20	35,26	14,64
pas	26,56	0,01	7,56	85,86	14,36	3,63	13,29	0,07	20,67
les	31,92	43,69	13,95	37,64	63,72	3,28	24,70	122,65	220,66
d'	0,10	9,20	4,85	201,82	19,97	12,62	1,19	7,14	17,82
est	73,87	1,10	118,18	74,60	5,23	91,25	27,00	7,51	12,21
ne	21,88	2,34	8,94	6,22	41,98	94,45	25,88	80,95	28,53
j'	2,07	0,28	31,24	162,07	339,86	11,88	33,55	18,74	251,34
qu'	15,07	21,04	12,67	50,34	2,95	0,74	2,37	7,09	0,75
en	40,98	3,88	0,14	23,20	19,79	5,15	2,41	0,57	38,21
elle	3,10	102,18	156,56	311,06	53,28	1,79	6,79	86,65	4,27
une	11,70	34,83	0,43	70,04	20,95	4,40	17,93	53,29	33,33
me	4,00	30,13	43,30	13,21	3,25	104,93	190,95	74,66	160,16
a	13,98	4,18	0,76	359,16	1131,50	63,74	137,23	60,78	81,13
qui	0,00	8,57	0,13	50,65	37,79	12,43	22,44	10,77	1,37
ce	7,13	0,98	5,59	21,12	7,75	91,91	7,70	53,51	15,02
n'	10,04	2,84	0,42	0,67	0,03	19,49	0,00	10,59	8,17
des	12,24	2,20	0,20	0,88	45,65	15,11	38,83	171,86	99,27
c'	45,50	6,86	151,10	216,08	0,41	58,03	22,32	40,49	43,70
dans	6,41	0,17	22,07	22,37	11,08	4,07	31,01	7,97	25,60
pour	1,19	0,00	17,79	75,88	15,53	0,80	0,01	3,37	6,76
était	0,42	0,05	6,74	276,16	92,50	2,08	0,87	21,78	57,26
m'	2,77	0,93	3,07	2,14	290,55	73,66	164,26	189,22	196,63
vous	343,32	13,81	18,73	3,36	95,59	10,35	15,17	67,79	0,08
plus	2,24	0,72	27,06	4,57	20,01	0,15	132,07	37,56	6,62
mais	0,31	1,80	1,10	19,27	14,81	2,66	4,08	9,60	23,34
ai	0,33	11,80	25,91	240,53	780,05	16,24	58,27	0,41	109,03
nous	66,46	5,97	55,20	118,47	0,03	3,32	49,68	126,92	0,22
avait	2,43	0,04	36,20	290,20	49,62	0,31	21,16	23,99	10,55
du	0,09	0,58	0,02	33,28	1,91	8,45	0,39	0,82	15,39
s'	9,69	23,67	20,61	1,78	2,23	14,27	1,22	0,80	36,47
on	54,49	0,13	97,41	131,68	0,43	73,47	0,77	73,70	14,55
se	5,05	3,43	6,56	6,67	60,03	26,57	1,76	22,88	77,42
mon	6,92	30,50	3,07	50,28	11,53	36,23	45,75	378,94	91,07
au	1,58	1,88	2,63	53,10	2,09	5,31	0,88	0,97	5,16
tout	0,38	6,71	8,64	4,45	0,67	10,33	0,54	0,83	0,45
moi	3,50	7,97	6,97	0,97	1,49	20,13	24,75	49,32	53,31

y	89,87	30,83	40,92	58,35	6,43	13,93	0,70	8,27	42,69
avec	9,75	2,59	107,69	5,03	0,13	14,03	18,14	17,23	0,30
lui	23,62	39,75	28,29	0,23	31,13	1,27	6,43	0,13	20,86
tu	0,15	48,63	109,01	79,73	89,89	460,79	64,52	118,20	1,13
sur	12,70	4,17	8,19	65,84	0,11	7,33	0,88	18,67	68,63
comme	4,20	15,84	1,62	54,67	27,03	0,17	2,87	3,80	1,38
son	25,00	10,87	42,58	18,65	1,21	1,58	7,27	0,16	64,01
même	3,27	5,73	15,46	34,19	2,89	0,08	0,12	0,16	12,63
si	0,07	0,70	23,40	14,88	0,47	4,43	0,92	64,12	0,08
ça	13,92	46,68	214,08	613,01	62,79	82,76	94,42	99,03	48,33
bien	1,42	0,54	1,04	1,62	7,14	7,33	2,45	0,80	0,52
dit	1,04	11,22	0,02	0,63	211,26	48,83	8,15	2,07	0,42
ils	44,13	68,70	17,39	15,54	16,09	0,02	33,13	95,01	100,59
fait	1,31	1,36	16,98	26,68	4,19	14,49	0,97	1,59	4,48
suis	12,78	2,82	63,06	72,80	16,52	7,67	23,16	5,28	20,50
χ^2	1.258,67	921,64	1.817,71	5.157,73	3.980,68	1.941,05	1.891,56	3.089,81	3.230,05

ddl: 472, à savoir (60-1)(9-1);

seuil de rejet d'après la correction de Bonferroni: 0,009 (à savoir, 5/540);

valeurs significatives: toute valeur supérieure à 149,4.

Appendice 11.9

Les Neuf Textes: Valeurs calculées (abstraction faite des marqueurs de personne)

TYPE	Clair	Tick	Câl	Vie	Étran	Vip	Imm	Porte	Eng
de	1231,20	2152,16	2055,2	2535,00	1173,50	2048,00	1377,53	1351,92	17131,20
et	743,78	1300,19	1241,6	1531,00	708,94	1237,50	832,22	816,74	10349,60
à	647,00	1131,00	1080,00	1332,00	617,00	1076,00	724,00	710,00	9001,00
que	508,89	889,60	849,51	1047,00	485,06	846,73	569,40	558,82	7081,18
un	479,12	837,55	799,81	986,50	456,68	797,19	536,09	526,12	6666,92
pas	386,00	674,77	644,36	794,80	367,92	642,25	431,90	423,87	5371,15
d'	401,34	701,59	669,97	826,40	382,55	667,78	449,06	440,72	5584,63
ne	336,84	588,84	562,30	693,60	321,07	560,46	376,90	369,89	4687,14
qu'	323,64	565,76	540,26	666,40	308,49	538,50	362,13	355,39	4503,45
en	355,60	621,61	593,60	732,20	338,94	591,66	397,88	390,48	4948,06
ce	249,47	436,10	416,45	513,70	237,79	415,09	279,13	273,95	3471,36
une	332,21	580,73	554,56	684,00	316,65	552,75	371,71	364,80	4622,60
qui	283,56	495,70	473,36	583,90	270,28	471,81	317,28	311,38	3945,76
n'	252,37	441,16	421,28	519,60	240,55	419,90	282,37	277,12	3511,63
des	320,43	560,15	534,90	659,80	305,43	533,16	358,53	351,87	4458,77
c'	222,95	389,74	372,18	459,10	212,51	370,96	249,46	244,82	3102,33

dans	266,76	466,31	445,30	549,30	254,26	443,85	298,47	292,93	3711,87
pour	227,59	397,85	379,92	468,60	216,93	378,68	254,65	249,92	3166,87
plus	197,90	345,94	330,35	407,50	188,63	329,27	221,43	217,31	2753,70
mais	187,19	327,23	312,48	385,40	178,43	311,46	209,45	205,56	2604,76
du	185,92	325,01	310,37	382,80	177,22	309,35	208,03	204,16	2587,11
au	163,61	286,00	273,11	336,90	155,94	272,22	183,06	179,66	2276,55
tout	150,92	263,82	251,93	310,70	143,85	251,11	168,86	165,73	2100,03
y	130,50	228,13	217,85	268,70	124,39	217,14	146,02	143,31	1815,94
avec	147,11	257,17	245,58	302,90	140,22	244,78	164,61	161,55	2047,07
sur	158,13	276,43	263,98	325,60	150,73	263,11	176,94	173,65	2200,42
comme	137,24	239,91	229,10	282,60	130,82	228,35	153,56	150,71	1909,72
même	106,68	186,48	178,08	219,70	101,68	177,50	119,36	117,14	1484,42
si	110,48	193,14	184,43	227,50	105,31	183,83	123,62	121,32	1537,37
ça	96,45	168,61	161,01	198,60	91,93	160,48	107,92	105,91	1342,10
bien	103,15	180,32	172,19	212,40	98,32	171,63	115,42	113,27	1435,32

Appendice 11.10

Les Neuf Textes: χ^2 (abstraction faite des marqueurs de personne)

TYPE	Clair	Tick	Câl	Vie	Étran	Vip	Imm	Porte	Eng
de	8,83	6,31	11,87	407,95	12,17	10,06	0,15	0,67	68,57
et	19,40	0,08	2,21	136,71	76,62	100,43	0,00	14,21	18,16
à	2,09	0,05	1,05	0,01	11,19	0,01	0,19	0,07	0,04
que	0,80	2,39	1,32	2,91	48,85	86,27	75,69	265,50	134,85
un	3,41	12,99	11,48	14,24	4,70	8,07	4,68	33,18	13,33
pas	1,14	24,37	3,67	87,84	1,69	3,55	11,31	0,20	22,05
d'	3,20	0,27	8,84	199,83	3,88	12,46	2,00	8,25	16,48
ne	0,18	20,04	13,55	5,83	18,98	94,01	29,31	84,80	30,02
qu'	30,69	16,31	17,87	51,70	18,98	0,71	3,36	8,09	1,01
en	9,20	42,96	1,28	22,44	4,47	5,24	3,46	0,34	36,32
une	22,09	10,41	1,91	68,85	5,48	4,48	15,83	51,30	31,63
qui	14,82	0,02	1,06	51,94	17,76	12,30	20,31	9,85	1,08
ce	0,00	6,26	8,77	20,52	0,58	91,53	9,27	56,18	15,96
n'	6,65	8,99	1,64	0,80	5,83	19,33	0,08	11,67	8,88
des	6,25	13,36	1,35	1,04	22,24	15,25	35,95	169,02	96,29
c'	2,84	43,12	132,21	218,83	1,98	58,22	20,47	39,09	45,17
dans	1,84	7,18	15,73	21,73	1,61	4,14	28,65	7,20	24,27
pour	0,80	1,52	12,56	77,36	4,41	0,83	0,11	2,90	7,38
plus	0,00	2,65	32,33	4,87	7,91	0,16	139,21	39,53	7,19
mais	4,47	0,18	0,23	19,90	36,39	2,61	5,06	10,49	24,33
du	0,00	0,19	0,49	32,66	0,19	8,52	0,17	0,61	14,53

au	0,35	1,89	1,17	52,40	0,06	5,24	0,55	0,76	4,70
tout	10,58	0,24	5,75	4,71	6,74	10,40	0,87	1,06	0,58
y	22,64	86,90	34,07	59,36	1,23	14,00	0,44	7,74	43,77
avec	0,74	10,40	94,60	5,31	4,38	14,11	16,81	16,45	0,41
sur	7,47	13,46	11,04	65,11	1,98	7,40	0,56	17,83	66,86
comme	20,96	4,64	2,97	55,66	15,35	0,18	3,58	3,42	1,17
même	2,97	3,63	11,84	34,74	0,13	0,07	0,27	0,08	13,09
si	2,10	0,02	26,90	14,53	0,56	4,32	1,24	66,29	0,05
ça	37,39	13,10	196,77	618,28	50,20	83,10	92,51	98,06	49,63
bien	1,66	1,68	0,35	1,81	2,09	7,40	2,99	1,02	0,39
χ^2	245,54	355,62	666,91	2.359,85	388,64	684,37	525,10	1.025,9	798,22

ddl: 242, à savoir $(31-1) \times (9-1)$;

seuil de rejet d'après la correction de Bonferroni: 0,179 (à savoir $5/279$);

valeurs significatives - toute valeur supérieure à 135,8.

Appendice 11.11

Les Huit Romans: Valeurs calculées

Type	Clair	Tick	Câl	Vie	Étran	Vip	Imm	Porte
de	1121,34	1891,68	1831,41	2339,13	1234,27	1882,62	1285,87	1254,56
je	857,19	1446,07	1400,00	1788,12	943,52	1439,15	982,97	959,03
et	772,49	1303,17	1261,65	1611,42	850,28	1296,93	885,83	864,26
la	614,09	1035,96	1002,96	1281,00	675,94	1031,00	704,20	687,05
à	612,86	1033,88	1000,94	1278,42	674,58	1028,93	702,78	685,66
il	571,38	963,90	933,19	1191,90	628,92	959,28	655,21	639,26
que	587,97	991,89	960,29	1226,51	647,18	987,14	674,24	657,82
le	534,35	901,44	872,72	1114,66	588,16	897,12	612,75	597,83
l'	458,38	773,27	748,63	956,18	504,54	769,57	525,63	512,83
un	447,20	754,42	730,38	932,86	492,23	750,80	512,81	500,33
pas	411,31	693,87	671,76	857,99	452,73	690,54	471,66	460,17
les	381,53	643,63	623,12	795,87	419,95	640,55	437,51	426,85
d'	369,91	624,04	604,16	771,64	407,17	621,05	424,19	413,86
est	369,04	622,57	602,73	769,82	406,21	619,58	423,19	412,88
ne	365,46	616,53	596,88	762,35	402,27	613,57	419,08	408,88
j'	366,86	618,88	599,16	765,27	403,80	615,92	420,69	410,44
qu'	325,55	549,20	531,70	679,11	358,34	546,57	373,32	364,23
en	314,20	530,05	513,16	655,42	345,84	527,51	360,30	351,53
elle	319,79	539,48	522,29	667,08	351,99	536,89	366,71	357,78
une	294,73	497,20	481,36	614,80	324,41	494,82	337,97	329,74
me	296,82	500,73	484,78	619,17	326,71	498,34	340,37	332,08
a	276,91	467,15	452,26	577,64	304,80	464,91	317,54	309,81

qui	274,38	462,87	448,13	572,36	302,01	460,66	314,64	306,98
ce	266,96	450,35	436,00	556,87	293,84	448,19	306,13	298,67
n'	264,69	446,52	432,29	552,14	291,34	444,38	303,52	296,13
des	259,27	437,39	423,45	540,84	285,38	435,29	297,31	290,07
c'	252,90	426,63	413,04	527,55	278,37	424,59	290,00	282,94
dans	237,27	400,26	387,51	494,94	261,16	398,35	272,08	265,45
pour	238,14	401,74	388,94	496,76	262,12	399,81	273,08	266,43
était	237,35	400,41	387,65	495,12	261,26	398,49	272,18	265,55
m'	225,65	380,67	368,54	470,71	248,38	378,85	258,76	252,46
vous	205,30	346,34	335,31	428,27	225,98	344,69	235,43	229,69
plus	207,75	350,47	339,30	433,37	228,67	348,79	238,23	232,43
mais	206,88	349,00	337,88	431,55	227,71	347,32	237,23	231,45
ai	205,83	347,23	336,17	429,36	226,56	345,57	236,03	230,28
nous	163,39	275,63	266,85	340,83	179,84	274,31	187,36	182,80
avait	181,99	307,01	297,23	379,63	200,32	305,54	208,69	203,61
du	166,71	281,23	272,27	347,75	183,49	279,88	191,17	186,51
s'	160,59	270,92	262,29	335,00	176,77	269,62	184,16	179,67
on	171,42	289,19	279,97	357,59	188,68	287,80	196,57	191,79
se	159,11	268,41	259,86	331,90	175,13	267,13	182,45	178,01
mon	160,42	270,62	262,00	334,63	176,57	269,33	183,96	179,48
au	152,56	257,36	249,16	318,24	167,92	256,13	174,94	170,68
tout	152,12	256,63	248,45	317,33	167,44	255,40	174,44	170,19
moi	151,77	256,04	247,88	316,60	167,06	254,81	174,04	169,80
y	153,52	258,98	250,73	320,24	168,98	257,74	176,05	171,76
avec	147,84	249,41	241,46	308,40	162,73	248,21	169,54	165,41
lui	134,66	227,16	219,93	280,90	148,22	226,08	154,42	150,65
tu	140,51	237,03	229,48	293,10	154,66	235,90	161,12	157,20
sur	122,69	206,98	200,39	255,94	135,05	205,99	140,70	137,27
comme	131,43	221,71	214,65	274,16	144,66	220,65	150,71	147,04
son	106,97	180,46	174,71	223,15	117,75	179,60	122,67	119,68
même	117,62	198,43	192,10	245,36	129,47	197,48	134,88	131,60
si	108,39	182,86	177,03	226,11	119,31	181,98	124,30	121,27
ça	117,74	198,62	192,29	245,60	129,60	197,67	135,01	131,73
bien	99,80	168,36	162,99	208,18	109,85	167,55	114,44	111,65
dit	87,64	147,85	143,14	182,82	96,47	147,14	100,50	98,05
ils	84,93	143,28	138,71	177,17	93,49	142,59	97,39	95,02
fait	76,54	129,12	125,01	159,66	84,25	128,50	87,77	85,63
suis	76,96	129,82	125,69	160,53	84,71	129,20	88,25	86,10

Appendice 11.12

Les Huit Romans: χ^2

Type	Clair	Tick	Câl	Vie	Étran	Vip	Imm	Porte
de	0,42	83,85	2,48	288,25	26,33	50,84	4,63	12,95
je	1,76	0,66	0,02	164,36	44,33	3,39	105,50	87,67
et	1,72	18,63	0,83	88,47	9,89	130,84	3,27	27,89
la	3,31	26,34	7,89	62,52	1,07	13,50	22,12	1,88
à	16,93	0,81	14,40	53,05	15,60	32,57	9,16	17,76
il	13,13	2,11	1,25	4,86	148,96	2,85	30,87	127,29
que	22,09	16,75	21,68	12,44	0,10	17,08	15,66	124,50
le	1,96	11,41	0,64	14,17	25,65	5,44	19,48	0,59
l'	0,23	1,84	1,49	90,51	1,96	1,13	39,65	1,55
un	27,45	24,73	0,95	4,51	0,24	1,52	1,40	22,60
pas	12,50	3,17	0,67	47,09	26,60	0,00	25,49	1,60
les	0,81	0,12	7,84	0,99	10,67	25,76	0,07	45,17
d'	1,20	25,02	0,21	160,25	9,80	30,64	7,08	18,35
est	52,32	0,18	83,45	46,00	11,45	114,70	38,83	14,69
ne	7,84	0,50	24,89	22,98	63,06	50,73	9,45	46,66
j'	33,50	29,40	0,82	20,99	128,96	8,17	0,04	1,02
qu'	17,07	24,16	15,13	44,02	1,98	0,24	1,50	5,50
en	21,50	0,48	5,44	4,03	6,08	0,14	15,49	2,15
elle	1,22	88,48	140,91	359,80	45,10	0,22	3,68	104,35
une	31,45	77,90	3,43	35,53	7,52	0,14	5,46	31,39
me	29,66	0,01	1,31	72,71	5,07	18,75	69,34	14,30
a	0,53	28,88	7,00	181,01	781,96	116,68	185,26	99,77
qui	0,16	6,04	0,02	60,21	33,78	16,56	19,16	8,46
ce	1,82	0,40	14,68	38,21	15,66	58,41	1,86	33,03
n'	4,71	7,93	3,22	0,27	0,61	9,69	0,90	4,84
des	0,07	9,25	16,88	38,96	13,64	0,14	9,21	114,28
c'	18,34	0,03	79,28	117,01	7,39	94,76	43,26	65,31
dans	0,86	3,37	51,66	6,10	2,82	0,02	16,05	1,28
pour	3,57	1,18	9,28	52,99	22,11	3,77	0,63	7,08
était	8,29	7,67	28,25	144,93	41,99	2,65	2,52	45,19
m'	28,83	18,39	42,76	48,25	110,44	5,15	46,12	58,51
vous	349,05	14,82	19,90	2,90	94,30	9,65	14,50	66,61
plus	5,16	0,06	37,17	0,80	27,07	2,06	105,81	25,89
mais	0,93	10,31	0,84	3,98	4,30	0,15	0,10	1,82
ai	12,55	48,84	0,42	88,24	467,49	56,37	100,52	14,25
nous	63,19	6,91	57,48	121,90	0,00	2,60	46,80	121,78
avait	6,35	1,30	50,26	222,10	32,50	3,47	30,43	33,52
du	1,06	6,50	2,43	16,94	0,00	1,71	0,61	0,23
s'	1,51	6,18	4,75	2,69	14,85	1,73	1,20	1,67
on	34,21	3,58	62,26	85,07	3,50	94,37	4,45	89,19
se	0,50	4,46	1,56	3,10	28,08	2,56	3,86	3,80
mon	25,87	1,69	3,91	105,21	0,01	3,05	9,16	204,38
au	0,28	5,80	7,02	41,01	0,47	11,33	0,02	0,04
tout	0,16	8,11	6,95	3,16	0,34	12,02	0,25	0,46
moi	0,22	0,06	0,14	16,19	1,36	1,60	4,82	16,67

y	45,39	6,50	11,32	17,45	19,21	35,57	8,22	22,21
avec	10,74	1,87	101,48	3,88	0,03	15,59	19,53	18,56
lui	12,89	21,67	13,22	7,90	58,08	0,74	1,00	4,26
tu	0,00	41,32	115,04	86,36	94,13	428,95	68,59	107,18
sur	0,93	2,79	0,46	22,53	8,04	0,82	4,92	2,70
comme	2,87	12,53	0,63	65,34	23,79	0,01	4,59	2,47
son	7,32	0,01	13,03	0,38	3,48	5,84	0,01	7,18
même	7,93	0,70	5,38	15,61	7,66	1,33	0,72	2,39
si	0,12	0,51	22,45	13,97	0,35	5,00	1,13	66,32
ça	1,72	12,21	111,56	374,17	86,22	117,66	119,79	123,86
bien	0,94	0,19	1,78	2,67	6,02	5,91	3,38	1,33
dit	0,67	9,67	0,17	1,24	222,76	46,10	7,06	1,51
ils	16,12	23,13	0,00	107,32	0,46	27,33	7,17	56,16
fait	3,12	0,07	9,68	16,03	1,61	20,44	2,76	3,66
suis	3,70	10,97	30,91	34,54	5,33	18,57	34,58	12,62
χ^2	980,75	782,46	1.280,99	3.768,16	2.832,30	1.749,01	1.359,07	2.160,33

ddl: 413, à savoir $(60-1) \times (8-1)$;

seuil de rejet d'après la correction de Bonferroni: 0,0104 (à savoir $5/480$);

valeurs significatives - toute valeur supérieure à 135,8.

Appendice 11.13

Les Huit Romans: Valeurs calculées (abstraction faite des marqueurs de personne)

Type	Clair	Tick	Câl	Vie	Étran	Vip	Imm	Porte
de	1.122,12	1.975,05	1.905,49	2.306,33	1.079,81	1.910,67	1.281,01	1.260,52
et	773,02	1.360,58	1.312,67	1.588,80	743,87	1.316,24	882,47	868,36
à	613,27	1.079,42	1.041,41	1.260,48	590,15	1.044,24	700,11	688,91
que	588,37	1.035,59	999,12	1.209,29	566,18	1.001,83	671,68	660,94
un	447,50	787,65	759,91	919,77	430,63	761,98	510,87	502,70
pas	411,59	724,44	698,92	845,95	396,07	700,82	469,87	462,35
d'	370,17	651,53	628,58	760,81	356,21	630,29	422,58	415,82
ne	365,71	643,69	621,02	751,65	351,92	622,71	417,49	410,81
qu'	325,77	573,40	553,20	669,57	313,49	554,71	371,90	365,95
en	314,41	553,40	533,91	646,23	302,56	535,36	358,94	353,19
une	294,93	519,10	500,82	606,17	283,81	502,18	336,69	331,30
qui	274,57	483,26	466,24	564,32	264,21	467,51	313,44	308,43
ce	267,14	470,19	453,63	549,06	257,07	454,86	304,97	300,09
n'	264,87	466,19	449,77	544,39	254,88	451,00	302,37	297,53

des	259,45	456,66	440,57	533,25	249,67	441,77	296,19	291,45
c'	253,07	445,43	429,74	520,14	243,53	430,91	288,90	284,28
dans	237,43	417,90	403,18	487,99	228,47	404,27	271,05	266,71
pour	238,30	419,43	404,66	489,79	229,32	405,76	272,04	267,69
plus	207,89	365,91	353,02	427,28	200,05	353,98	237,33	233,53
mais	207,02	364,37	351,54	425,49	199,21	352,49	236,33	232,55
du	166,82	293,62	283,28	342,87	160,53	284,05	190,44	187,39
au	152,66	268,70	259,24	313,77	146,91	259,94	174,28	171,49
tout	152,23	267,93	258,50	312,88	146,49	259,20	173,78	171,00
y	153,62	270,39	260,87	315,75	147,83	261,58	175,38	172,57
avec	147,94	260,40	251,23	304,07	142,37	251,91	168,89	166,19
sur	122,78	216,10	208,49	252,35	118,15	209,06	140,16	137,92
comme	131,52	231,48	223,33	270,31	126,56	223,94	150,14	147,74
même	117,62	207,03	199,73	241,75	113,19	200,28	134,28	132,13
si	108,45	190,88	184,15	222,89	104,36	184,65	123,80	121,82
ça	117,88	207,49	200,18	242,29	113,44	200,72	134,58	132,42
bien	99,88	175,80	169,61	205,29	96,12	170,07	114,03	112,20

Appendice 11.14

Les Huit Romans: χ^2 (abstraction faite des marqueurs de personne)

Type	<i>Clair</i>	<i>Tick</i>	<i>Câl</i>	<i>Vie</i>	<i>Étran</i>	<i>Vip</i>	<i>Imm</i>	<i>Porte</i>
de	0,39	50,22	0,02	269,46	0,62	41,42	5,25	11,71
et	1,77	7,12	0,27	100,81	52,77	141,29	2,89	29,25
à	17,05	5,13	6,08	46,65	0,56	26,95	9,81	16,65
que	22,23	28,76	33,56	9,34	9,37	13,24	16,51	121,23
un	27,29	13,56	4,11	2,91	12,16	2,66	1,21	23,50
pas	12,39	8,28	0,05	53,66	7,11	0,17	24,76	1,86
d'	1,17	14,58	2,01	152,67	0,42	26,28	7,53	17,45
ne	7,77	3,10	34,33	19,69	33,71	44,94	9,97	45,15
qu'	17,16	33,89	22,35	49,70	16,31	0,02	1,69	5,06
en	21,60	0,10	1,93	2,76	0,02	0,00	16,12	1,89
une	31,29	58,93	0,90	31,95	0,27	0,00	5,16	32,21
qui	0,15	11,11	0,50	66,47	15,12	13,86	18,64	8,91
ce	1,79	2,34	21,01	34,71	3,76	52,91	2,05	31,94
n'	4,66	13,45	6,67	0,04	2,10	7,72	0,78	4,47
des	0,08	4,31	10,32	43,76	2,85	0,00	8,85	115,47
c'	18,23	1,13	62,79	125,86	0,46	99,35	42,57	66,29
dans	0,88	0,87	39,26	4,72	0,13	0,03	15,61	1,46

pour	3,60	3,71	4,86	58,46	8,18	4,94	0,53	7,46
plus	5,20	1,08	44,99	1,43	12,52	2,89	107,42	25,04
mais	0,95	15,59	2,65	5,30	17,95	0,44	0,14	1,63
du	1,04	3,14	0,76	15,06	3,15	2,39	0,70	0,17
au	0,29	2,77	3,89	38,40	0,99	9,64	0,01	0,07
tout	0,15	12,10	3,84	4,17	5,55	13,52	0,30	0,37
y	45,26	3,24	7,13	19,89	8,68	37,91	7,97	22,69
avec	10,78	0,43	85,74	4,98	3,60	17,24	19,16	19,00
sur	0,95	1,03	0,01	20,74	21,03	0,47	5,14	2,88
comme	2,90	16,86	1,85	70,14	13,00	0,02	4,81	2,64
même	7,97	0,04	2,95	17,61	2,04	1,86	0,64	2,49
si	0,12	1,67	26,72	12,55	0,72	4,05	1,20	65,29
ça	1,69	7,91	96,27	388,26	70,52	120,81	119,06	124,54
bien	0,98	0,93	0,64	3,48	1,53	6,83	3,50	1,24
χ^2	267,78	327,40	528,48	1.675,63	327,19	693,85	459,99	809,99

ddl: 270, à savoir (31-1)(7-1);

seuil de rejet d'après la correction de Bonferroni: 0,02 (à savoir, 5/248);

valeurs significatives- toute valeur supérieure à 131,1

Appendice 11.15

Les Quatre Romans de Gary/Ajar: Valeurs calculées

Type	Clair	Tick	Câl	Vie
de	1.069,24	1.803,80	1.746,32	2.230,46
je	782,68	1.320,36	1.278,29	1.632,67
et	855,11	1.442,55	1.396,59	1.783,76
la	574,78	969,64	938,75	1.199,00
à	570,54	962,49	931,82	1.190,15
il	598,48	1.009,63	977,46	1.248,44
que	508,26	857,42	830,10	1.060,23
le	490,15	826,87	800,53	1.022,45
l'	421,93	711,79	689,12	880,16
un	471,57	795,53	770,19	983,71
pas	449,87	758,93	734,75	938,44
les	392,74	662,55	641,44	819,27
d'	336,08	566,96	548,90	701,07
est	453,46	764,99	740,61	945,93
ne	331,40	559,06	541,25	691,30
j'	344,82	581,71	563,17	719,30
qu'	308,92	521,14	504,54	644,41
en	304,08	512,98	496,63	634,31
elle	316,72	534,31	517,28	660,69

une	323,75	546,16	528,76	675,34
me	253,35	427,39	413,78	528,49
a	302,36	510,08	493,83	630,73
qui	296,59	500,34	484,39	618,68
ce	233,05	393,16	380,63	486,15
n'	253,19	427,13	413,52	528,16
des	304,39	513,50	497,14	634,96
c'	329,99	556,69	538,95	688,37
dans	254,28	428,97	415,30	530,44
pour	264,90	446,88	432,64	552,58
était	247,26	417,12	403,83	515,79
m'	156,88	264,65	256,22	327,25
vous	265,52	447,93	433,66	553,88
plus	187,32	316,00	305,93	390,75
mais	199,18	336,02	325,31	415,49
ai	209,80	353,92	342,65	437,64
nous	121,29	204,61	198,09	253,01
avait	199,81	337,07	326,33	416,80
du	167,49	282,56	273,56	349,39
s'	150,95	254,64	246,53	314,88
on	226,19	381,57	369,41	471,83
se	174,05	293,62	284,26	363,07
mon	119,42	201,45	195,03	249,10
au	146,26	246,74	238,88	305,11
tout	157,19	265,18	256,73	327,90
moi	138,15	233,05	225,63	288,18
y	192,94	325,48	315,11	402,47
avec	174,83	294,94	285,54	364,70
lui	116,14	195,92	189,68	242,26
tu	105,83	178,54	172,85	220,77
sur	114,42	193,02	186,87	238,68
comme	139,24	234,89	227,41	290,46
son	93,97	158,53	153,48	196,02
même	129,38	218,25	211,30	269,88
si	88,84	149,87	145,09	185,32
ça	197,83	333,74	323,11	412,68
bien	103,75	175,02	169,45	216,42
dit	83,67	141,15	136,65	174,53
ils	91,74	154,77	149,84	191,38
fait	87,91	148,31	143,58	183,39
suis	95,05	160,35	155,25	198,28

Appendice 11.16

Les Quatre Romans de Gary/Ajar: χ^2

Type	<i>Clair</i>	<i>Tick</i>	<i>Câl</i>	<i>Vie</i>
de	5,09	131,03	13,32	227,57
je	16,41	18,58	10,66	91,58
et	16,59	0,19	7,54	23,62
la	0,06	55,28	0,65	33,69
à	6,21	1,88	38,41	24,90
il	5,92	0,00	6,30	0,31
que	2,31	0,04	0,24	1,73
le	0,29	0,87	2,94	1,09
l'	1,61	13,83	12,52	54,07
un	15,84	11,46	5,69	13,61
pas	2,44	16,51	2,37	15,49
les	2,10	1,15	4,14	0,03
d'	8,97	58,45	3,54	112,68
est	6,56	30,60	10,08	0,15
ne	23,16	2,85	8,11	5,44
j'	22,88	16,41	0,34	41,46
qu'	10,86	14,57	7,75	66,88
en	17,09	2,13	9,69	1,45
elle	0,88	85,16	137,07	372,83
une	13,97	40,02	0,09	64,27
me	10,01	13,38	22,38	27,93
a	0,59	49,61	19,38	115,81
qui	0,82	16,31	2,30	31,37
ce	13,43	4,89	1,59	11,62
n'	8,65	3,77	0,83	0,27
des	8,01	0,30	0,24	4,10
c'	0,24	32,11	5,62	11,16
dans	3,85	0,15	31,13	15,42
pour	11,80	10,01	0,62	20,49
était	11,91	12,47	36,15	118,49
m'	0,90	3,95	0,68	0,16
vous	162,12	2,00	0,64	46,73
plus	0,81	2,85	20,37	9,60
mais	0,19	6,58	0,06	7,96
ai	14,31	52,97	0,08	79,36
nous	170,28	3,67	15,32	53,19
avait	13,43	7,44	70,18	153,82
du	0,93	6,08	2,18	17,59
s'	0,23	2,39	1,55	7,98
on	2,10	40,67	4,91	7,67
se	0,21	0,30	0,06	0,00
mon	4,59	40,70	6,27	41,85
au	0,00	9,83	11,37	33,51
tout	0,00	11,07	4,31	1,36
moi	0,45	1,54	1,19	6,47

y	10,06	1,99	0,39	0,14
avec	25,55	1,94	44,29	1,29
lui	4,61	7,73	2,96	30,34
tu	11,68	138,87	64,82	34,10
sur	0,05	7,47	2,86	14,43
comme	5,33	18,49	2,62	47,57
son	2,39	3,48	4,57	1,65
même	13,82	0,29	0,79	5,17
si	6,07	3,62	6,67	1,28
ça	21,92	22,09	0,76	44,86
bien	1,79	0,86	0,66	1,08
dit	0,16	6,86	0,95	3,13
ils	20,92	30,81	0,78	79,93
fait	8,18	1,78	1,83	3,94
suis	0,02	29,07	6,92	6,80
χ^2	751,66	1.111,37	682,73	2.252,47

ddl: 177, à savoir (60-1)(4-1);

seuil de rejet d'après la correction de Bonferroni: 0,02 (à savoir, 5/240);

valeurs significatives: toute valeur supérieure à 131,1.

Appendice 11.17

Les Quatre Romans de Gary/Ajar: Valeurs calculées (abstraction faite des marqueurs de personne)

Type	<i>Clair</i>	<i>Tick</i>	<i>Câl</i>	<i>Vie</i>
de	1.051,750	1.846,030	1.764,270	2.181,220
et	841,093	1.476,289	1.410,902	1.744,335
à	630,275	1.106,261	1.057,263	1.307,122
que	499,927	877,473	838,609	1.036,794
un	463,845	814,142	778,082	961,964
pas	442,503	776,682	742,282	917,702
d'	330,572	580,221	554,522	685,570
ne	325,966	572,136	546,795	676,017
qu'	303,856	533,329	509,707	630,164
en	299,096	524,975	501,723	620,293
une	318,442	558,931	534,175	660,415
qui	291,726	512,039	489,360	605,009
ce	229,235	402,355	384,534	475,409

n'	249,042	437,120	417,759	516,486
des	299,403	525,514	502,238	620,930
c'	324,584	569,711	544,477	673,152
dans	250,117	439,006	419,562	518,715
pour	260,558	457,332	437,076	540,368
plus	184,248	323,393	309,070	382,111
mais	195,917	343,875	328,644	406,311
du	164,749	289,167	276,360	341,671
au	143,867	252,516	241,332	298,365
tout	154,615	271,381	259,361	320,655
y	189,776	333,095	318,342	393,574
avec	171,965	301,833	288,465	356,637
sur	112,545	197,539	188,790	233,406
comme	136,958	240,389	229,742	284,036
même	127,131	223,141	213,258	263,656
si	87,364	153,342	146,550	181,184
ça	194,689	341,719	326,583	403,764
bien	102,104	179,214	171,276	211,753

Appendice 11.18

Les Quatre Romans de Gary/Ajar: χ^2 (abstraction faite des marqueurs de personne)

Type	<i>Clair</i>	<i>Tick</i>	<i>Câl</i>	<i>Vie</i>
de	7,92	106,77	10,29	201,66
et	13,13	0,20	9,69	34,32
à	0,18	0,53	0,12	0,63
que	1,34	0,24	0,61	4,23
un	19,11	7,26	7,05	9,18
pas	3,71	21,65	3,27	21,76
d'	11,05	49,10	2,67	102,87
ne	26,55	1,26	9,43	3,13
qu'	9,19	18,50	8,99	78,09
en	15,05	0,84	8,23	0,43
une	16,53	32,64	0,28	56,65
qui	0,39	20,33	3,01	38,69
ce	15,58	2,98	2,12	8,73
n'	10,43	5,75	1,24	1,07
des	6,59	1,14	0,07	6,82

e'	0,04	37,78	4,50	15,71
dans	2,94	0,01	28,55	11,95
pour	10,20	13,08	0,33	26,04
plus	0,46	1,58	21,79	12,78
mais	0,04	8,76	0,18	10,95
du	1,41	4,20	1,69	14,62
au	0,03	7,49	10,22	29,85
tout	0,04	13,43	3,62	2,51
y	11,75	3,29	0,65	0,01
avec	23,79	3,15	41,59	0,52
sur	0,00	5,67	2,38	12,22
comme	4,55	21,20	3,11	54,10
même	12,67	0,77	0,54	7,13
si	6,95	2,52	7,23	0,69
ça	20,19	25,70	0,47	52,24
bien	1,43	1,47	0,44	1,94
χ^2	253,26	419,29	194,36	821,48

ddl: 90, à savoir $(31-1)(4-1)$;

seuil de rejet d'après la correction de Bonferroni: 0,04 (à savoir, $5/124$);

valeurs significatives: toute valeur supérieure à 124,3.

Appendices au chapitre 12

Appendice 12.1

Limites de confiance de r pour la corrélation entre *Clair de femme* et les autres oeuvres (lorsque $n=60$)

<i>Clair</i>	Corr	r_1	r_2
<i>Ticket</i>	0,9538	0,9722	0,9236
<i>Gros-Câlin</i>	0,9409	0,9643	0,9028
<i>La Vie</i>	0,8182	0,8875	0,7127
<i>L'Étranger</i>	0,7914	0,8702	0,6730
<i>Vipères</i>	0,8889	0,9323	0,8204
<i>L'Immoraliste</i>	0,8793	0,9263	0,8055
<i>La Porte étroite</i>	0,8341	0,8977	0,7365
<i>Engwall</i>	0,8887	0,9321	0,8201

Appendice 12.2

Limites de confiance de r pour la corrélation entre *Clair de femme* et les autres oeuvres (lorsque $n=31$)

<i>Clair</i>	Corr	r_1	r_2
<i>Ticket</i>	0,9763	0,9886	0,9513
<i>Gros-Câlin</i>	0,9504	0,9759	0,8993
<i>La Vie</i>	0,8430	0,9215	0,6983
<i>L'Étranger</i>	0,9283	0,9649	0,8561
<i>Vipères</i>	0,9246	0,9631	0,8490
<i>L'Immoraliste</i>	0,9192	0,9604	0,8386
<i>La Porte étroite</i>	0,8856	0,9434	0,7755
<i>Engwall</i>	0,9485	0,9749	0,8955

Appendice 12.3

Limites de confiance de r pour la corrélation entre *Au-delà de cette limite votre ticket n'est plus valable* et les autres oeuvres (lorsque $n=60$)

<i>Ticket</i>	Corr	r_1	r_2
<i>Clair</i>	0,9538	0,9722	0,9236
<i>Gros-Câlin</i>	0,9628	0,8718	0,6766
<i>La Vie</i>	0,7938	0,9776	0,6766
<i>L'Étranger</i>	0,8135	0,8845	0,7057
<i>Vipères</i>	0,9394	0,9634	0,9004
<i>L'Immoraliste</i>	0,9282	0,9566	0,8825
<i>La Porte étroite</i>	0,8783	0,9256	0,8040
<i>Engwall</i>	0,9352	0,9608	0,8936

Appendice 12.4

Limites de confiance de r pour la corrélation entre *Au-delà de cette limite votre ticket n'est plus valable* et les autres oeuvres (lorsque $n=31$)

<i>Ticket</i>	Corr	r_1	r_2
<i>Clair</i>	0,9763	0,9886	0,9513
<i>Gros-Câlin</i>	0,9738	0,9873	0,9463
<i>La Vie</i>	0,8281	0,9137	0,6720
<i>L'Étranger</i>	0,9481	0,9747	0,8948
<i>Vipères</i>	0,9329	0,9672	0,8650
<i>L'Immoraliste</i>	0,9458	0,9736	0,8902
<i>La Porte étroite</i>	0,8971	0,9493	0,7969
<i>Engwall</i>	0,9817	0,9912	0,9622

Appendice 12.5

Limites de confiance de r pour la corrélation entre *Gros-Câlin* et les autres oeuvres (lorsque $n=60$)

<i>Gros-Câlin</i>	Corr	r_1	r_2
<i>Clair</i>	0,9409	0,9643	0,9028
<i>Ticket</i>	0,9628	0,9776	0,9384
<i>La Vie</i>	0,8334	0,8972	0,7354
<i>L'Étranger</i>	0,8184	0,8876	0,7130
<i>Vipères</i>	0,8909	0,9335	0,8236
<i>L'Immoraliste</i>	0,9004	0,9394	0,8384
<i>La Porte étroite</i>	0,8351	0,8983	0,7380
<i>Engwall</i>	0,9041	0,9417	0,8443

Appendice 12.6

Limites de confiance de r pour la corrélation entre *Gros-Câlin* et les autres oeuvres (lorsque $n=31$)

<i>Gros-Câlin</i>	Corr	r_1	r_2
<i>Clair</i>	0,9504	0,9759	0,8993
<i>Ticket</i>	0,9738	0,9873	0,9463
<i>La Vie</i>	0,8696	0,9353	0,7461
<i>L'Étranger</i>	0,9420	0,9718	0,8829
<i>Vipères</i>	0,9109	0,9562	0,8228
<i>L'Immoraliste</i>	0,9183	0,9599	0,8369
<i>La Porte étroite</i>	0,8758	0,9385	0,7575
<i>Engwall</i>	0,9645	0,9828	0,9274

Appendice 12.7

Limites de confiance de r pour la corrélation entre *La Vie devant soi* et les autres oeuvres (lorsque $n=60$)

<i>La Vie</i>	Corr	r_1	r_2
<i>Clair</i>	0,8182	0,8875	0,7127
<i>Ticket</i>	0,7938	0,8718	0,6766
<i>Gros-Câlin</i>	0,8334	0,8972	0,7354
<i>L'Étranger</i>	0,8272	0,8933	0,7261
<i>Vipères</i>	0,7160	0,8205	0,5652
<i>L'Immoraliste</i>	0,7258	0,8270	0,5789
<i>La Porte étroite</i>	0,6810	0,7969	0,5168
<i>Engwall</i>	0,7711	0,8570	0,6435

Appendice 12.8

Limites de confiance de r pour la corrélation entre *La Vie devant soi* et les autres oeuvres (lorsque $n=31$)

<i>La Vie</i>	Corr	r_1	r_2
<i>Clair</i>	0,8430	0,9215	0,6985
<i>Ticket</i>	0,8281	0,9137	0,6720
<i>Gros-Câlin</i>	0,8696	0,9353	0,7461
<i>L'Étranger</i>	0,8873	0,9443	0,7786
<i>Vipères</i>	0,7309	0,8614	0,5099
<i>L'Immoraliste</i>	0,7953	0,8964	0,6157
<i>La Porte étroite</i>	0,7429	0,8680	0,5292
<i>Engwall</i>	0,7886	0,8928	0,6044

Appendice 12.9

Limites de confiance de r pour la corrélation entre *L'Étranger* et les autres oeuvres (lorsque $n=60$)

<i>L'Étranger</i>	Corr	r_1	r_2
<i>Clair</i>	0,7914	0,8702	0,6730
<i>Ticket</i>	0,8135	0,8845	0,7057
<i>Gros-Câlin</i>	0,8184	0,8876	0,7130
<i>La Vie</i>	0,8272	0,8933	0,7261
<i>Vipères</i>	0,7668	0,8541	0,6373
<i>L'Immoraliste</i>	0,7748	0,8594	0,6489
<i>La Porte étroite</i>	0,7251	0,8266	0,5780
<i>Engwall</i>	0,7910	0,8700	0,6726

Appendice 12.10

Limites de confiance de r pour la corrélation entre *L'Étranger* et les autres oeuvres (lorsque $n=31$)

<i>L'Étranger</i>	Corr	r_1	r_2
<i>Clair</i>	0,9283	0,9649	0,8561
<i>Ticket</i>	0,9481	0,9747	0,8948
<i>Gros-Câlin</i>	0,9420	0,9718	0,8829
<i>La Vie</i>	0,8873	0,9443	0,7786
<i>Vipères</i>	0,8922	0,9468	0,7877
<i>L'Immoraliste</i>	0,9429	0,9722	0,8845
<i>La Porte étroite</i>	0,9050	0,9533	0,8117
<i>Engwall</i>	0,9304	0,9704	0,8777

Appendice 12.11

Limites de confiance de r pour la corrélation entre *Le Noeud de vipères* et les autres oeuvres (lorsque $n=60$)

<i>Vipères</i>	Corr	r_1	r_2
<i>Clair</i>	0,8889	0,9323	0,8204
<i>Ticket</i>	0,9394	0,9634	0,9004
<i>Gros-Câlin</i>	0,8909	0,9335	0,8236
<i>La Vie</i>	0,7160	0,8205	0,5652
<i>L'Étranger</i>	0,7668	0,8541	0,6373
<i>L'Immoraliste</i>	0,9495	0,9696	0,9168
<i>La Porte étroite</i>	0,9432	0,9657	0,9065
<i>Engwall</i>	0,9125	0,9469	0,8574

Appendice 12.12

Limites de confiance de r pour la corrélation entre *Le Noeud de vipères* et les autres oeuvres (lorsque $n=31$)

<i>Vipères</i>	Corr	r_1	r_2
<i>Clair</i>	0,9246	0,9631	0,8490
<i>Ticket</i>	0,9329	0,9672	0,8650
<i>Gros-Câlin</i>	0,9109	0,9562	0,8228
<i>La Vie</i>	0,7309	0,8614	0,5099
<i>L'Étranger</i>	0,8922	0,9468	0,7877
<i>L'Immoraliste</i>	0,9650	0,9830	0,9284
<i>La Porte étroite</i>	0,9696	0,9853	0,9376
<i>Engwall</i>	0,9509	0,9761	0,9004

Appendice 12.13

Limites de confiance de r pour la corrélation entre *L'Immoraliste* et les autres oeuvres (lorsque $n=60$)

<i>L'Immoraliste</i>	Corr	r_1	r_2
<i>Clair</i>	0,8793	0,9263	0,8055
<i>Ticket</i>	0,9282	0,9566	0,8825
<i>Gros-Câlin</i>	0,9004	0,9394	0,8384
<i>La Vie</i>	0,7258	0,8270	0,5789
<i>L'Étranger</i>	0,7748	0,8594	0,6489
<i>Vipères</i>	0,9495	0,9696	0,9168
<i>La Porte étroite</i>	0,9558	0,9734	0,9270
<i>Engwall</i>	0,8703	0,9206	0,7915

Appendice 12.14

Limites de confiance de r pour la corrélation entre *L'Immoraliste* et les autres oeuvres (lorsque $n=31$)

<i>L'Immoraliste</i>	Corr	r_1	r_2
<i>Clair</i>	0,9192	0,9604	0,8386
<i>Ticket</i>	0,9458	0,9376	0,8902
<i>Gros-Câlin</i>	0,9183	0,9599	0,8369
<i>La Vie</i>	0,7953	0,8964	0,6157
<i>L'Étranger</i>	0,9429	0,9722	0,8845
<i>Vipères</i>	0,9650	0,9830	0,9284
<i>La Porte étroite</i>	0,9767	0,9887	0,9521
<i>Engwall</i>	0,9472	0,9743	0,8931

Appendice 12.15

Limites de confiance de r pour la corrélation entre *La Porte étroite* et les autres oeuvres (lorsque $n=60$)

<i>La Porte étroite</i>	Corr	r_1	r_2
<i>Clair</i>	0,8341	0,8977	0,7365
<i>Ticket</i>	0,8783	0,9256	0,8040
<i>Gros-Câlin</i>	0,8351	0,8983	0,7380
<i>La Vie</i>	0,6810	0,7969	0,5168
<i>L'Étranger</i>	0,7251	0,8266	0,5780
<i>Vipères</i>	0,9432	0,9657	0,9065
<i>L'Immoraliste</i>	0,9558	0,9734	0,9270
<i>Engwall</i>	0,8014	0,8767	0,6878

Appendice 12.16

Limites de confiance de r pour la corrélation entre *La Porte étroite* et les autres oeuvres (lorsque $n=31$)

<i>La Porte étroite</i>	Corr	r_1	r_2
<i>Clair</i>	0,8856	0,9434	0,7755
<i>Ticket</i>	0,8971	0,9493	0,7969
<i>Gros-Câlin</i>	0,8758	0,9385	0,7575
<i>La Vie</i>	0,7429	0,8680	0,5292
<i>L'Étranger</i>	0,9050	0,9533	0,8117
<i>Vipères</i>	0,9696	0,9853	0,9376
<i>L'Immoraliste</i>	0,9767	0,9887	0,9521
<i>Engwall</i>	0,8966	0,9490	0,7959

Appendice 12.17

Limites de confiance de r pour la corrélation entre Engwall et les autres oeuvres
(lorsque $n=60$)

Engwall	Corr	r_1	r_2
<i>Clair</i>	0,8887	0,9321	0,8201
<i>Ticket</i>	0,9352	0,9608	0,8936
<i>Gros-Câlin</i>	0,9041	0,9417	0,8443
<i>La Vie</i>	0,7711	0,8570	0,6435
<i>L'Étranger</i>	0,7910	0,8700	0,6726
<i>Vipères</i>	0,9125	0,9469	0,8574
<i>L'Immoraliste</i>	0,8703	0,9206	0,7915
<i>La Porte étroite</i>	0,8014	0,8767	0,6878

Appendice 12.18

Limites de confiance de r pour la corrélation entre Engwall et les autres oeuvres
(lorsque $n=31$)

Engwall	Corr	r_1	r_2
<i>Clair</i>	0,9485	0,9749	0,8955
<i>Ticket</i>	0,9817	0,9912	0,9622
<i>Gros-Câlin</i>	0,9645	0,9828	0,9274
<i>La Vie</i>	0,7886	0,8928	0,6044
<i>L'Étranger</i>	0,9304	0,9704	0,8777
<i>Vipères</i>	0,9509	0,9761	0,9004
<i>L'Immoraliste</i>	0,9472	0,9743	0,8931
<i>La Porte étroite</i>	0,8966	0,9490	0,7959

Appendices au chapitre 13

Appendice 13.1.1

Clair de femme (33.755 vocables)

<i>A</i>	<i>Fr. Abs.</i>	<i>Fr. Rel.</i>	<i>B</i>	<i>Fr. Abs.</i>	<i>Fr. Rel.</i>
Alors	44	1,30	Donc	18	0,53
Aucun(e)	19	0,56	Nul(le)	4	0,12
Ça	132	3,90	Cela	62	1,84
Commencer	10	0,29	Débuter	0	0
Commencement	1	0,03	Début	6	0,18
Croire	52	1,54	Penser	28	0,83
Devoir	62	1,84	Falloir	79	2,34
Demeurer	6	0,18	Habiter	5	0,15
Lorsque	35	1,04	Quand	45	1,33
Parce qu'(que)	57	1,69	Puisqu'(que)	3	0,09

Appendice 13.1.2

Au-delà de cette limite votre ticket n'est plus valable (57.915 vocables)

<i>A</i>	<i>Fr. Abs.</i>	<i>Fr. Rel.</i>	<i>B</i>	<i>Fr. Abs.</i>	<i>Fr. Rel.</i>
Alors	66	1,39	Donc	26	0,45
Aucun(e)	24	0,41	Nul(le)	01	0,02
Ça	249	4,30	Cela	34	0,59
Commencer	27	0,50	Débuter	0	0,00
Commencement	1	0,02	Début	13	0,22
Croire	62	1,07	Penser	43	0,74
Devoir	17	0,29	Falloir	01	0,02
Demeurer	118	2,03	Habiter	106	1,83
Lorsque	85	1,47	Quand	67	1,16
Parce qu'(que)	73	1,26	Puisqu'(que)	9	0,16

Appendice 13.1.3

Gros-Câlin (53.264 vocables)

<i>A</i>	<i>Fr. Abs.</i>	<i>Fr. Rel.</i>	<i>B</i>	<i>Fr. Abs.</i>	<i>Fr. Rel.</i>
Alors	64	1,20	Donc	50	0,94
Aucun(e)	30	0,56	Nul(le)	5	0,09
Ça	340	6,38	Cela	54	1,01
Commencer	29	0,54	Débuter	0	0,00
Commencement	3	0,06	Début	7	0,13
Croire	74	1,39	Penser	71	1,33
Devoir	13	0,24	Falloir	6	0,11
Demeurer	65	1,22	Habiter	113	2,12
Lorsque	87	1,63	Quand	76	1,42
Parce qu'(que)	86	1,61	Puisqu'(que)	4	0,08

Appendice 13.1.4

La Vie devant soi (65.095 vocables)

<i>A</i>	<i>Fr. Abs.</i>	<i>Fr. Rel.</i>	<i>B</i>	<i>Fr. Abs.</i>	<i>Fr. Rel.</i>
Alors	78	1,20	Donc	20	0,31
Aucun(e)	24	0,37	Nul(le)	9	0,14
Ça	549	8,43	Cela	6	0,09
Commencer	49	0,75	Débuter	0	0
Commencement	1	0,02	Début	9	0,14
Croire	88	1,35	Penser	99	1,52
Devoir	3	0,05	Falloir	5	0,08
Demeurer	86	1,32	Habiter	150	2,30
Lorsque	45	0,69	Quand	324	4,97
Parce qu'(que)	280	4,30	Puisqu'(que)	5	0,08

Appendice 13.1.5

L'Étranger (34.623 vocables)

<i>A</i>	<i>Fr. Abs.</i>	<i>Fr. Rel.</i>	<i>B</i>	<i>Fr. Abs.</i>	<i>Fr. Rel.</i>
Alors	107	3,09	Donc	10	0,29
Aucun(e)	14	0,40	Nul(le)	0	0,00
Ça	24	0,69	Cela	98	2,83
Commencer	5	0,46	Débuter	0	0,00
Commencement	0	0,00	Début	5	0,14
Croire	36	1,04	Penser	62	1,79
Demeurer	1	0,03	Habiter	1	0,03
Devoir	46	1,33	Falloir	51	1,47
Lorsque	10	0,29	Quand	77	2,22
Parce qu' (que)	63	1,81	Puisqu'(que)	9	0,26

Appendice 13.1.6

Le Noeud de Vipères (59.426 vocables)

<i>A</i>	<i>Fr. Abs.</i>	<i>Fr. Rel.</i>	<i>B</i>	<i>Fr. Abs.</i>	<i>Fr. Rel.</i>
Alors	27	0,45	Donc	17	0,29
Aucun(e)	65	1,09	Nul(le)	1	0,02
Ça	45	0,76	Cela	65	1,09
Commencer	18	0,30	Débuter	1	0,02
Commencement	3	0,05	Début	4	0,07
Croire	101	1,69	Penser	67	1,13
Demeurer	48	0,80	Habiter	10	0,17
Devoir	82	1,38	Falloir	123	2,07
Lorsque	56	0,94	Quand	77	1,30
Parce qu' (que)	43	0,72	Puisqu'(que)	14	0,24

Appendice 13.1.7

L'Immoraliste (41.260 vocables)

<i>A</i>	<i>Fr. Abs.</i>	<i>Fr. Rel.</i>	<i>B</i>	<i>Fr. Abs.</i>	<i>Fr. Rel.</i>
Alors	52	1,26	Donc	67	1,62
Aucun(e)	22	0,53	Nul(le)	7	0,17
Ça	0	0,00	Cela	59	1,42
Commencer	28	0,68	Débuter	0	0,00
Commencement	0	0,00	Début	4	0,09
Croire	67	1,62	Penser	52	1,26
Demeurer	12	0,29	Habiter	4	0,09
Devoir	78	1,89	Falloir	37	0,89
Lorsque	13	0,32	Quand	59	1,43
Parce qu' (que)	99	2,39	Puisqu'(que)	2	0,05

Appendice 13.1.8

La Porte étroite (39.953 vocables)

<i>A</i>	<i>Fr. Abs.</i>	<i>Fr. Rel.</i>	<i>B</i>	<i>Fr. Abs.</i>	<i>Fr. Rel.</i>
Alors	36	0,90	Donc	36	0,90
Aucun(e)	14	0,35	Nul(le)	4	0,10
Ça	4	0,10	Cela	67	1,17
Commencer	5	0,46	Débuter	0	0,00
Commencement	1	0,03	Début	01	0,03
Croire	61	1,53	Penser	45	1,13
Demeurer	9	0,23	Habiter	9	0,23
Devoir	62	1,55	Falloir	34	0,85
Lorsque	25	0,63	Quand	56	51,40
Parce qu' (que)	15	0,38	Puisqu'(que)	8	0,20

Appendice 13.2.1

Les Neuf Textes: Valeurs observées et table de contingence

Lemmes	Clair	Tick	Câl	Vie	Étran	Vip	Imm	Porte	Eng	
Alors	44	66	64	78	107	27	52	36	578	1.052
Donc	18	26	50	20	10	17	67	36	279	523
Ça	132	249	340	549	24	45	0	4	1084	2.427
Cela	62	34	54	6	98	65	59	67	584	1.029
Croire	52	62	74	88	36	101	67	61	595	1.136
Penser	28	43	71	99	62	67	52	45	577	1.044
Devoir	62	118	65	86	46	82	78	62	782	1.381
Falloir	79	106	113	150	51	123	37	34	712	1.405
Lorsque	35	85	87	45	10	56	13	25	160	516
Quand	45	67	76	324	77	77	59	56	953	1.734
Parce qu' (que)	57	73	86	280	63	43	9	15	379	1.005
Puisqu' (que)	3	9	4	5	9	14	2	8	112	166
	617	938	1.084	1.730	593	717	495	449	6.795	13.418

Appendice 13.2.2

Les Neuf Textes: Valeurs calculées (aucune valeur inférieure à 5)

Lemmes	Clair	Tick	Câl	Vie	Étran	Vip	Imm	Porte	Eng
Alors	48,37	73,54	84,99	135,64	46,49	56,21	38,81	35,20	532,74
Donc	24,05	36,56	42,25	67,43	23,11	27,95	19,29	17,50	264,85
Ça	111,60	169,66	196,07	312,92	107,26	129,69	89,53	81,21	1.229,06
Cela	47,32	71,93	83,13	132,67	45,48	54,99	37,96	34,43	521,10
Croire	52,24	79,41	91,77	146,47	50,20	60,70	41,91	38,01	575,28
Penser	48,01	72,98	84,34	134,60	46,14	55,79	38,51	34,93	528,69
Devoir	63,50	96,54	111,57	178,05	61,03	73,79	50,95	46,21	699,35
Falloir	64,61	98,22	113,51	181,15	62,09	75,08	51,83	47,01	711,51
Lorsque	23,73	36,07	41,69	66,53	22,80	27,57	19,04	17,27	261,31
Quand	79,73	121,22	140,08	223,57	76,63	92,66	63,97	58,02	878,11
Parce qu'(que)	46,21	70,26	81,19	129,58	44,42	53,70	37,08	33,63	508,94
Puisqu'(que)	7,63	11,60	13,41	21,40	7,34	8,87	6,12	5,55	84,06

Les lemmes *aucun(e)/nul(le)*, *commencer/débuter*, *commencement/début*, *demeurer/habiter* sont exclus de l'appendice 13.2.2, étant donné que leurs valeurs calculées sont inférieures à 5.

Appendice 13.2.3

Les Neuf Textes: Valeurs du χ^2

Lemmes	Clair	Tick	Câl	Vie	Étran	Vip	Imm	Porte	Eng
Alors	0,40	0,77	5,18	24,49	78,75	15,18	4,48	0,02	3,84
donc	1,52	3,05	1,42	33,36	7,44	4,29	117,96	19,55	0,76
Ça	3,73	37,10	105,66	178,12	64,63	55,30	89,53	73,41	17,12
Cela	4,56	20,00	10,21	120,94	60,66	1,82	11,66	30,80	7,59
Croire	0,00	3,82	3,44	23,34	4,02	26,75	15,02	13,90	0,68
Penser	8,34	12,32	2,11	9,42	5,45	2,25	4,72	2,90	4,41
Devoir	0,04	4,77	19,44	47,59	3,70	0,91	14,37	5,39	9,77
Falloir	3,21	0,62	0,00	5,36	1,98	30,59	4,24	3,60	0,00
Lorsque	5,36	66,37	49,26	6,97	7,19	29,31	1,91	3,46	39,28
Quand	15,13	24,25	29,32	45,12	0,00	2,65	0,39	0,07	6,39
Parce qu'(que)	2,52	0,11	0,28	174,63	7,78	2,13	21,26	10,32	33,18
Puisqu' (que)	2,81	0,58	6,60	12,57	0,38	2,97	2,78	1,08	9,28
χ^2	47,60	173,76	232,92	681,90	241,98	174,16	288,33	164,50	132,30

ddl: 88, à savoir (12-1)(9-1);

seuil de rejet d'après la correction de Bonferroni: 0,046 (à savoir, 5/108);

valeurs significatives: toute valeur supérieure à 124,3..

Appendice 13.3.1

Les Huit Romans: Valeurs observées et table de contingence

Lemmes	<i>Clair</i>	<i>Tick</i>	<i>Câl</i>	<i>Vie</i>	<i>Étran</i>	<i>Vip</i>	<i>Imm</i>	<i>Porte</i>	
Alors	44	66	64	78	107	27	52	36	474
Donc	18	26	50	20	10	17	67	36	244
Ça	132	249	340	549	24	45	0	4	1343
Cela	62	34	54	6	98	65	59	67	445
Croire	52	62	74	88	36	101	67	61	541
Penser	28	43	71	99	62	67	52	45	467
Devoir	62	118	65	86	46	82	78	62	599
Falloir	79	106	113	150	51	123	37	34	693
Lorsque	35	85	87	45	10	56	13	25	356
Quand	45	67	76	324	77	77	59	56	781
	557	856	994	1.445	521	660	484	426	5.943

Appendice 13.3.2

Les Huit Romans: Valeurs calculées

Lemmes	<i>Clair</i>	<i>Tick</i>	<i>Câl</i>	<i>Vie</i>	<i>Étran</i>	<i>Vip</i>	<i>Imm</i>	<i>Porte</i>
Alors	44,43	68,27	79,28	115,25	41,55	52,64	38,60	33,98
Donc	22,87	35,14	40,81	59,33	21,39	27,10	19,87	17,49
Ça	125,87	193,44	224,62	326,54	117,74	149,15	109,37	96,27
Cela	41,71	64,10	74,43	108,20	39,01	49,42	36,24	31,90
Croire	50,70	77,92	90,49	131,54	47,43	60,08	44,06	38,78
Penser	43,77	67,26	78,11	113,55	40,94	51,86	38,03	33,48
Devoir	56,14	86,28	100,19	145,64	52,51	66,52	48,78	42,94
Falloir	64,95	99,82	115,91	168,50	60,75	76,96	56,44	49,67
Lorsque	33,37	51,28	59,54	86,56	31,21	39,54	28,99	25,52
Quand	73,20	112,49	130,63	189,89	68,47	86,73	63,60	55,98

Les lemmes *aucun(e)/nul(le)*, *commencer/débiter*, *commencement/début*, *demeurer/habiter*, *Parce qu'(que)/puisque'(que)* sont exclus, étant donné que leur valeurs calculées sont inférieures à 5.

Appendice 13.3.3

Les Huit Romans: Valeurs du χ^2

Lemmes	<i>Clair</i>	<i>Tick</i>	<i>Câl</i>	<i>Vie</i>	<i>Étran</i>	<i>Vip</i>	<i>Imm</i>	<i>Porte</i>
Alors	0,00	0,08	2,94	12,04	103,08	12,49	4,65	0,12
Donc	1,04	2,38	2,07	26,07	6,07	3,76	111,77	19,59
Ça	0,30	15,96	59,26	151,55	74,63	72,72	109,37	88,43
Cela	9,87	14,13	5,61	96,53	89,20	4,91	14,29	38,63
Croire	0,03	3,25	3,00	14,41	2,75	27,87	11,94	12,73
Penser	5,68	8,75	0,65	1,86	10,83	4,42	5,13	3,97
Devoir	0,61	11,66	12,36	24,42	0,81	3,60	17,50	8,46
Falloir	3,04	0,38	0,07	2,03	1,57	27,54	6,69	4,95
Lorsque	0,08	22,18	12,66	19,95	14,41	6,86	8,82	0,01
Quand	10,86	18,40	22,84	94,71	1,06	1,09	0,33	0,00
χ^2	31,52	97,17	121,47	443,58	304,40	165,27	290,51	176,89

ddl: 63, à savoir (10-1)(8-1);

seuil de rejet d'après la correction de Bonferroni: 0,06 (à savoir, 5/80);

valeurs significatives: toute valeur supérieure à 83,3.

Appendice 13.4.1

Porte et Immoraliste: Valeurs observées et table de contingence

Lemmes	<i>Imm</i>	<i>Porte</i>	
Alors	52	36	88
Donc	67	36	103
Aucun(e)	22	14	36
Nul(le)	7	4	11
Croire	67	61	128
Penser	52	45	97
Demeurer	12	9	21
Habiter	4	9	13
Devoir	78	62	140
Falloir	37	34	71
Lorsque	13	25	38
Quand	59	56	115
	470	391	861

Appendice 13.4.2

Porte et Immoraliste: Valeurs calculées

Lemmes	<i>Imm</i>	<i>Porte</i>	Lemmes	<i>Imm</i>	<i>Porte</i>
Alors	48,04	39,96	Donc	56,23	46,77
Aucun(e)	19,65	16,35	Nul(le)	6,00	5,00
Croire	69,87	58,13	Penser	52,95	44,05
Demeurer	11,46	9,54	Habiter	7,10	5,90
Devoir	76,42	63,58	Falloir	38,76	32,24
Lorsque	20,74	17,26	Quand	62,78	52,22

Les lemmes *ça/cela*, *commencer/débuter*, *commencement/début*, *parce que/puisque* sont exclus, étant donné que leurs valeurs calculées sont au-dessous de 5.

Appendice 13.4.3

Porte et Immoraliste: Valeurs du χ^2

Lemmes	<i>Imm</i>	<i>Porte</i>	Lemmes	<i>Imm</i>	<i>Porte</i>
Alors	0,33	0,39	Donc	2,06	2,48
Aucun(e)	0,28	0,34	Nul(le)	0,16	0,20
Demeurer	0,03	0,03	Habiter	1,35	1,62
Croire	0,12	0,14	Penser	0,02	0,02
Devoir	0,03	0,04	Falloir	0,08	0,10
Lorsque	2,89	3,47	Quand	0,23	0,27
χ^2	<i>Imm</i>	7,58	<i>Porte</i>	9,11	

ddl: 20, à savoir (6-1)(5-1);

seuil de rejet d'après la correction de Bonferroni: 0,167 (à savoir, 5/30);

valeurs significatives: toute valeur supérieure à 26,5.

Appendice 13.5.1

Clair et Ticket : Valeurs observées et table de contingence

Lemmes	<i>Clair</i>	<i>Ticket</i>	
Alors	44	66	110
Donc	18	26	44
Ça	132	249	381
Cela	62	34	96
Croire	52	62	114
Penser	28	43	71
Devoir	62	118	180
Falloir	79	106	185
Lorsque	35	85	120
Quand	45	67	112
	557	856	1.413

Appendice 13.5.2

Clair et Ticket: Valeurs calculées

Lemmes	Clair	Tick	Lemmes	Clair	Tick
Alors	43,36	66,64	Donc	17,34	26,66
Ça	150,19	230,81	Cela	37,84	58,16
Croire	44,94	69,06	Penser	27,99	43,01
Devoir	70,96	109,04	Falloir	72,93	112,07
Lorsque	47,30	72,70	Quand	44,15	67,85

Les lemmes *aucun(e)/nul(le)*, *commencer/débuter*, *commencement/début*, *demeurer/habiter*, *parce que/puisque* sont exclus étant donné que leurs valeurs calculées sont au-dessous de 5.

Appendice 13.5.3

Clair et Ticket: Valeurs du χ^2

Lemmes	Clair	Ticket	Lemmes	Clair	Ticket
Alors	0,01	0,01	Donc	0,02	0,02
Ça	2,20	1,43	Cela	15,42	10,03
Croire	1,11	0,72	Penser	0,00	0,00
Devoir	1,13	0,74	Falloir	0,51	0,33
Lorsque	3,20	2,08	Quand	0,02	0,01
χ^2	Clair 23,62		Ticket 15,37		

ddl: 12, à savoir (5-1)(4-1);

seuil de rejet d'après la correction de Bonferroni: 0,25 (à savoir, 5/20);

valeurs significatives: toute valeur supérieure à 25,04.

Appendice 13.6.1

Câlin et Vie: Valeurs observées et table de Contingence

Lemmes	<i>Câl</i>	<i>Vie</i>	
Alors	64	78	142
Donc	50	20	70
Aucun(e)	30	24	54
Nul(le)	5	9	14
Ça	340	549	889
Cela	54	6	60
Croire	74	88	162
Penser	71	99	170
Demeurer	13	3	16
Habiter	6	5	11
Devoir	65	86	151
Falloir	113	150	263
Lorsque	87	45	132
Quand	76	324	400
Parce qu'(que)	86	280	366
Puisqu' (que)	4	5	9
	1.138	1.771	2.909

Appendice 13.6.2

Câlin et Vie: Valeurs calculées

Lemmes	<i>Câl</i>	<i>Vie</i>	Lemmes	<i>Câl</i>	<i>Vie</i>
Alors	55,55	57,16	Donc	27,38	28,18
Aucun(e)	21,12	21,74	Nul(le)	5,48	5,64
Ça	347,78	357,86	Cela	23,47	24,15
Croire	63,37	65,21	Penser	66,50	68,43
Devoir	59,07	60,78	Falloir	102,89	105,87
Lorsque	51,64	53,14	Quand	156,48	161,02

Les lemmes *commencer/débuter*, *commencement/début*, *demeurer/habiter*, *parce que/puisque* sont exclus, étant donné que leurs valeurs calculées sont inférieures à 5.

Appendice 13.6.3

Câl et *Vie*: Valeurs du χ^2

Lemmes	<i>Câl</i>	<i>Vie</i>	Lemmes	<i>Câl</i>	<i>Vie</i>
Alors	1,29	7,60	Donc	18,68	2,37
Aucun(e)	3,73	0,24	Nul(le)	0,04	2,01
Ça	0,17	102,09	Cela	39,71	13,64
Croire	1,78	7,96	Penser	0,30	13,65
Devoir	0,60	10,46	Falloir	0,99	18,40
Lorsque	24,22	1,25	Quand	41,39	164,97
χ^2	<i>Câl</i> 139,90	<i>Vie</i> 344,64			

ddl: 15, à savoir (6-1)(4-1);

seuil de rejet d'après la correction de Bonferroni: 0,167 (à savoir, 5/30);

valeurs significatives: toute valeur supérieure à 20,6.

BIBLIOGRAPHIE

I. OEUVRES DE ROMAIN GARY

Gary, Romain. Adieu Gary Cooper. Paris: Gallimard, 1969.

---. Au-delà de cette limite votre ticket n'est plus valable. Paris: Gallimard, 1975.

---. La Bonne Moitié. Comédie en deux actes. Paris: Gallimard, 1979.

---. Les Cerfs-volants. Paris: Gallimard, 1980.

---. Charge d'âme. Paris: Gallimard, 1978.

---. Chien Blanc. Paris: Gallimard, 1970.

---. Clair de femme. Paris: Gallimard, 1977.

---. Les Clowns lyriques. Paris: Gallimard, 1979.

---. Les Couleurs du jour. Paris: Gallimard, 1952.

---. La Danse de Gengis Cohn. Paris: Gallimard, 1967.

---. Éducation européenne. Paris: Gallimard, 1945.

---. Les Enchanteurs. Paris: Gallimard, 1973.

---. Europa. Paris: Gallimard, 1972.

- . Gloire à nos illustres pionniers. Paris: Gallimard, 1962.
- . Le Grand Vestiaire. Paris: Gallimard, 1949.
- . L'Homme à la colombe [Sinibaldi, Fosco]. Paris: Gallimard, 1958.
- . Lady L. Paris: Gallimard, 1963.
- . Les Mangeurs d'étoiles. Paris: Gallimard, 1966.
- . La Nuit sera calme. Paris: Gallimard, 1974.
- . La Promesse de l'aube. Paris: Gallimard, 1960.
- . Les Racines du ciel. Paris: Gallimard, 1956.
- . La Tête coupable. Paris: Gallimard, 1968.
- . Les Têtes de Stéphanie [Bogat, Shatan]. Paris: Gallimard, 1974.
- . Les Trésors de la Mer Rouge. Paris: Gallimard, 1971.
- . Tulipe. Paris: Gallimard, 1946.
- . Vie et mort d'Émile Ajar. Paris: Gallimard, 1981.

II. SOUS LE PSEUDONYME D'ÉMILE AJAR

- Ajar, Émile. L'Angoisse du Roi Salomon. Paris: Mercure de France, 1979.
- . Gros-Câlin. Paris: Mercure de France, 1974.
 - . Pseudo. Paris: Mercure de France, 1976.
 - . La Vie devant soi. Paris: Mercure de France, 1975.

---. La Vie devant soi. Paris: Mercure de France, 1975.

III. FILMOGRAPHIE

Gary, Romain. Kill avec Jean Seberg, James Mason, Stephen Boyd et Curt Jurgens.

---. Les Oiseaux vont mourir au Pérou avec Jean Seberg, Maurice Ronet, Danielle

Darieux et Pierre Brasseur.

IV. ROMANS DE GARY/AJAR ADAPTÉS À L'ÉCRAN

Au-delà de cette limite votre ticket n'est plus valable, de George Kaczender, avec Richard

Harris, Jeanne Moreau et George Peppard, 1982.

Clair de femme, de Costa-Gavras, avec Yves Montand, Romy Schneider et Lila Kedrova,

1979.

Dressé pour tuer, de Samuel Fuller, d'après Chien Blanc, avec Kristy Mc Nichol, Paul

Winfield et Burl Yves, 1982.

Gros-Câlin, de Jean-Pierre Rawson, adaptation et dialogue de Age et Scarpelli, avec Jean

Carnet, et Nino Manfredi, 1979.

La Promesse de l'aube, de Jules Dassin, pour la MGM avec Melina Mercouri, Assaf

Dayan et Didier Haudepin, 1970.

La Vie devant soi, de Moshe Mizrahi, avec Simone Signoret et Claude Dauphin, 1977.

Lady L, de Peter Ustinov produit par Carlo Ponti pour la MGM, avec Sophia Loren, Paul Newman, David Niven, Claude Dauphin et Michel Piccoli, 1965.

Les Cerfs-Volants, téléfilm de Pierre Badel en quatre épisodes, avec Jacques Penot, Anne Gauthier, Paul Crauchet, Jean-Marc Thibault et Rosy Varte.

Les Racines du ciel, de John Huston, produit par Darryl Zanuck pour la 20th Century Fox, avec Trevor Howard, Errol Flynn, Juliette Greco et Orson Welles, 1958.

V. OUVRAGES CITÉS

ARTFL (1997). American and French Research on the Treasury of The French Language,

Chicago: U. of Chicago. URL: <http://humanities.uchicago.edu/ARTFL/ARTFL.html>.

Baby, Yvonne. "Rencontre avec l'auteur de La Vie devant soi". Le Monde 150 (10 octobre 1975): 20.

Bally, Charles. Traité de stylistique française. 2e édition. Paris: Klincksieck, 1937.

Barthes, Roland. Le Bruissement de la langue. Paris: Seuil, 1993.

Bayard, Pierre. Il était deux fois Romain Gary. Paris: Presses Universitaires de France, 1990.

Boisen, Jørn. Un Picaro métaphysique, Romain Gary et l'art du roman. Thèse de doctorat, Université d'Odense, 1993.

Bona, Dominique. Romain Gary. Paris/Saint Laurent: Mercure de France/Lacombe, 1987.

Brainerd, Barron. "On the Distinction Between a Novel and a Romance: A Discriminant Analysis". Computers and the Humanities 7 (1973): 259-70.

—. Weighing Evidence in Language and Literature: A Statistical Approach. Toronto: University of Toronto Press, 1974.

Brinegar, Claude. "Mark Twain and the *Quintus Curtius Snodgrass Letters*: A Statistical Test of Authorship". Journal of American Statistical Association 57 (1963): 85-96.

Brooke, Tucker C.F. The Authorship of the Second and Third Parts of King Henry VI. New Haven: Yale University Press, 1912.

Brunet, Étienne. Le Vocabulaire de Jean Giraudoux, structure et évolution. Genève: Slatkine, 1978.

Burrows, John F. "'An Ocean Where Each Kind...': Statistical Analysis and some Major Determinants of Literary Style". Computers and the Humanities 23 (1989): 309-21.

—. Computation into Criticism: A Study of Jane Austen's Novels and an Experiment in Method. Oxford: Clarendon Press, 1987.

—. "Modal Verbs and Moral Principles: An Aspect of Jane Austen's Style". Literary and Linguistic Computing 1.1 (1986): 8-23.

- Camus, Albert. L'Étranger. Paris: Gallimard, 1942.
- Caradec, François. La Farce et le sacré: Fêtes et farceurs, Mythes et Mystificateurs. Paris: Casterman, 1977.
- Catonné, Jean-Marie. Romain Gary/Émile Ajar. Paris: Les Dossiers Belfond, 1990.
- Cohen, Jean. Structure du langage poétique. Paris: Flammarion, 1966.
- Corneille, Pierre. "Le Cid" Théâtre complet. Paris: Presses universitaires de France, 1984.
- Dellofre, François. Stylistique et poétique françaises. Paris: SEDES, 1970.
- Desanti, Dominique. "Masquer son nom". Corps Écrit 8 (1983): 91-8.
- Dugast, Daniel. La statistique lexicale. Genève: Slatkine, 1980.
- Ellegård, Alvar. A Statistical Method for Determining Authorship: The Junius Letters 1769 -1772. Gothenburg: Gothenburg Press, 1962.
- Engwall, Gunnel. Fréquence et distribution du vocabulaire dans un choix de romans français. Stockholm: Scriptor, 1974.
- Fortier, Dominique. Une Étude stylistique des romans d'Émile Ajar. Mémoire de maîtrise, Université McGill, 1998.
- Francis, Ivor S. "An Exposition of a Statistical Approach to the Federalist Dispute". The Computer and Literary Style. Ed. Jacob Leed. Kent, Ohio: Kent State University Press, 1966: 38-78.
- Fucks, William. "On the Mathematical Analysis of Style". Biometrika 39 (1952): 122-9.

- Galdi, László. Précis de Stylistique française. Budapest: Tankönyvkiadó, 1969.
- Gallagher, Guy. L'Univers du double reflet: Étude thématique de l'oeuvre de Gary. Thèse de doctorat. Québec: Université Laval, 1979.
- Genette, Gerard. Figures II. Paris: Seuil, 1968.
- Gide, André. Journal 1939-1949 Souvenirs. Paris, Gallimard, 1954.
- . L'Immoraliste. Paris: Mercure de France, 1902.
- . La Porte Étroite. Paris: Mercure de France, 1909.
- Gougenheim, Georges. L'Élaboration de français élémentaire. Étude sur l'établissement d'un vocabulaire et d'une grammaire de base. Paris: Didier, 1956.
- Guiraud, Pierre et Joshua Whatmouth. Bibliographie critique de la statistique linguistique. Utrecht-Anvers: Mouton, 1954.
- Guyau, Jean-Marie. L'Art du point de vue sociologique. Paris: F. Alcan, 1930.
- Hart, Adam. Shakespeare and the Vocabulary of *The Two Noble Kinsmen*. Melbourne: Melbourne University Press, 1934.
- Henmon, Vivian A. C. A French Word Book Based on a Count of 400,000 Running Words. Madison Wisconsin, 1924.
- Herdan, Gustav. Quantitative Linguistics. London: Butterworths, 1964.
- Hoffman, Calvin. The Murder of the Man Who Was Shakespeare. London: Max Parrish, 1955.

Holmes, David. I. "Authorship Attribution". Computers and the Humanities 28 (1994): 87-106.

Horton, Thomas B. The Effectiveness of the Stylometry of Function Words in Discriminating between Shakespeare and Fletcher. Edinburgh: University of Edinburgh Press, 1987.

Imbs, Paul et al. Trésor de la langue française du XIXe et du XXe siècles (1789-1960). Paris: Éditions du Centre national de la recherche scientifique, 1971-9.

Ionesco, Eugène. "Tueur sans gages". Théâtre Complet. Ed. Emmanuel Jacquot. Paris: Gallimard, 1991.

Jeandillou, Jean-François. Supercherries Littéraires, La Vie et l'Oeuvre des auteurs supposés. Florence: Usher, 1989.

Juilland, Alphonse. Frequency Dictionary of French Words. The Hague/Paris: Mouton, 1970.

—. Frequency Dictionary of Italian Words. The Hague: Mouton, 1973.

—. Frequency Dictionary of Spanish Words. The Hague: Mouton, 1964.

—. Frequency Dictionary of Rumanian Words. The Hague: Mouton, 1965.

Kjetsaa, Geir. "And Quiet Flows the Don Through the Computer". Association for Literary and Linguistic Computing Bulletin 7 (1979): 248-56.

Laclos, Pierre. Les Liaisons dangereuses 1782. Ed. Maurice Allem. Paris: Garnier, 1959.

Lacroix, Paul. Histoire des mystifications et des mystifiés. Bruxelles et Liepzig: A.

Schnée, 1856-1858.

Lafond, Hellène. L'univers romanesque d'Émile Ajar ou le refus de la norme. Mémoire de maîtrise, Université McGill, 1991.

Lane, Philippe. La périphérie du texte. Paris: Nathan, 1992.

Ledger, Gerard. Recounting Plato: A Complete Analysis of Plato's Style. Oxford: Oxford University Press, 1989.

Ledger, Gerard et Thomas Merriam. "Shakespeare, Fletcher, and Two Noble Kinsmen". Literary and Linguistic Computing 9.3 (1994): 234-48.

Legris, Georges. Statistique Vol 1. Statistique descriptive. Paris: Economica, 1975.

Lindner, Jean-Guillaume Sigismond. Dictionnaire des écrivains pseudonymes allemands. Leipzig, 1830.

Lord, Robert D. "Studies in the History of Probability and Statistics VIII: De Morgan and the Statistical Study of Literary Style". Biometrika 45 (1958): 250-82.

Maingueneau, Dominique. Pragmatique pour un discours littéraire. Paris: Bordas. 1990.

Marouzeau, Jules. Précis de stylistique française. Paris: Masson, 1941.

Matthews, Robert A. J. et Thomas V. N. Merriam. "Neural Computation in Stylometry 1: An Application to the Works of Shakespeare and Fletcher". Literary and Linguistic Computing 8.4 (1993): 202-9.

- Mauriac, François. Le Noeud de Vipères. Paris: Grasset, 1932.
- Maxwell, James Coutts (ed.) King Henry VIII. Cambridge: Cambridge University Press, 1962.
- McKee, Rebecca J. The Humanism of Romain Gary. Mémoire de maîtrise, Trinity College, Dublin 1978.
- McLeman, James et Andrew Q. Morton. Paul, the Man and the Myth: A Study in the Authorship of Greek Prose. London: Hodder and Stoughton, 1966.
- Mendenhall, Thomas C. "A Mechanical Solution of a Literary Problem". The Popular Science Monthly 60 (1901): 97-105.
- . "The Characteristic Curves of Composition". Science 9.214 Supplement 11. March 1887: 237-49.
- Merriam, Thomas V.M. Modelling a Canon: A Stylometric Examination of Shakespeare's First Folio. Thèse de doctorat, University of London, 1992.
- Merriam, Thomas V. M. et Robert A. Matthews. "Neural Computation in Stylometry II: An Application to the Works of Shakespeare and Marlowe". Literary and Linguistic Computing 9.1 (1994): 1-6.
- Milic, Louis T. A Quantitative Approach to the Style of Jonathan Swift. The Hague: Mouton, 1967.
- Mitterrand, Henri. Dictionnaire des oeuvres du XXe siècle: littérature française et francophone. Paris: Dictionnaires Le Robert, 1995.

Molière. "L'Avare" Oeuvres complètes. Ed. Gabriel D. Bonno 1898. Paris: Larousse, 1941.

---. "L'École des Femmes" Oeuvres complètes. Ed. Gabriel D. Bonno 1898. Paris: Larousse, 1941.

Morier, Henri. La psychologie des styles. Genève: Georg, 1959.

Morton, Andrew Q. "The Authorship of the Pauline Corpus". The New Testament in Historical and Contemporary Perspective. Oxford: Blackwell, 1965.

---. Literary Detection. New York: Scribner's, 1978.

---. "The Authorship of Greek Prose". Journal of the Royal Statistical Society, Series A, 128 (1965), 169-224.

Mosteller, Frederick et David Wallace. Applied Bayesian and Classical Inference: The Case of the Federalist Papers. Reading, Mass.: Addison-Wesley, 1984.

---. Inference and Disputed Authorship: The Federalist. Reading, Mass.: Addison-Wesley, 1964.

Muller, Charles. Initiation aux méthodes de la statistique linguistique. Paris: Hachette, 1973.

---. Étude de statistique lexicale. Le Vocabulaire du théâtre de Pierre Corneille. Paris: Larousse, 1967.

---. Principes et Méthodes de statistique lexicale. Paris: Hachette, 1977.

- Nøjgaard, Morten. "La Mort d'un artiste". Moderna Sprach 82. 1 (1990): 31-6.
- Östman, Anne-Charlotte. L'Utopie et l'ironie, Étude sur Gros-Câlin et sa place dans l'oeuvre de Romain Gary. Thèse de doctorat, Université de Stockholm. Stockholm: Almqvist et Wiksell, 1994.
- Outzen, Vagn. "Romain Gary - Portrait d'un humaniste en saltimbanque". Actes du Xe. Congrès des Romanistes scandinaves, Études romanes. Lund: Lund University Press, 1990: 56-60.
- Pavlowitch, Paul. L'Homme que l'on croyait. Paris: Fayard, 1981.
- Pawlowski, Adam. Séries Temporelles en linguistique avec application à l'attribution de textes: Romain Gary et Émile Ajar. Paris: Champion, 1998.
- Piatier, Jacqueline. Le Monde 134, 27 septembre 1974: 10.
- Picard, Roger. Artifices et mystifications littéraires. Montréal: Variétés, 1945.
- Poirot-Delpech, Bertrand. "Ajar alias Gary". Le Monde (10 juillet 1981): 2.
- Proudfoot, G. R. (ed.). The Two Noble Kinsmen. London: Edward Arnold, 1970.
- Proust, Marcel. Contre Sainte-Beuve. Paris: Gallimard 1971.
- Quérard, Jean-Marie. Les Supercheries littéraires dévoilées: Galerie des auteurs apocryphes, supposés, déguisés, plagiaires, et des éditeurs infidèles de la littérature française 1847-1853. Paris: Maisonneuve et Larose, 1864.
- Revon, Maxime (ed.) Le Théâtre de Clara Gazul, comédienne espagnole. Paris: Garnier Frères, 1928.

Rinaldo, Angelo. Paris-Match 9 octobre 1987: 4.

Robert, Paul. Dictionnaire alphabétique et analogique de la langue française. Paris: société du Nouveau Littré, Le Robert, 1981.

Royer, Jean. "Romain Gary accuse". Le Devoir 24 octobre 1980: 23.

Sainte-Beuve, Charles-Augustin. Critiques et Portraits littéraires (1852). Paris: Garnier, 1924.

Sartre, Jean-Paul. Situations II. Paris: Gallimard, 1947.

Saussure Ferdinand de. Cours de Linguistique générale. Ed. Charles Bally. Paris: Payot, 1964-65.

Sherman, Lucius A. "Some Obvervations upon the sentence-length in English prose". University [of Nebraska] Studies I (1888): 119-30.

Shoenbaum, S. (ed.). The Famous History of the Life of King Henry the Eighth. New York: New American Library, 1967.

Smith, Michael William. "An Investigation of Morton's Method to Distinguish Elizabethan Playwrights". Computers and the Humanities 19.1 (1984): 3-21.

---. "A Stylometric Analysis of *Hero and Leander*". The Bard 3 (1982): 105-32.

---. "Recent Experience and New Developments of Methods for the Determination of Authorship". Association for Literary and Linguistic Computing Bulletin 11 (1983): 73-82.

Somers, H. H. "Statistical Methods in Literary Analysis". Computer and Literary Style.

Ed. J. Leed. Kent, Ohio: Kent State University Press, 1966: 45-49

Souriau, Étienne. La Correspondance des arts, éléments esthétiques comparés. Paris:

Flammarion, 1947.

Spitzer, Leo. Études de style. Paris: Gallimard, 1970.

Starobinski, Jean. L'oeil vivant. Paris: Gallimard, 1961.

Tadié, Jean-Yves. La Critique littéraire au XXe siècle. Paris: Belfond, 1987.

Taylor, Gary. The Canon and Chronology of Shakespeare's Plays. William Shakespeare:

A Textual Companion. Oxford: Clarendon Press, 1987.

Thorndike, Edward L. The Teacher's Word Book. New York: Columbia University,

1921.

Trahard, Pierre. La jeunesse de Prosper Mérimée 1803-1834 : Essai sur sa formation

intellectuelle et sur ses premiers ouvrages. Thèse de lettres. Paris: Champion,

1924.

Yovanowich, Voyslav. "La Guzla" de Prosper Mérimée. Les Origines du livre, ses

sources, sa fortune. Étude d'histoire romantique. Thèse de lettres. Grenoble:

Allié Frères, 1910.

Yule, G. Udny. "On Sentence-Length as a Statistical Characteristic of Style in Prose, with

Application to Two Cases of Disputed Authorship". Biometrika 30 (1938): 369-

90.

Zipf, George Kingsley. Human Behaviour and the Principle of Least Effort. New York: Hafner, 1969.

VI. OUVRAGES CONSULTÉS

Allaire, Denis. Comprendre la statistique. Québec: Presses de l'Université du Québec, 1995.

Amacker, René. Linguistique saussurienne. Genève: Droz, 1975.

Androuard, André. Le Canard enchaîné 1er octobre 1975: 12

Anorthy, Christine. Le Parisien libéré 29 octobre 1974: 23.

Anski, Shalom. Le Dibbouk. Paris: L'Arche, 1957.

Antosch, F. "The Diagnosis of Literary Style with the Verb-Adjective Ratio." Statistics and Style. Eds. Lubomír Dolezel and Richard W. Bailey. New York: American Elsevier, 1969: 221-8.

Autrand, Dominique. "Pour Ajar, l'écriture c'est le règne du pseudo". Quinzaine littéraire 247 (1977): 9.

Baer, Colette. "Pseudo: un chant d'humour". Les Nouveaux Cahiers 48 (1977): 75-7.

Bailey, Richard W. "Authorship Attribution in a Forensic Setting". Advances in Computer-aided Literary and Linguistic Research. Eds. David E. Ager et al. Birmingham: AMLC, 1979: 67-80.

—. "Statistics and Style: A Historical Survey". Statistics and Style. Eds. Lubomír

Dolezel and Richard W. Bailey. New York: American Elsevier, 1969: 217-36.

Baker, F.M. "A Test of Authorship Based on the Rate at Which New Words Enter an

Author's Text". Journal of the Association for Literary and Linguistic Computing

3:1 (1988): 36-9.

Bally, Charles. Linguistique Générale et linguistique française. Berne: Francke, 1944.

Baroche, Christiane. "Émile Ajar: *Gros-Câlin*". Nouvelle Revue Française 265 (1974):

105-6.

—. "Émile Ajar: *La Vie devant soi*". Nouvelle Revue Française 227 (1975): 989.

Bayard, Pierre. "Romain[G] et Mohammed: du roman parental au récit d'enfance".

Revue des Sciences humaines (avril-juin 1975): 105-120.

Beauvoir, Simone de. Les Belles Images. Paris: Gallimard, 1966.

Bélisle, Jean-Pierre et Jacques Desrosiers, Introduction à la statistique. Quebec: Gaëtan

Morin, 1983.

Belle-Isle, Francine. "Autobiographie et analyse: là où le rêve prend corps". Études

littéraires 17.2 (août 1984): 371-80.

Benveniste, Émile. Problèmes de linguistique générale. Paris: Gallimard, 1966.

Berdes, Charles. "Winter tales - Fiction about Aging". Gerontologist 21.2 (1981): 121-4.

Bernet, Charles. "La richesse lexicale de la tragédie classique". Le Français Moderne 1

(janv. 1978): 44-53.

- Boisen, Jørn. "La conception du temps chez Romain Gary". Revue Romane 4.51 (1989): 51-70.
- Bondy, François. "A Man and his double". Encounter 57.2 (Aug. 1981): 42-3.
- Borel, Émile. Probabilités et Certitude. Paris: P.U.F. 1961.
- . Le Hasard. Paris: Presses Universitaires de France, 1961.
- Boujout, Michel. "Les Clowns lyriques". Nouvelles Littéraires 2692 (21 juin 1979): 23.
- Bourgeade, Pierre. "La Vie et mort d'Émile Ajar". Nouvelle Revue Française 345 (oct. 1981): 149-50.
- Bourgeaud, Georges. Voyage à l'étranger. Paris: Gallimard, 1974.
- Boursin, Jean-Louis. Les Structures du hasard. Paris: Seuil, 1966.
- Bouzerand, Jacques. "Énigme: Qui est Ajar". Le Point (10 novembre 1975): 61.
- . "Émile Ajar parle". Le Point (17 novembre 1975): 174-6.
- Brainerd, Barron. "Statistical Analysis of Lexical Data Using Chi-squared and Related Distributions". Computers and the Humanities 9 (1975): 161-78.
- Bruno, Agnes M. Toward a Quantitative Methodology for Stylistic Analyses. Berkley: University of California Press, 1974.
- Burrows, John F. "Not Unless You Ask Nicely: The Interpretive Nexus between Analysis and Information". Literary and linguistic Computing 7. 2 (1992): 91-109.
- . "Word-Patterns and Story-Shapes: The Statistical Analysis of Narrative Style". Literary and Linguistic Computing 2. 2 (1987): 61 -70.

Burton, Dolores, M. "Aspects of Word Order in Two Plays of Shakespeare". Computers and the Humanities 3 (1970): 34-9.

—. Shakespeare's Grammatical Style, a Computer-assisted analysis of *Richard II* and *Anthony and Cleopatra*. Austin: University of Texas Press, 1973.

Carroll, John B. "Vectors of Prose Style". Statistics and Style. Eds Lubomír Dolezel and Richard W. Bailey. New York: American Elsevier. 1969: 80-91.

Chalon, Jean. "Avec les ajarophiles (*Pseudo*)". Figaro Littéraire 1596 (18 dec. 1976): 19.

Chardin, Jean-Marie. Journal de Voyage. Paris, 1711.

Cheydleur, Frédéric D. A French Idiom List Based on a Running Count of 1,183,000 Words. New York: Macmillan, 1929.

Cluny, Claude, M. "La Vie devant soi". Magazine Littéraire 106 (nov. 1975): 32.

Combes, Patrick. La Littérature et le mouvement de mai 1968: Écritures, Mythes, Critiques, Essai 1968-1981. Paris: Seghers, 1984.

Costaz, Gille. "Ajar, alias Romain Gary". Le Matin (2 juillet 1981): 22.

Damereau, F. J. "The Use of Function Word Frequencies as Indicators of Style". Computers and the Humanities 9 (1975): 271-80.

Dauberton, Annie. "Le soleil, le clown, les chiens. (*La vie devant soi*)". Nouvelles Littéraires 2500 (29 sept. 1975): 5.

Day, Leroy T. "Gary-Ajar and the rhetoric of non-communication". French Review (Oct. 1991): 75-83.

Delcourt, Christian. "On Vocabulary Curves". Association for Literary and Linguistic

Computing Journal 2 (1981): 13-24.

Dellepierre, Octave. Supercherries littéraires, pastiches, suppositions d'auteur dans les

lettres et dans les arts. London: N. Trüber, 1872.

Dolezel, Lubomír. "A Framework for the Statistical analysis of Style". Statistics and

Style. Eds. Lubomír Dolezel and Richard W. Bailey. New York: American

Elsevier, 1969: 10-25.

---. "Un modèle statistique de codage linguistique". Études de linguistique appliquée 3

(1965): 51-63.

Dolezel, Ludomír et Richard W. Bailey. An Annotated Bibliography of Statistical

Stylistics. Michigan: University of Michigan, 1969.

Dormann, Geneviève. "Pourquoi mon ami Romain s'est-il suicidé"? Figaro Littéraire 98

(6 dec 1980): 102-3.

Ducretet, Pierre. "Quantitative Stylistics: An Essay in Methodology". Computers and the

Humanities 4 (1970): 187-91.

Ducrot, Oswald et al. Les Mots du discours. Paris: Éditions de Minuit, 1980.

Dufrunoy, Marie-Louise. "Analyse statistique du langage". Journal de la société de

statistique de Paris 87 (1946): 208-19.

Dugast, Daniel. "La Mesure de la richesse lexicale". Verbum 1 (1980): 89-93.

- . "Sur quoi se fonde la notion d'étendue théorique du vocabulaire"? Le Français moderne 46.1 (janvier 1978): 56-67.
- . La statistique lexicale. Genève: Slatkine, 1980.
- . Vocabulaire et Discours: Essai de lexicométrie organisationnelle. Genève: Slatkine, 1979.
- . Vocabulaire et Stylistique: Théâtre et dialogue. Genève: Slatkine 1979.
- Duras, Marguerite. Le vice-consul. Paris: Gallimard, 1966.
- Eerde, John. A. von. "Au-delà de cette limite votre ticket n'est plus valable". Books Abroad 76 (1975): 826-7.
- Enckell, Pierre. "Ajar est mort, vive Ajar! Sept ans de vie derrière soi". Nouvelles littéraires (2 juillet 1981): 32.
- Erdman, David V. & E. G. Fogel eds. "Evidence for Authorship". Essays & Problems of Attribution. Ithaca: Cornell University Press, 1966.
- Ezine, Jean-Louis. "Ajar ou l'art de se faire peur". Nouvelles Littéraires 2370 (27 mars 1980): 19.
- . "Après la mort de Gary, le suicide est-il un geste littéraire"? Nouvelles Littéraires 2766 (11 déc. 1980): 60.
- . "Lettre ouverte à mon ami de longue date". Nouvelles Littéraires 2795 (9 juillet 1981): 24.

---. "Voyage au pays d'Ajar ou Ajar malgré lui. *Pseudo*". Nouvelles Littéraires 2502 (9 déc. 1976): 3-4.

Fabre, Dominique. "Le refus d'Ajar: de quoi rire". La Suisse (21 novembre 1975): 35.

Fortier, Paul A. "THEME: A System for Computer-Aided Theme Searches of French Texts". Research Reports. Winnipeg: University of Manitoba, 1975.

---. "The Importance of Computer-Aided Analysis of Literature to Comparative Literature". Actes du Congrès de l'Association Internationale de Littérature comparée. Budapest: Akadémia Kiadó, 1972: 315-318.

---. "Vers un indice de thématization dans le roman français". Revue Informatique et statistique dans les sciences humaines 1-4 (1988): 181-91.

Fouletier-Smith, Nicole. "Les Nord-africains en France: réalités et représentations littéraires". French Review 51.5 (1978): 683-91.

Freund, John E. et Frank Williams. Elementary Business Statistics. London: Prentice-Hall, 1975.

Freund, John E. et Ronald Walpole. Mathematical Statistics. New Jersey: Prentice-Hall, 1987.

Freustié, Jean. "La défense Ajar". Le Nouvel Observateur (27 décembre 1976): 51.

---. "La Manière Ajar". Le Nouvel Observateur (26 février 1979): 69.

---. "Le gamin et la monstresse: *La Vie devant soi*". Le Nouvel Observateur (22 sept. 1975): 74.

Freyburger, Henri. "Émile Ajar: *L'Angoisse du Roi Salomon*". French Review 52 (1979-80): 972-3.

Froger, Jacques. "La Critique des textes et l'ordinateur". Vigilae Christianae 24 (1970): 210-7.

Garcin, Jérôme. "Gary derrière le masque". Nouvelles Littéraires 2795 (9 juillet. 1981): 25.

Gary, Romain. "Interview". Paris-Normandie 15 février 1977: 12.

---. "Romain Gary s'attaque au nouveau roman". Figaro littéraire 103 (sept. 1965): 23-9.

Garzorolli, Richard. "Émile Ajar, l'angoisse du roi Salomon". Le Matin (23 février 1979). 28.

Genette Gérard. Introduction à l'architexte. Paris: Seuil, 1979.

---. Nouveau discours du récit. Paris: Seuil, 1983.

---. Palimpsestes. Paris: Seuil, 1982.

---. Figures I. Paris: Seuil, 1966.

---. Figures III. Paris: Seuil, 1972.

Glascow, Janis. "*La Nuit sera calme*", "*Europa*", "*Les Enchanteurs*". French Review 49 (1975-76): 442-3.

Gracq, Julien. Le Rivage des Syrtes. Paris: Corti, 1951.

Greenwood, H. H. "Common Word Frequencies and Authorship in Luke's Gospel and Acts". Literary and Linguistic Computing 10.3 (1995): 182-7.

- , "St. Paul Revisited: a Computational Result". Literary and Linguistic Computing 7 (1992): 43-7.
- , "St. Paul Revisited: Word Clusters in Multidimensional Space". Literary and Linguistic Computing 8 (1995): 211-9.
- Greimas, Algirdas J. Sémantique structurale. Paris: Larousse, 1972.
- Guiraud, Pierre. Essai de stylistique. Paris: Klincksiek, 1980.
- , Problèmes et Méthodes de la statistique linguistique. Paris: Presses Universitaires de France, 1960.
- Guiraud, Pierre et Pierre Kuentz. La Stylistique. Initiation à la linguistique. Série A. Paris: Klincksiek, 1970.
- Harkin, David. "The History of Word Counts". Babel 3.3 (1957): 113-24.
- Hayes, Curtis. "A Study in Prose Styles: Edward Gibbon and Ernest Hemingway". Statistics and Style. Eds. Lubomír Dolezel and Richard W. Bailey. New York: American Elsevier, 1969: 80-91.
- Hecquet, Stephen. Le Bulletin de Paris 18 octobre 1956: 56-7.
- Henmon, Vivian A. C. A New Derivation and Interpretation of Yule's 'Characteristic K'. Journal of Applied Mathematics and Physics 6 (1955): 332-4.
- , The Advanced Theory of Language as Choice and Chance. New York: Springer-Verlag, 1966.

Hockey, Susan. A Guide to Computer Applications in the Humanities. London: John Hopkins University Press, 1980.

Holmes, David. I. et Richard S. Forsyth. "The Federalist Revisited: New Directions in Authorship Attribution". Literary and Linguistic Computing 8.2 (1993): 110-27.

Honoré, A. "Some Simple Measures of Richness of Vocabulary". Association for Literary and Linguistic Computing Bulletin 7.2 (1979): 172-7.

Käding, F. W. Häufigkeitwörterbuch der deutschen Sprache. Steglitz bei Berlin: der Herausgeber, 1897.

Kaufmann, Judith. "La Danse de Romain Gary ou Gengis Cohn et la valse-hora des mythes de l'occident". Études Littéraires 17.1 (avril 1981): 71-94.

Kenny, Anthony. A Stylometric Study of the New Testament. Oxford: Oxford University Press, 1986.

Laan, Nancy M. "Stylometry and Method. The Case of Euripides". Literary and Linguistic Computing 10.4 (1995): 271-9.

Lainé, Pascal. La Dentellière. Paris: Gallimard, 1974.

Lalande, Bernard. "Ajar. Pour une lecture de *La vie devant soi*". Le Français dans le Monde 158 (1981): 61-2.

Lamont, Rosette. "Émile Ajar: *La Vie devant soi*". French Review 50:6 (1977): 952-3.

Laugaa, Maurice. "Autobiographie d'un pseudonyme, George Sand, R. Abellio, R. Gary". Textuel 22 (1989): 111-35.

Leclerc, Pierre-Robert. "Les deux morts de Romain Gary". Magazine Littéraire 256

(1988): 61-2.

Ledger, Gerard. "An Exploration of Difference in the Pauline Epistles using Multivariate

Statistical analysis". Literary and Linguistic Computing 10.2 (1995): 84-97.

Leich, David. "Vanishing Manhood: Au-delà de cette limite votre ticket n'est plus

valable". Times Literary Supplement 75 (1976): 158.

Lejay, Paul. Préface à la syntaxe latine d'O. Reimann, Paris, 1894.

Lorian, Alexandre. "Les Raisonnements déraisonnables d'Émile Ajar". Hebrew

University Studies in Literature and the Arts (1987): 120-48.

Maingueneau, Dominique. Initiation aux méthodes de l'analyse du discours. Paris:

Hachette, 1976.

Martinet, André. Éléments de linguistique générale. Paris: Armand Colin, 1966.

Mauraux, Jean-Baptiste. "Un gentil divertissement: *La Vie devant soi*". Quinzaine

Littéraire 220 (1er nov. 1975): 10.

McKinnon, Alistair et Robert. Webster. "A Method for Author Identification". The

Computer in Literary and Linguistic Research. Éd. R.A. Wisbey. Cambridge:

Cambridge University Press, 1971 .

Medvedev, Roy A. Problems in the Literary biography of Mikhail Sholokov. Tr. A. D. P.

Briggs. Cambridge: Cambridge University Press, 1977.

Merriam, Thomas V.M. "What Shakespeare Wrote in Henry VIII, part 1". The Bard 2 (1979): 81-94.

Metz, George Harold (ed). The Sources of Four Plays Ascribed to Shakespeare.
Columbia: University of Missouri Press, 1989.

Michaelson, Sidney et Andrew Q. Morton. "Identifying Aristotle". Computer Calepraxis I (1972): 13-42.

Michalczyk, John J. "Clair de femme". Film Costa-Gavras. French Review 53 (1979-80): 638-9.

Mogui, Jean-Pierre. "Émile Ajar avait répondu d'avance à nos questions". Le Figaro (18 novembre 1975): 42.

Molinié, Georges. La Stylistique. Paris: Presses Universitaires du théâtre de France, 1989.

Monjallon, A. Introduction à la méthode statistique. Paris: Vuibert, 1958.

Monselet, Charles. De A à Z: Portraits contemporains. Paris: Charpentier, 1888.

Moore, David. S. et George P. McCabe. Introduction to the Practice of Statistics. New York, London: W. H. Freeman, 1989.

Mordillat, Gerard. "Ajar, Gary, Pavlowitch: Pour qui sonne le gag?" Libération (3 juillet 1981): 23.

Moreau, Jean A. "Le langage devant soi: La Vie devant soi". Critique 30 (1975): 1197-9.

Morton, Andrew Q. Literary Detection. New York: Scribner's, 1978.

Muller, Charles. "Baudelaire: *Les Fleurs du Mal*. Concordances, index et relevé statistique établis d'après l'édition de Crépet-Blin par le Centre d'Étude du Vocabulaire Français de la faculté des lettres de Besançon, avec la collaboration de K. Menemcioglu". Cahiers de lexicologie 7.2 (1965): 33-42.

---. "La longueur moyenne du mot dans le théâtre classique". Cahiers de lexicologie 5 (1964): 29-44.

---. "Peut-on estimer l'étendue d'un lexique"? Cahiers de Lexicologie 27 (1975): 3-29.

O'Donnell, Bernard. An Analysis of Prose Style to Determine Authorship: The O'Ruddy, a Novel by Stephen Crane and Robert Barr. The Hague: Mouton, 1970.

Oakman, Robert. Computer Methods for Literary Research. Athens: University of Georgia Press, 1984.

Oates, James C. Success and the Pseudonymous Writer: Turning Over a New Self. Princeton: Princeton University Press, 1987.

Orsenna, Eric. L'Exposition coloniale. Paris: Le Seuil, 1989.

---. La Vie commune à Lausanne. Paris: Le Seuil, 1977.

---. Loyola's Blues. Paris: Le Seuil: 1974.

---. Une Comédie française. Paris: Le Seuil, 1988.

Östman, Anne-Charlotte. "The disguises of French writer Romain Gary". Bonniers Literara Magasin 57.1 (1988): 38-41.

Pineau, Vincent. "Une lecture de l'imaginaire dans L'Angoisse du Roi Salomon de

Romain Gary/Émile Ajar". Recherches sur l'imaginaire, Cahier 24 (1993):

255-65.

Pons, Anne. "Ajar, seul avec tous". Le Point (5 février 1979): 82.

Pudlowski Gilles. "Émile Ajar a-t-il assassiné Gary"? Nouvelles Littéraires 2795 (2

juillet 1981): 23-4.

Pudlowski, Gilles et Jérôme Garvin. "Ajar est mort, vive Ajar! Sept ans de vie derrière

soi". Nouvelles Littéraires 2794 (2 juillet 1981): 32.

Queneau, Raymond. Les Oeuvres complètes de Sally Mara. Paris: Gallimard, 1962.

Radday, Yehuda T. et Haim Shore. Genesis: An Authorship Study in Computer-assisted

Statistical Linguistics. Rome: Biblical Institute Press, 1985.

Raspiengas, Jean-Claude. "La bonne moitié". Nouvelles Littéraires 2685 (3 mai 1979):

23.

Reeb, George et Alain Fuchs. Statistique commentée. Paris: Gauthier-Villars, 1967.

Reuchlin, Maurice. Précis de statistique: présentation notionnelle. Paris: Presses

Universitaires de France, 1976.

Robert, Paul. Dictionnaire de la langue française. Paris: Nouveau Littre, 1979.

- Rosengren, I. "The Quantitative Concept of Language and its Relation to the Structure of Frequency Dictionaries". Études de linguistique appliquée, nouvelle série 1 (1971): 3-27.
- Sagan, Françoise. Bonjour tristesse. Paris: Juillard, 1954.
- . La Chamade. Paris: Juillard, 1965.
- Sangolowski, Joseph. "La Judéité dans l'oeuvre de Romain Gary. De l'ambiguïté à la transparence symbolique". Études littéraires 17.5 (été 1993): 111-27.
- Sedelow, Sally Y. "The Use of the Computer for Stylistics Studies of Shakespeare". Computer Studies in the Humanities and Verbal Behaviour 4 (1973): 33-6.
- Sincich, Terry. Statistics by Example. San Francisco: Dellen, 1982.
- Sipriot, Pierre. "Harold et Maud des bas-fonds, *La Vie devant soi*". Figaro Littéraire 1530 (13 sept. 1975): 14.
- Smith, John Bristow. Imagery and the Mind of Stephen Dedalus: A computer-assisted study of James Joyce's *A Portrait of the Artist as a young Man*. Lewisburg (Pa.): Bucknell University Press, 1988.
- Smith, Michael William. "Attribution by Statistics: A Critique of Four Recent Studies". Revue Informatique et Statistique dans les Sciences Humaines 26 (1990): 233-51.
- . "The Authorship of *Pericles*. An Initial Investigation". The Bard 3 (1982): 143-76.
- . "The Authorship of the *Revenger's Tragedy*". Notes and Queries 38.4 (1991) 166-74.

Spencer, Samia. "Madame Rosa, film de Moshe Mizrahi d'après le roman de Momo".

French Review 54 (1980-81): 337-42.

Steele, R. B. "Non-recurrence of Vocabulary as a Test of Authorship". Philological

Quarterly 4 (1925): 267-80.

Tallentire, David. R. "Confirming Intuitions about Style Using Concordances". The

Computer in Literary and Linguistic Studies. Eds. Allan. Jones et Robert. F.

Churchouse. University of Wales Press. 1976.

---. "The Mathematics of Style Thinking by Numbers, 2". Times Literary Supplement

13 (August 1977): 73-4.

Tharp, James Burton et al., "A Basic French Vocabulary". Modern Language Journal 28

(1934): 238-74.

Tortrat, Alain. Principes de statistique mathématique. Paris: Dunod, 1961.

Tournier, Michel. "Émile Ajar ou le vie derrière soi". Nouvelle Revue Française 325

(1975): 329-44.

---. Le Vol du Vampire. Paris: Mercure de France, 1981.

Triolet, Elsa. Le premier accroc coûte 200 francs. Paris: Gallimard, 1944.

Ullmann, Stephen. Style in the French novel. Cambridge: Cambridge University Press,

1957.

Vander Beke, George E. French Word Book. New York: Macmillan, 1929.

- Vinay, Jean-Pierre et Jean Darbelnet. Stylistique Comparée du français et de l'anglais. Montréal: Beauchemin, 1958.
- Vincent, E. R. "Mechanical Aids for the Study of Literature and Style". Literature and Science: Proceedings of the Sixth Triennial Congress of the International Federation for Modern Languages and Literature. Oxford: Oxford University Press, 1955: 56-60.
- Wagner, Robert-Léon et Pierre Guirard. "La méthode statistique en lexicologie". Revue d'Enseignement Supérieur 1 (1959): 154-9.
- Wagner, Susan. Introduction to Statistics. New York: Harper Collins, 1992.
- Wake, William Charles. "Sentence-Length Distributions of Greek Authors". Journal of the Royal Statistical Society A. 120 (1957): 331-46.
- Wardi, Charlotte. "Le génocide des Juifs dans la fiction romanesque française: 1945-70. Une expression originale, *La Danse de Gengis Cohn*". Diagraphe 25 (Printemps 1981): 215-25.
- West, Michael Phillip et Otto Bond. A Grouped Frequency French Word List. Based on the French Word Book of Vander Beke. Chicago: University of Chicago Press, 1939.
- West-Sooby, John. "Romain Gary et le goût du paradoxe: l'exemple de *Clair de femme* et de *La Vie devant soi*". Studi Francesi 23 (maggio-agosto 1992): 307-13.

Williams, C.B. "A Note on the Statistical Analysis of Sentence-Length as a Criterion of Literary Style". Statistics and Style. Eds. Lubomír Doležel and Richard W.

Bailey. New York: American Elsevier, 1969: 69-75.

—. Style and Vocabulary: Numerical Studies. London: Charles Griffin, 1970.

Wills, G. The Federalist papers of Alexander Hamilton, James Madison and John Jay.

Toronto: Bantam Press. 1982.

Wolfzettel, F. "After the Nouveau Roman, the Relationship of History and Myth in the contemporary French novel". Beiträge Zur Romanischen Philologie 25.2 (1988): 217-27.