

# Etude de Statistique Lexicale : Accroissement Lexical dans l'Œuvre Littéraire de Julien Gracq\*

Elarbi ELBAKKALI\*\*

**Résumé**— Le présent travail vise l'examen de l'accroissement du vocabulaire de l'auteur à travers toute sa production littéraire. Ce travail s'inscrit dans le cadre de la statistique lexicale. Il s'agit de déterminer l'apport du vocabulaire au fil du temps : pour un segment déterminé du texte c'est le nombre d'unités nouvelles qui n'avaient pas été employées antérieurement et qui apparaissent dans ce segment. Cette analyse permet de constater que l'apport de nouvelles unités dans la seconde période artistique de l'auteur (1967-1992) est intense par rapport à celui remarqué lors de la première période littéraire (1938-1967). Aussi, l'apport lexical à travers l'œuvre de Gracq n'est pas stable puisqu'il change automatiquement à mesure que l'auteur passe d'un genre littéraire à un autre. En effet, on peut constater que la deuxième période artistique de notre auteur Julien Gracq est plus riche que la première et qu'elle connaît aussi un apport lexical plus intense que celui de la première. Le vocabulaire des œuvres de la seconde période est très particulier. Cela émane principalement du style d'écriture adopté par l'auteur dans la seconde moitié de sa vie littéraire dans la mesure où *Lettrines I* de Julien Gracq a inauguré un style d'écriture qui échappe à une définition classique, il ne semble pas exagéré de penser qu'il renouvelle une forme d'expression originale qui d'autres écrivains emprunteront après lui. A vrai dire, les œuvres de la seconde période font partie de ce que nous appelons communément la 'Littérature Fragmentée' qui se définit comme un ensemble très libre, une mosaïque de mots pour lire, réflexions et souvenirs (souvent de l'enfance).

**Mots-clés**— Accroissement, Apport, Lexicométrie, Lexique, Tranche, Vocabulaire.



## Lexical Statistics Study: Lexical Increase in Julien Gracq's Literary Work\*

Elarbi Elbakkali\*\*

**Extended abstract**— The present work aims to examine the develop process of author's vocabulary throughout his literary production. This work study the framework of lexical statistics. It is a question of determining the contribution of the vocabulary over time: for a given segment of the text, it is the number of new units which had not been used previously and which appear in this segment. In principle, lexical statistics are mainly based on the notion of frequency. Thus, the frequency of words in lexicometric studies is a dynamic character, as important as their meaning or their phonetic figure : the concentration corresponds, at the stylistic level, to the motivation of the author for such use of a word and the dispersion leads to the characterization and specification desired by the author. These two defined characters are likely to provide criteria for the identification of texts and for the analysis of the stylistic use of words and their expressive value, both in the language and in a particular writer. The effectiveness of lexical statistics is having reliable data on the lexicon of a text ; these data are multiple and they can be either digits when it comes to the number of each occurrence of a given text, or percentages which come from the ratio of the number of each occurrence to the set of occurrences of the text, or graphics: representations of a word according to parameters chosen beforehand. All this is achievable by resorting to quantitative statistics, especially the use of the Hyperbase software (developed by Professor Etienne Brunet from the University of Nice Sophia Antipolis in France) which is considered a centerpiece of lexical statistics in recent years. It is then sufficient to submit the computerized literary base to the various functionalities guaranteed by this software in order to have all the necessary lexical quantitative data. Admittedly, the principle of lexical statistics has not changed since its appearance and its position among the branches of linguistics, but we note a great evolution at the level of the processing tools intended for this purpose and in particular the software responsible for this task.

In this study we have worked on the entire literary production of Julien Gracq who is one of the masters of contemporary literature in France. Louis Poirier (his real name) or Julien Gracq (professor of geography) was able to mark his writing career without subscribing to any specific literary movement. His literary production, which spanned half a century (1938-1992), encompassed seventeen works and brought together several literary genres, notably literary essays which are richly represented with ten works. Indeed, his production has known two major literary periods: the first period (1938-1967) took surrealism and fiction as a main framework for literary creation while the second period (1967-1992) was focused on fragments of autobiography, reflections on literature and geographic mediations. It is important to underline that the common point between these two artistic periods of the author is the omnipresence of the landscape (fictitious or real) where the edge constitutes for Gracq a springboard which makes pass the intrigue of the story from a closed place (forest or castle), to another open (river or sea). The edge then embodies a climate of expectation or a sort of confused and ambiguous projection into the future.

It is to say that, in this work we have to highlight lexical growth or how the vocabulary progresses and regresses through the seventeen works of Julien Gracq which are: (At Argol castle(1938), A

*Archive of SID*

beautiful dark ( 1945), Great freedom(1945), The Fisher King (1948), André Breton, aspects of a writer (1948), The Shore of Syrtis (1951), Penthesilea (1954), A balcony in the forest (1958), Preferences (1961), Initials 1 (1967), The peninsula (1970), Initials 2 (1974), Narrow Waters(1976), Reading while writing (1980), The shape of a city (1985), Around of seven hills (1988) and Diaries of the great road (1992).)

The study of lexical growth determines the contribution of vocabulary over time; this increase is, for a given segment of the text, the number of new units which had not been used previously, which appear in that segment. To perform the necessary lexical growth calculations, we need to slice our corpus. In this study, we will adopt a distribution that follows a so-called natural division of the corpus, that is to say that this division is carried out in relation to the number of works that the corpus contains. Therefore, we will study the lexical growth of Gracq's corpus in seventeen slices where each slice represents one of the works of the author's production. Indeed, the various lexicon growth calculations are made by the Hyperbase software. This also takes care of the graphs and histograms that correspond to each lexical calculation table. The Hyperbase software therefore, first of all, performs the normal lexical growth calculations according to the chronological order of the production of each work in the corpus. Then he calculates the lexical increase of the different parts of the corpus by following a reverse chronological order where he takes the last work of Gracq entitled Diaries of the great road as the starting point of calculation and consequently the first work, At Argol castle, will become the last work to deal with its lexical increase.

In short, the analysis of the lexical contribution in Gracq's work allows us to carry out the following interpretations: at the beginning, the contribution of new units in the author's second artistic period (1967- 1992), which begins with the booklet entitled Initials 1, is intense compared to that noticed during the first literary period (1938 -1967) where the increase was above the average except for the significant value of accusation observed in the collection of prose poems: great freedom. Second, the lexical contribution through Gracq's work is not stable since it changes automatically as the author moves from one literary genre to another. Finally, the works in which Gracq describes real landscapes that are encountered during his life have an intense lexical contribution, on the other hand the works that were written within the framework of surrealism are judged to be weak in terms of the contribution of new words. Alain-Michel Boyer in a way justifies our interpretations made at the end of a study of lexical growth through the seventeen texts of Julien Gracq by saying:

*“But what strikes the reader of Julien Gracq is the singular evolution of a work which, from book to book, from At Argol castle, published in 1938, to Diaries of the great road, published half a century ago. Later, follows a curve, performs a backward course, which leads from a surrealist writing game to a meditation on Rome and which gives an increasingly important role, not only to spatial configurations, but also to real places.”* (Boyer 35)

Also, the lexical contribution through Gracq's work is not stable since it automatically changes as the author moves from one literary genre to another. Indeed, we can see that the second artistic period of our author Julien Gracq is richer than the first and that it also knows a more intense lexical contribution than that of the first. The vocabulary of works from the second period is very particular. This emanates mainly from the writing style adopted by the author in the second half of his literary life insofar as from Initials 1 Julien Gracq inaugurated a writing style which escapes a classical definition, it does not seem exaggerated to think that he renews an original form of expression that other writers will borrow after him. To tell the truth, the works of the second period are part of what we commonly call "Fragmented Literature" which is defined as a very free set, a mosaic of words to read, reflections and memories (often from childhood).

**Keywords**— Increase, Contribution, Lexicometry, Lexicon, Slice, Vocabulary.

## SELECTED REFERENCES

- [1] Boyer, Alain. *Paysages et mémoires /des eaux étroites à un balcon en forêt*. Nantes : Editions Cécile défaut, 2007.
- [2] Brunet, Etienne. *Le vocabulaire français de 1789 à nos jours*. Genève-Paris : Slaktine-Champion, 1981.
- [3] Brunet, Etienne. « *Le traitement des faits linguistiques et stylistiques. Texte d'application : Giraudoux* ». *Statistiques et linguistique*, Actes et colloque, Klincksieck, n. 15, 1974, pp. 105-136.
- [4] Canovas, Frédéric. *L'écriture révée*. Paris : l'Harmattan, 2000.
- [5] Gracq, Julien. *Au château d'Argol*. Paris : José Corti, 1938.
- [6] Gracq, Julien. *Un beau ténébreux*. Paris : José Corti, 1945.

## مطالعه آمار واژگانی:

# «افزایش واژگان» در آثار ادبی ژولین گراک\*

الربی الباکالی\*\*

**چکیده** — مقاله حاضر در چارچوب مطالعات آمار واژگانی، به بررسی افزایش واژگان در آثار ادبی ژولین گراک می‌پردازد. بررسی‌ها نشان می‌دهد که استفاده از واژگان جدید در دوره دوم هنری این نویسنده (۱۹۹۸-۱۹۹۲) بیشتر و غنی‌تر از دوره اول هنری وی (۱۹۶۷-۱۹۳۸) است. به علاوه، افزایش واژگان در آثار گراک ثابت نبوده و با گذر از ژانری به ژانر دیگر تغییر پیدا می‌کند. در این مقاله خواهیم دید که لغات استفاده شده توسط گراک در دوره دوم هنری، بسیار خاص و ناشی از سبک نوشتاری ویژه نویسنده در این دوره می‌باشد. سبک نوشتاری جدیدی که گراک با نگارش *لترین*، ارائه می‌کند، با تعاریف کلاسیک مطابقت ندارد. در حقیقت، اگر بیان کنیم این نویسنده فرم بیانی اصلی را تجدید نموده است که نویسندگان پس از او، از آن وام می‌گیرند، اغراق نکرده‌ایم. آثار گراک در دوره دوم بخشی از ادبیات به عبارتی «تکه تکه» را شامل می‌شوند که به‌عنوان مجموعه‌ای کاملاً آزاد، موزاییکی از واژگان، تأملات و خاطرات (غالباً دوران کودکی) تعریف می‌شود.

**کلمات کلیدی** — ژولین گراک، آمار واژگان، افزایش، سبک نوشتاری.

## I. INTRODUCTION

Ce travail vise à étudier l'accroissement du vocabulaire dans l'œuvre littéraire de Julien Gracq et établir ce que cela signifie sur le plan sémantique. La tâche paraît difficile et exige une très longue et épineuse période de travail, mais le recours à une discipline appelée 'statistique lexicale', une telle étude est réalisable en un peu de temps dans la mesure où les données quantitatives plausibles à l'exploitation sont établies automatiquement sous forme de bases de données informatisées.

'La statistique lexicale' (linguistique quantitative, lexicométrie ou bien lexicologie statistique) consiste à associer le nombre au discours et la méthode à l'objet. Toutes les dénominations que l'on a proposées pour cette branche de la linguistique font référence au lexique et parfois plus largement à la linguistique mais jamais à la littérature ni à la critique littéraire.

L'importance d'une étude de statistique lexicale est évoquée au départ par Charles Muller, père fondateur de cette branche en France vers les années cinquante, qui affirme que :

*« L'important est de constater que le vocabulaire d'un texte, qui est un échantillon d'un lexique virtuel, obéit dans sa structure quantitative à des impulsions qui ne sont pas fortuites, et se construit suivant des lois complexes et mal connues encore, mais assez impérieuses pour que nous ayons pu en définir quelques effets, le matériel lexical est une « matière » qui a ses lois propres, sa mécanique particulière, dont la découverte ne peut laisser aucun linguiste indifférent » ( Muller, Initiation 165).*

En effet, les mots de l'écrivain constituent bien la matière première de ces travaux lexicométriques, mais immédiatement soumis à des transformations et à des classements, ils perdent dès l'entrée leur authenticité, leur individualité, leur sens et leur saveur, dans la mesure où la structure lexicale met en évidence des classes, des catégories, des effectifs, c'est-à-dire des nombres.

En principe, la statistique lexicale se fonde principalement sur la notion de fréquence. Ainsi, la fréquence des mots dans les études lexicométriques est un caractère dynamique, aussi important que leur signification ou leur figure phonétique : la concentration correspond, au niveau stylistique, à la motivation de l'auteur pour tel emploi d'un mot et la dispersion mène à la caractérisation et à la spécification voulue par l'auteur. Ces deux caractères ainsi définis sont susceptibles de fournir des critères à la fois pour l'identification des textes et pour l'analyse de l'emploi stylistique des mots et de leur valeur expressive, à la fois dans la langue et chez un écrivain particulier.

L'efficacité de la statistique lexicale c'est avoir des données fiables sur le lexique d'un texte ; ces données sont multiples et elles peuvent être soit des chiffres quand il s'agit du nombre de chaque occurrence d'un texte donné, soit des pourcentages qui proviennent du rapport du nombre de chaque occurrence à l'ensemble des occurrences du texte, soit des graphiques : des représentations d'un vocable selon des paramètres choisis<sup>1</sup> au préalable.

Tout cela est réalisable par le recours à la statistique lexicale surtout l'utilisation du logiciel Hyperbase<sup>2</sup> qui est considéré comme une pièce maîtresse de la statistique lexicale, ces dernières années. Il suffit alors de soumettre la base littéraire informatisée aux différentes fonctionnalités garanties par ce logiciel pour avoir toutes les données quantitatives lexicales que nous venons d'évoquer ci-dessus. Certes, le principe de la statistique lexicale n'a pas changé depuis son apparition et sa position parmi les branches de la linguistique, mais l'on remarque une grande évolution au niveau des outils de traitements destinés à cette fin et notamment les logiciels chargés de cette tâche.

Dans cette étude nous travaillons sur toute la production littéraire de Julien Gracq qui est l'un des maîtres de la littérature contemporaine en France. Louis Poirier (son vrai nom) ou Julien Gracq (professeur de géographie) a pu marquer sa carrière d'écrivain sans qu'il s'inscrive à aucun courant littéraire précis. Sa production littéraire, qui s'étale sur un demi-siècle (1938-1992), englobe dix-sept



ouvrages et rassemble plusieurs genres littéraires, notamment les essais littéraires qui sont richement représentés avec dix ouvrages. Ainsi l'œuvre de Gracq est protéiforme puisqu'il y a les essais, les romans, les pièces de théâtre et les poèmes en prose. En effet, la production littéraire de Gracq a connu deux grandes périodes littéraires : la première (1938-1967) prenait du surréalisme et de la fiction un cadre principal de création littéraire tandis que la seconde période (1967-1992) était axée sur des fragments d'autobiographie, réflexions sur la littérature et médiations géographiques. Il est important de souligner que le point commun entre ces deux périodes artistiques de l'auteur est l'omniprésence du paysage (fictif ou réel) où la lisière constitue chez Gracq un tremplin qui fait passer l'intrigue du récit d'un endroit clos (forêt ou château), à un autre ouvert (rivière ou mer). La lisière alors incarne un climat d'attente ou une sorte de projection confuse et ambiguë vers l'avenir.

Ceci dit, dans ce travail nous essayerons de mettre en exergue l'accroissement lexical ou comment le vocabulaire progresse et régresse à travers les dix-sept ouvrages de Julien Gracq qui sont : (*Au château d'Argol* (1938), *Un beau ténébreux* (1945), *Liberté grande* (1945), *Le Roi pêcheur* (1948), *André Breton, aspects d'un écrivain* (1948), *Le Rivage des syrtes* (1951), *Penthésilée* (1954), *Un balcon en forêt* (1958), *Préférences* (1961), *Lettrines 1* (1967), *La Presqu'île* (1970), *Lettrines 2* (1974), *Les eaux étroites* (1976), *En lisant en écrivant* (1980), *La forme d'une ville* (1985), *Autour de sept collines* (1988) et *Carnets du grand chemin* (1992).)

Selon Charles Muller : « *On admet qu'à tout moment de l'acte de discours, le locuteur a un certain lexique « en jeu », dans lequel il puise et dont chaque unité a une probabilité d'emploi non nulle. L'étendue et la structure de ce lexique déterminent l'apport lexical, qui est plus grand si le lexique est plus étendu, mais qui est plus faible si les probabilités d'emploi sont inégalement réparties* » (Muller, *Langue* 130).

L'étude de l'accroissement lexical détermine l'apport du vocabulaire au fil du temps ; cet accroissement est, pour un segment déterminé du texte, le nombre d'unités nouvelles qui n'avaient pas été employées antérieurement, qui apparaissent dans ce segment. Pour effectuer les calculs nécessaires de l'accroissement lexical, nous devons découper notre corpus en tranches. Dans cette étude, nous adopterons une répartition qui suit un découpage dit naturel du corpus c'est-à-dire que ce découpage s'effectue par rapport au nombre d'ouvrages que le corpus contient. Par conséquent, nous étudierons l'accroissement lexical du corpus de Gracq dans dix-sept tranches où chaque tranche représente un des ouvrages de la production de l'auteur.

En effet, les différents calculs de l'accroissement du lexique sont élaborés par le logiciel Hyperbase. Celui-ci s'occupe aussi des graphiques et histogrammes qui correspondent à chaque tableau de calcul lexical. Le logiciel Hyperbase donc, dans un premier temps, effectue les calculs d'accroissement lexical normal selon l'ordre chronologique de la production de chaque ouvrage du corpus. Ensuite, il calcule l'accroissement lexical des différentes parties du corpus en suivant un ordre chronologique inverse où il prend le dernier ouvrage de Gracq intitulé *Cahiers du grand chemin* comme point de départ de calcul et conséquemment le premier ouvrage *Au Château d'Argol* deviendra le dernier ouvrage à traiter son accroissement lexical.

## II. ACCROISSEMENT LEXICAL EN ORDRE NORMAL

Le logiciel Hyperbase assure les différents calculs nécessaires à l'étude de l'accroissement lexical dans les dix-sept ouvrages qui composent l'œuvre de Julien Gracq. L'élaboration de ces calculs fait appel à un ajustement des deux séries parallèles (vocabulaire cumulé et étendue cumulée) grâce à une fonction-puissance de type :  $y = ax^b$  pour  $x =$  vocabulaire cumulé et  $y =$  étendue cumulée théorique. L'écart entre étendue théorique et étendue réelle est alors calculé pour chaque texte, puis pondéré par l'étendue de chaque texte comme le montre le tableau suivant :



| ACCROISS. CHRONO | Acc  | Vocab | VocCum | Occur  | OccCum | Ecart     | Pondéré |
|------------------|------|-------|--------|--------|--------|-----------|---------|
| Argol            | 6297 | 6297  | 6297   | 36464  | 36464  | -2139.16  | -0.59   |
| Ténébreux        | 6421 | 9790  | 12718  | 73161  | 109625 | 5973.02   | 0.82    |
| Liberté          | 1881 | 4879  | 14599  | 19983  | 129608 | 10015.83  | 5.01    |
| Pêcheur          | 1239 | 3943  | 15838  | 29055  | 158663 | -7736.85  | -2.66   |
| Breton           | 2855 | 7108  | 18693  | 48229  | 206892 | 5426.57   | 1.13    |
| Syrtes           | 4757 | 12711 | 23450  | 124998 | 331890 | -22224.11 | -1.78   |
| Penthésilé       | 961  | 3805  | 24411  | 33093  | 364983 | -10384.71 | -3.14   |
| Balcon           | 2390 | 8434  | 26801  | 65379  | 430362 | -6164.42  | -0.94   |
| Préférence       | 3071 | 10923 | 29872  | 80421  | 510783 | 1263.10   | 0.16    |
| 1Lettrines       | 2083 | 8571  | 31955  | 44863  | 555646 | 14030.33  | 3.13    |
| Presque          | 1488 | 8161  | 33443  | 58002  | 613648 | -14241.82 | -2.46   |
| 2Lettrines       | 3256 | 12028 | 36699  | 68159  | 681807 | 32406.66  | 4.75    |
| Étroites         | 289  | 2948  | 36988  | 11644  | 693451 | -2403.37  | -2.06   |
| Lisant           | 3066 | 12734 | 40054  | 94142  | 787593 | 6993.46   | 0.74    |
| Ville            | 1675 | 8946  | 41729  | 52718  | 840311 | 4897.52   | 0.93    |
| Collines         | 841  | 5338  | 42570  | 22929  | 863240 | 6620.33   | 2.89    |
| Chemin           | 2419 | 12780 | 44989  | 75914  | 939154 | 11368.17  | 1.50    |

Fonction  $y=a(x \text{ exposant } b)$ :  $a=1.18584999682773e-002$   $b=1.70080320218268$   
 $r^2=0.997262023244488$   $r=0.998630073272625$

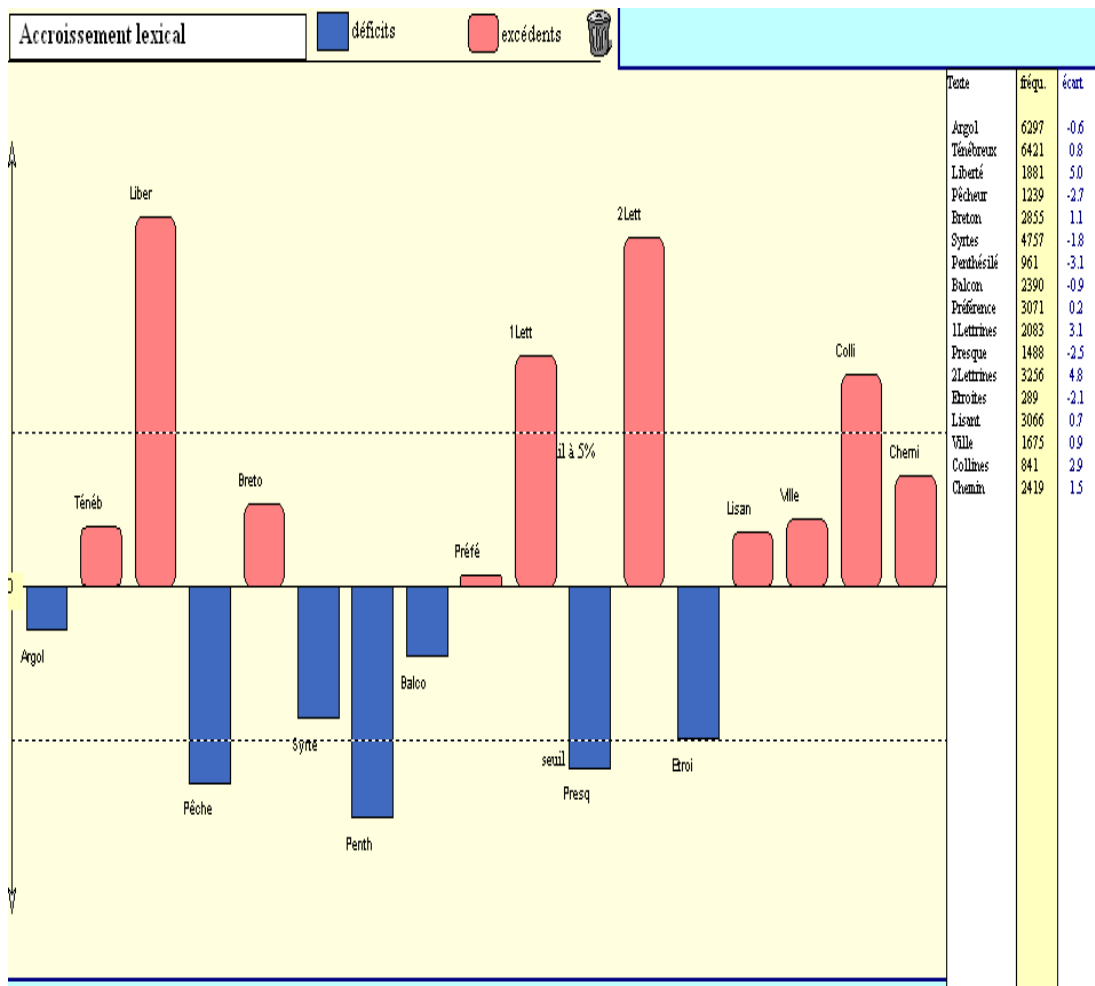
**Tableau N°1 : l'accroissement lexical de corpus Gracq à partir des formes**

En principe, selon les expériences de Brunet et celles de Muller, l'accroissement du vocabulaire diminue en passant du début de l'œuvre vers sa fin. Alors, il a fallu que les textes (*André Breton, Le Rivage des Syrtes, Un Balcon en forêt, Préférences, Lettrines 2, En lisant en écrivant et Carnets du grand chemin*) ayant des valeurs d'accroissement inférieures, les raisons qui sont derrière cette chute sont tributaires essentiellement du genre littéraire dans le sens où la production de Gracq est protéiforme : elle regroupe les romans, les essais, les pièces de théâtre et un recueil de poèmes. Le texte : *André Breton* fait exception de ce que nous venons de signaler puisque l'accroissement lexical est 2855 après avoir chuté dans *Le Roi pêcheur* (1239).

Parmi les raisons qui expliquent la progression des valeurs d'accroissement dans ces textes, nous soulevons la question des hapax dans les textes concernés. Après avoir jeté un œil sur les résultats obtenus dans l'étude des hapax<sup>3</sup> à travers l'œuvre de Gracq, nous avons trouvé que l'effectif des hapax augmente brusquement dans les textes où il y a un accroissement lexical anormal : *André Breton, Le rivage des Syrtes, Un Balcon en forêt, Préférences, Lettrines 2, En lisant en écrivant et Carnets du grand chemin*. Nous constatons alors que la présence des hapax dans ces textes explique nos remarques concernant le mouvement de l'accroissement lexical dans ces textes. Sachons que les hapax constituent un bon signe de mots rares ou spécifiques dans un texte. Les hapax et l'accroissement lexical se complètent pour trancher sur l'originalité du vocabulaire d'un texte.

En outre, le phénomène de l'accroissement nous permet de mettre en place les variations stylistiques qui se manifestent au sein d'un corpus.

En ce qui concerne l'œuvre de Julien Gracq, ces variations stylistiques se dégagent manifestement de la représentation graphique ci-dessous qui permet une observation plus aisée de l'accroissement lexical :



Graphique N° 1 : l'accroissement lexical du corpus des formes de Gracq

D'après ce graphique, nous constatons que l'accroissement lexical suit un mouvement normal c'est-à-dire les effectifs diminuent en fonction de la chronologie dans les couples : (Lib, Pê), (Br, Sy), (Sy, Pe), (L1, Pi), (L2, Et) et (Co, Ch). Dans les couples de textes, l'écart réduit baisse quand nous partons immédiatement d'un texte à un autre: (5 vs -2,7), (1,1 vs -1,8), (-1,8 vs -3,1), (3,1 vs -2,5), (4,8 vs -2,1) et (2,9 vs 1,5). Ainsi, Pê, Sy, Pe, Pi, Et et Ch ont un accroissement normal par rapport aux textes qui précèdent chacun d'eux. Par ailleurs, nous enregistrons qu'il y a des textes où les valeurs augmentent alors au moment qu'ils devraient diminuer.

Le vocabulaire s'accroît dès le début de la production et brusquement il atteint la valeur la plus significative du graphique dans le recueil de poèmes intitulé *Liberté Grande*. Puis nous remarquons une chute spectaculaire du vocabulaire avec la pièce de théâtre : *Le Roi pêcheur*. Cette décroissance du vocabulaire au niveau du théâtre est justifiée par l'abondance des mots de liaison et aussi la répétition qui caractérise le genre théâtral : il s'agit de la répétition au niveau de mots grammaticaux et des mots lexicaux. Avec l'essai sur André Breton, nous remarquons une augmentation significative du vocabulaire et ensuite il y a une pauvreté remarquable du vocabulaire qui va régner jusqu'à la fin de la première période littéraire c'est-à-dire jusqu'à l'ouvrage intitulé *Préférences*.

Les valeurs déficitaires de l'accroissement lexical qui caractérisent la fin de la première période signifient que notre auteur use seulement son vocabulaire ancien et par conséquent l'apport de nouvelles formes est presque absent et nous déduisons qu'il n'y a pas de nouveaux thèmes à soulever dans cette période. Ainsi, Julien Gracq entame la seconde période de son œuvre avec une richesse très significative

du vocabulaire. Le livre intitulé *Lettrines 1* marque un événement décisif dans la production de Gracq dans la mesure où notre auteur adopte un nouveau genre littéraire de cahiers et cela correspond à une nette rupture dans son écriture et à un tournant dans son œuvre. Dès le début de cette seconde période, nous nous trouvons devant une véritable explosion lexicale qui atteint une valeur de **3,1** et qui dépasse largement le seuil de **5%**.

*La Presqu'île*, le second livre de cette seconde période, coïncide avec une chute considérable au niveau de vocabulaire. Ce livre s'accapare de la valeur la plus faible de l'accroissement lexical dans la seconde période. Ce changement de l'accroissement lexical est justifié par le changement du genre littéraire puisque *La presqu'île* s'inscrit parmi les récits de corpus.

Il faut signaler que, à travers les textes de corpus, chaque changement de genre littéraire s'accompagne d'un changement considérable au niveau de l'accroissement lexical. Ce changement pourrait être positif aussi bien que négatif. Pour les cas positifs nous citons les changements effectués lors du passage d'*Un beau ténébreux* à *Liberté grande*, du *roi pêcheur* à *André Breton*, d'*Un balcon en forêt* à *Préférences*, de *Préférences* à *Lettrines 1*, de *La Presqu'île* à *Lettrines 2* et le changement effectué lors du passage des *eaux étroites* à *En lisant en écrivant*. En ce qui concerne les cas négatifs, nous soulignons les changements observés lors du passage de *Liberté grande* au *Roi pêcheur*, d'*André Breton* au *Rivage des Syrtes*, de *Lettrines 1* à *La Presqu'île*, de *Lettrines 2* aux *eaux étroites*. En somme, tous les ouvrages de cette période sont riches, à part deux qui sont considérés comme pauvres à savoir : *La Presqu'île* et *Les eaux étroites*.

### III. ACCROISSEMENT DU VOCABULAIRE (ORDRE INVERSE) :

Etienne Brunet propose une méthode inversible de l'accroissement lexical, il s'agit de l'accroissement opposé qui consiste toujours à examiner l'apport lexical dans chaque texte du corpus mais en prenant en considération comme point de départ le dernier texte de ce corpus.

En effet, le logiciel Hyperbase fournit également les calculs nécessaires à l'étude de l'accroissement lexical en ordre inverse tout en suivant le même principe du calcul<sup>4</sup> utilisé lors de l'accroissement normal.

Dans cette optique, Brunet dit : « *Les deux visées sont en fait complémentaires, la première désigne les thèmes qui apparaissent pour la première fois, la seconde ceux qui n'apparaîtront pas. La première marque le jaillissement d'une inspiration, son épuisement* » (Brunet, *Le vocab.* 81).

Dans cette analyse, nous adopterons le même découpage naturel du corpus utilisé lors l'étude de l'accroissement lexical normal qui suit l'ordre chronologique de la production de chaque texte de Gracq et nous nous baserons essentiellement sur les formes lexicales.

En se basant sur les fonctionnalités du logiciel Hyperbase, nous pouvons présenter les résultats de l'accroissement lexical inverse des formes graphiques selon le tableau suivant :

| ACCROISS. INVERSE | Acc Vocab  | VocCum | Occur  | OccCum | Ecart     | Pondéré |
|-------------------|------------|--------|--------|--------|-----------|---------|
| Chemin            | 2780 12780 | 12780  | 75914  | 75914  | -7055.61  | -0.93   |
| Collines          | 2316 5338  | 15096  | 22929  | 98843  | 4419.27   | 1.93    |
| Ville             | 3899 8946  | 18995  | 52718  | 151561 | 3680.28   | 0.70    |
| Lisant            | 5508 12734 | 24503  | 94142  | 245703 | 7716.09   | 0.82    |
| Étroites          | 502 2948   | 25005  | 11644  | 257347 | -1067.12  | -0.92   |
| 2Lettrines        | 4010 12028 | 29015  | 68159  | 325506 | 24095.49  | 3.54    |
| Presque           | 2141 8161  | 31156  | 58002  | 383508 | -3089.72  | -0.53   |
| 1Lettrines        | 1941 8571  | 33097  | 44863  | 428371 | 8325.21   | 1.86    |
| Préférence        | 2566 10923 | 35663  | 80421  | 508792 | -5132.64  | -0.64   |
| Balcon            | 1842 8434  | 37505  | 65379  | 574171 | -7839.73  | -1.20   |
| Penthésilé        | 800 3805   | 38305  | 33093  | 607264 | -7193.31  | -2.17   |
| Syrtes            | 2677 12711 | 40982  | 124998 | 732262 | -34323.14 | -2.75   |
| Breton            | 1038 7108  | 42020  | 48229  | 780491 | -11409.06 | -2.37   |
| Pêcheur           | 468 3943   | 42488  | 29055  | 809546 | -12150.43 | -4.18   |
| Liberté           | 606 4879   | 43094  | 19983  | 829529 | 2186.70   | 1.09    |
| Ténébreux         | 1212 9790  | 44306  | 73161  | 902690 | -27871.98 | -3.81   |
| Argol             | 683 6297   | 44989  | 36464  | 939154 | -10384.33 | -2.85   |

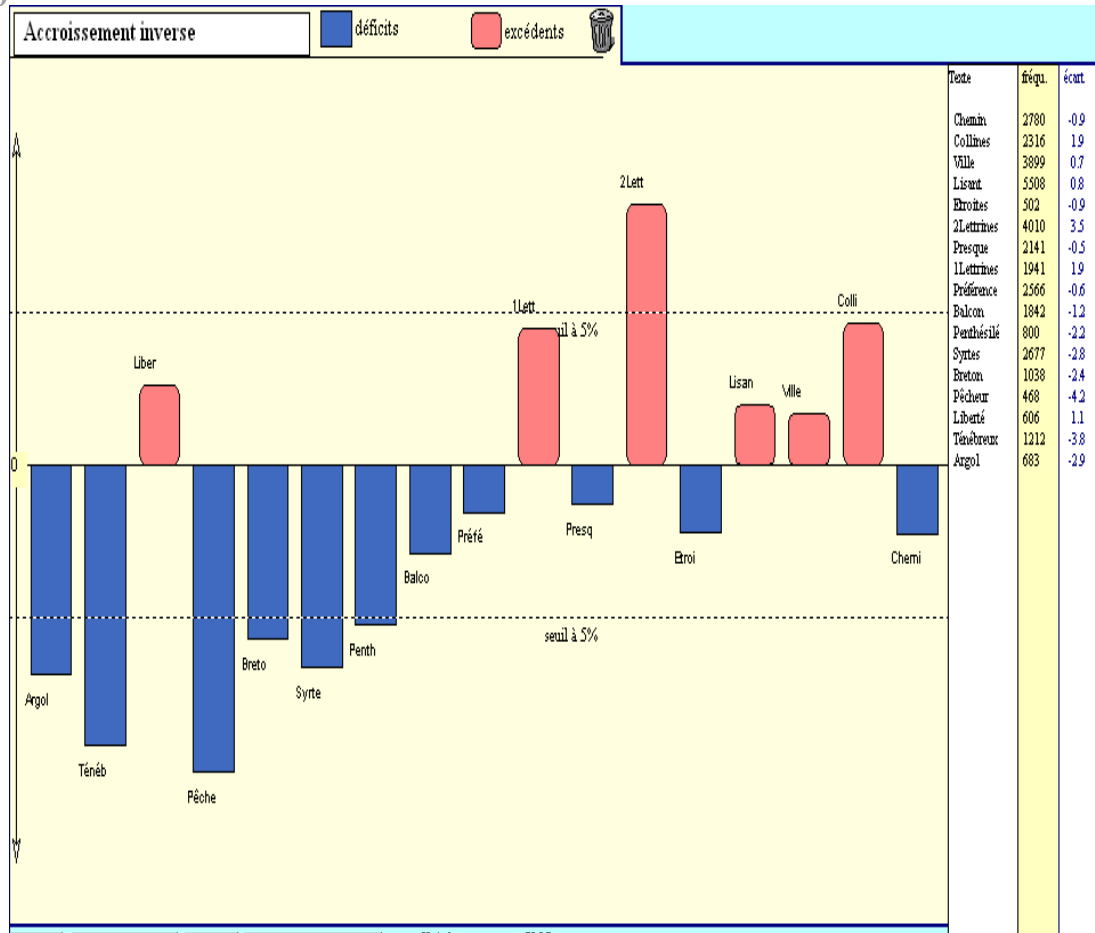
Fonction  $y=a(x \text{ exposant } b)$ :  $a=3.90493868632554e-004$   $b=2.00810438646947$   
 $r^2=0.993883113118819$   $r=0.996936865161891$

**Tableau N° 2 : Accroissement inverse du vocabulaire dans l'œuvre de Gracq à partir des formes**

En partant du dernier texte de la série, la tendance régulière quand il s'agit de mouvement de l'accroissement lexical est inversée dans six ouvrages à savoir : *La Forme d'une ville* (3899), *En Lisant en écrivant* (5508), *Lettrines 2* (4010), *Préférences* (2566), *Le Rivage des Syrtes* (2677), *Liberté grande* (606) et *Un beau Ténébreux* (1212). En prenant en considération les textes qui précèdent immédiatement chacun d'eux, ces derniers se caractérisent par un apport lexical très important.

Par conséquent, les constats déjà tirés se trouvent là aussi. En effet, ces textes se caractérisent par leur originalité par rapport aux autres textes du corpus. Nous présentons le graphique suivant pour clarifier davantage ce que nous avons constaté par le biais des données quantitatives mentionnées au tableau ci-dessus.

Le graphique suivant illustre les différents résultats de l'accroissement inverse sous forme d'histogramme :



**Graphique N° 2 : Accroissement inverse du vocabulaire dans l'œuvre de Gracq à partir des formes**

Ce graphique, alors, montre que la majorité des ouvrages qui avaient des valeurs positives d'accroissement lexical dans les deux graphiques précédents maintient presque les mêmes valeurs sauf une minorité de quatre ouvrages à savoir : *Un beau ténébreux*, *André Breton*, *Préférences* et *Carnets du Grand chemin*.

Il est évident que l'accroissement lexical de *Carnets du grand chemin* connaît cette régression spectaculaire pour une raison aussi évidente que simple : cet ouvrage fait le point de départ de cette analyse de l'accroissement inverse et par conséquent il n'y a pas de nouveaux vocables à souligner.

Suivant le graphique de l'accroissement lexical en ordre inverse, nous remarquons qu'il y a toujours les mêmes caractéristiques observées lors des deux autres graphiques concernant l'accroissement lexical en ordre normal des œuvres de Gracq. Autrement dit, l'apport lexical dans les ouvrages de la seconde période artistique de Gracq dépasse celui des ouvrages de la première période sauf *Liberté grande* (le seul recueil de poèmes dans l'œuvre gracquienne) qui connaît un apport lexical important.

Après avoir effectué l'accroissement lexical en tenant compte des formes graphiques, nous constatons que les résultats ne changent pas. L'étude de l'accroissement lexical vient pour confirmer la spécificité du vocabulaire des ouvrages de la seconde période puisqu'il s'agit d'un vocabulaire riche en hapax qui donne naissance à des thèmes spéciaux qui émanent spécialement de la formation géographique de l'auteur.

Alain-Michel Boyer confirme ce constat en disant : « *Quoi d'étonnant si le lexique géographique vient spontanément sous la plume de Gracq lorsque, dans Préférences ou Lettrines, il évoque même des*

phénomènes littéraires ? Des expressions comme « gisements profonds », « courbes de niveau », « beauté géodésique », « nappe phréatique profonde » sont relativement fréquentes » (Boyer 23).

Après avoir examiné l'apport de nouveaux mots dans l'œuvre de Gracq en suivant les ruptures naturelles de notre corpus, nous viserons jeter un coup d'œil rapide sur l'accroissement de lexique dans notre corpus mais cette fois-ci en se basant sur des ruptures artificielles c'est à dire des tranches de même taille.

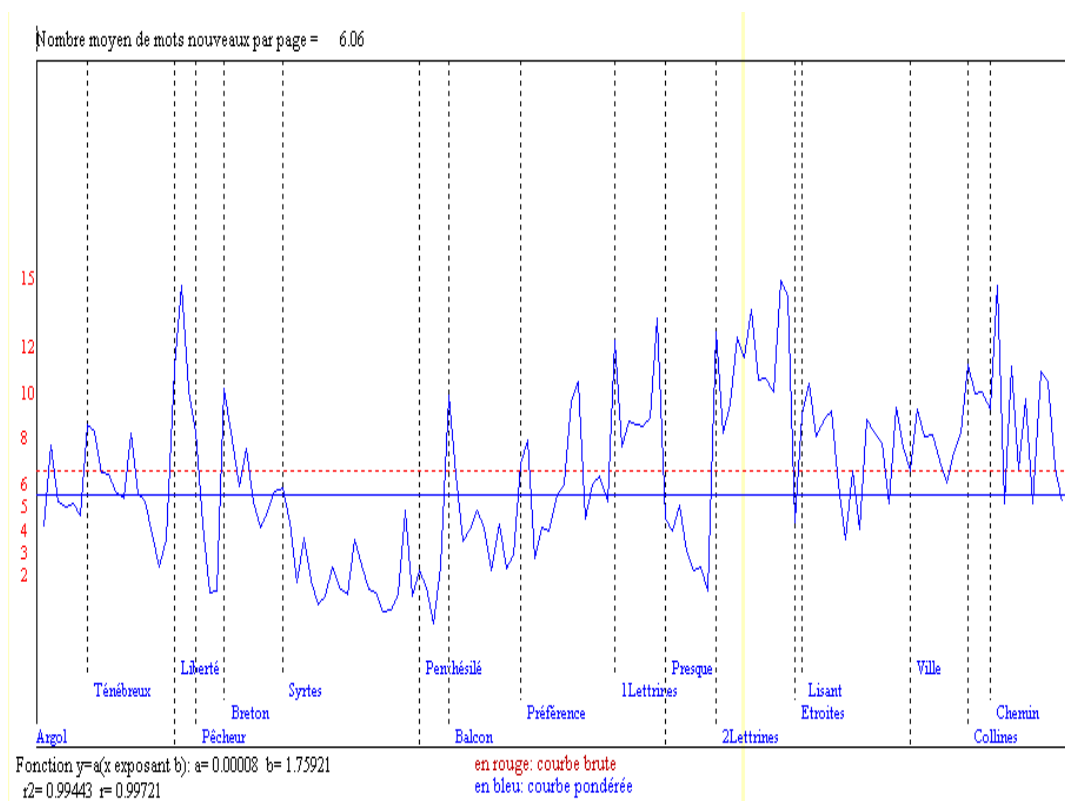
#### IV. L'ACCROISSEMENT LEXICAL MESURÉ A PARTIR DES TRANCHES EGALES

Le logiciel Hyperbase effectue également le calcul de l'accroissement lexical d'une autre manière. Pour établir la mesure, cette fois-ci, le logiciel s'attache à découper le même nombre de mots (1000 mots) et non par les césures naturelles qui se réalisent selon les différents ouvrages d'un corpus.

Cette nouvelle technique de l'examen de l'accroissement lexical consiste à localiser les ruptures thématiques dans le corpus là où se produit un afflux intense de nouveaux vocables (une pente ascendante du segment) ; les tranches où l'accroissement est inférieur aux valeurs théoriques évoquent l'épuisement d'un ou de plusieurs thèmes (une pente descendante du segment).

L'étude de l'accroissement selon des tranches égales se fait selon les deux approches adoptées auparavant : l'accroissement lexical en ordre normal et l'autre en ordre inverse.

ACCROISSEMENT EN ORDRE NORMAL— Le logiciel Hyperbase nous permet de présenter le graphique de l'accroissement lexical sur tranches égales de la manière suivante :



Graphique N° 3 : Accroissement lexical calculé à partir des tranches égales des formes du corpus Gracq

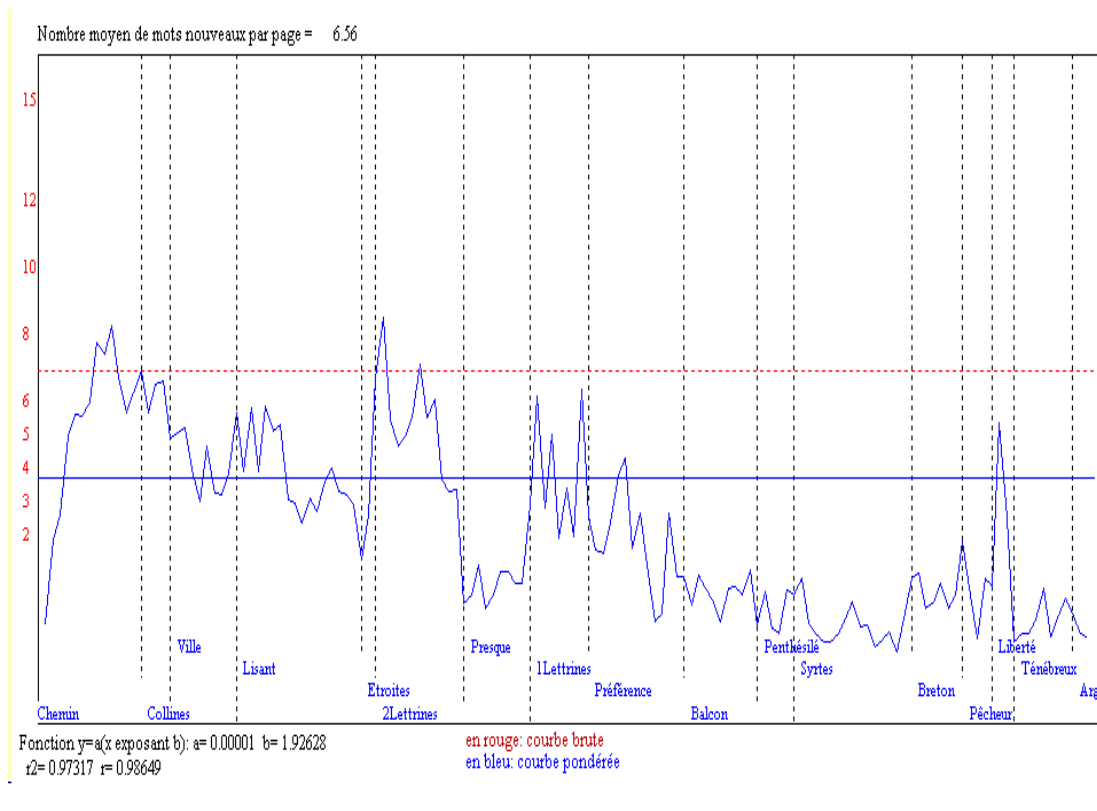
Nous pouvons constater un mouvement semblable à celui observé dans les histogrammes précédents. Ainsi, la courbe, observée dans la figure ci-dessus, nous permet de déduire qu'il y a toujours une ascendante maximale dans les ouvrages jugés riches en mots nouveaux dans l'analyse précédente. Il

s'agit de *Liberté grande*, *Préférences*, *Lettrines 1*, *Lettrines 2*, *En lisant en écrivant*, *Autour de sept collines* et *Carnets du grand chemin*.

La courbe fait encore apparaître les mêmes caractéristiques obtenues lors de la dernière analyse de l'apport lexical dans l'œuvre de Gracq dans la mesure où elle connaît des variations lexicales très significatives à travers tout le corpus. Cette variation provoque une évolution lexicale où se dégagent deux grandes périodes littéraires et chacune d'elles connaît un déséquilibre d'apport lexical qui s'effectue automatiquement quand nous passons d'un ouvrage à un autre. Le passage d'un texte à un autre coïncide souvent avec le changement de genre littéraire puisque le corpus de Gracq, comme nous l'avons signalé, est très varié et contient plusieurs genres.

ACCROISSEMENT EN ORDRE INVERSE— Pour accomplir notre étude d'accroissement lexical en ordre inverse à partir des tranches égales, nous mettons en relief ce qui se passe dans ce sens si nous traitons ce phénomène en commençant par le dernier ouvrage de Gracq.

Le graphique suivant dresse le mouvement de l'accroissement lexical inverse à travers des tranches de 1000 mots :



Graphique N° 4: Accroissement lexical inverse calculé à partir des tranches égales du corpus Gracq

Suivant la courbe qui représente les variations de l'accroissement du vocabulaire en ordre inverse, nous pouvons déduire que nous avons toujours les mêmes résultats obtenus auparavant : les intenses afflux de mots nouveaux s'effectuent dans les ouvrages *Liberté grande*, *Préférences*, *Lettrines 1*, *Lettrines 2*, *En lisant en écrivant*, *Autour de sept collines* et *Carnets du grand chemin*. Ceci a une influence remarquable sur l'apparition de nouveaux thèmes d'écriture.



Pour clore, l'analyse de l'apport lexical dans l'œuvre de Gracq nous permet d'effectuer les interprétations suivantes : au début, l'apport de nouvelles unités dans la seconde période artistique de l'auteur (1967-1992), qui débute avec le cahier intitulé *Lettrines I*, est intense par rapport à celui remarqué lors de la première période littéraire (1938-1967) où l'accroissement était au-dessus de la moyenne à l'exception de la valeur significative d'accoisement observée dans le recueil de poèmes en prose : *Liberté grande*. Ensuite, l'apport lexical à travers l'œuvre de Gracq n'est pas stable puisqu'il change automatiquement à mesure que l'auteur passe d'un genre littéraire à un autre. Enfin, les ouvrages où Gracq décrit des paysages réels qui sont rencontrés pendant sa vie connaissent un apport lexical intense en revanche les ouvrages qui ont été écrits dans le cadre du surréalisme sont jugés faibles au niveau de l'apport de nouveaux mots.

Alain-Michel Boyer justifie en quelque sorte nos interprétations faites à l'issue d'une étude de l'accroissement lexical à travers les dix-sept textes de Julien Gracq en disant :

« Mais ce qui frappe le lecteur de Julien Gracq, c'est la singulière évolution d'une œuvre qui, de livre en livre, du *Château d'Argol*, publié en 1938, à *Carnets du grand chemin*, publié d'un demi-siècle plus tard, suit une courbe, effectue un parcours à rebours, qui conduit du jeu d'écriture surréaliste à une méditation sur Rome et qui accorde un rôle de plus en plus important, non seulement aux configurations spatiales, mais aussi lieux réels » (Boyer 35)

En effet, nous pouvons constater que la seconde période artistique de notre auteur Julien Gracq est plus riche que la première et aussi elle connaît un apport lexical plus intense que celui de la première. Le vocabulaire des ouvrages de la seconde période est très spécial. Cela émane principalement du style d'écriture adopté par l'auteur à la seconde moitié de sa vie littéraire dans la mesure où dans *Lettrines I*, Julien Gracq a inauguré un style d'écriture qui échappe à une définition classique, il ne paraît pas exagéré de penser qu'il renouvelle une forme d'expression originale que d'autres écrivains vont emprunter après lui. A vrai dire, les ouvrages de la seconde période s'inscrivent dans le cadre de ce que nous appelons communément 'Littérature en fragment' qui se définit comme un ensemble très libre, une mosaïque de mots de lecture, de réflexions et de souvenirs (d'enfance souvent).

#### NOTES

- [1] Chaque représentation graphique se fait suivant l'un des paramètres de l'époque, du genre littéraire, etc.  
 [2] Ce logiciel, développé par Etienne Brunet, est destiné au traitement des textes littéraires. Il est considéré comme le fruit de l'évolution des méthodes de la statistique lexicale depuis son existence en France vers la fin des années cinquante.  
 [3] Mot utilisé une seule fois dans un texte c'est-à-dire sa fréquence d'emploi est 1.  
 [4] Ce principe fait référence à l'équation :  $y = ax^b$ .

#### BIBLIOGRAPHIE

- [1] Boyer, Alain. *Paysages et mémoires /des eaux étroites à un balcon en forêt*. Nantes : Editions Cécile défaut, 2007.  
 [2] Brunet, Etienne. *Le vocabulaire français de 1789 à nos jours*. Genève-Paris : Slatkine-Champion, 1981.  
 [3] Brunet, Etienne. « *Le traitement des faits linguistiques et stylistiques. Texte d'application : Giraudoux* ». *Statistiques et linguistique*, Actes et colloque, Klincksieck, n. 15, 1974, pp. 105-136.  
 [4] Brunet, Etienne. « *La structure lexicale dans l'œuvre de Hugo* ». *Etudes sur la richesse et la structure lexicales*, Paris-Genève, Champion-Slatkine, 1988, pp. 23-42.  
 [5] Brunet, Etienne. *Hyperbase, Logiciel hypertexte pour le traitement documentaire et statistique des corpus textuels, Manuel de Référence*. Version 7,1 pour Windows, 2006.  
 [6] Canovas, Frédéric. *L'écriture révée*. Paris : l'Harmattan, 2000.  
 [7] El Bakkali, El Arbi. « *Etude de statistique lexicale : quelques aspects ambigus des couleurs dans l'œuvre littéraire de Julien Gracq* ». *Ambiguïtés, Actes du XXXIII colloque d'Albi, Langages et Signification*, Le Mirail, Toulouse, Janvier 2012, pp. 253-264.  
 [8] Gracq, Julien. *Au château d'Argol*. Paris : José Corti, 1938.  
 [9] Gracq, Julien. *Un beau ténébreux*. Paris : José Corti, 1945.  
 [10] Gracq, Julien. *Liberté grande*. Paris : José Corti, 1946.

*Archive of SID*

- [11] Gracq, Julien. *Le roi pécheur*. Paris : José Corti, 1948.
- [12] Gracq, Julien. *André Breton*. Paris : José Corti, 1948.
- [13] Gracq, Julien. *La littérature à l'estomac*. Paris : José Corti, 1950.
- [14] Gracq, Julien. *Le rivage des Syrtes*. Paris : José Corti, 1951.
- [15] Gracq, Julien. *Penthésilée*. Paris : José Corti, 1954.
- [16] Gracq, Julien. *Un balcon en forêt*. Paris : José Corti, 1958.
- [17] Gracq, Julien. *Préférences*. Paris : José Corti, 1961.
- [18] Gracq, Julien. *Lettrines 1*. Paris : José Corti, 1967.
- [19] Gracq, Julien. *La Presqu'île*. Paris : José Corti, 1974.
- [20] Gracq, Julien. *Lettrines 2*. Paris : José Corti, 1974.
- [21] Gracq, Julien. *Les eaux étroites*. Paris : José Corti, 1976.
- [22] Gracq, Julien. *En lisant en écrivant*. Paris : José Corti, 1980.
- [23] Gracq, Julien. *Autour des sept collines*. Paris : José Corti, 1988.
- [24] Gracq, Julien. *Carnets du grand chemin*. Paris : José Corti, 1992.
- [25] Gracq, Julien. *Entretiens*. Paris : José Corti, 2002.
- [26] Guiraud, Pierre. *Problèmes et méthodes de la statistique*. Paris, P.U.F,1960.
- [27] Kastberg-Sjoblomilf, Margarita. *L'écriture de J.M.G. Le Clézio, Une approche lexicométrique*. Thèse de doctorat, Université de Nice-Sophia Antipolis, 2007.
- [28] Michel, Jacqueline. *Une Mise en récit du silence, Le Clézio-Bosco-Gracq*. Paris : José Corti, 1986.
- [29] Muller, Charles. *Initiation à la statistique linguistique*. Paris : Librairie Larousse, 1968.
- [30] Muller, Charles. *Langue française et linguistique quantitative*. Recueil d'articles, Genève : Editions Slatkine, 1979.