



**HAL**  
open science

## Le Vocabulaire de Giraudoux. Introduction

Étienne Brunet

► **To cite this version:**

Étienne Brunet. Le Vocabulaire de Giraudoux. Introduction. Céline Poudat. Ce qui compte, Champion, 2011, 978-2-7453-2225-8. hal-01575144

**HAL Id: hal-01575144**

**<https://hal.science/hal-01575144>**

Submitted on 17 Aug 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## Le Vocabulaire de Giraudoux. Introduction<sup>1</sup>

Etienne Brunet

Giraudoux n'a jamais vu d'ordinateur : il n'a vu que les lieux où allaient naître, trente ans après son passage à Harvard, le dernier calculateur mécanique et le premier électronique, le Fameux Mark II d'Aiken. A-t-il au moins connu la naissance du « messie » moderne ? C'est peu probable en ces années 40 où les bottes allemandes résonnaient trop douloureusement dans son cerveau pour que pussent l'atteindre les vagissements des monstres que la technique enfantait, monstres peu rassurants au reste, comme cet ENIAC de trente tonnes qui apparut à Philadelphie l'année où Giraudoux disparaissait, et qui servait à calculer des tables de tir. Eût-il connu l'existence de cette machine, Giraudoux eût pu difficilement la reconnaître comme instrument de progrès. Plus vraisemblablement l'eût-il classée parmi les engins de mort comme les V2 qui virent le jour à la même époque, ou la bombe atomique qui allait exploser l'année suivante.

D'ailleurs Giraudoux ne voue pas aux mathématiques et à leurs applications un culte aussi fervent que Valéry. Dans son musée des grands hommes, les savants ne sont certes pas ignorés ; Pasteur, Claude Bernard, Einstein<sup>2</sup>, reçoivent leur juste lot de vénération et dans le

---

<sup>1</sup> Ce texte est l'introduction de la thèse d'Etienne Brunet *Le vocabulaire de Giraudoux Structure et Evolution*, Slatkine, Genève, 1978, p 1-19. Il a été repris et publié par Céline Poudat dans *Ce qui Compte*, Champion, 2011, p. 27-43.

<sup>2</sup> « C'étaient les dernières réponses à Einstein, à Bergson et à d'autres auxquels il n'avait jamais été répondu aussi nettement encore, à Darwin, à Spencer... » (*Bella*, p. 19).

Panthéon intime de la Cour de l'École Normale, Giraudoux sait reconnaître le buste de Lavoisier ou de Cuvier<sup>3</sup>.

Mais plus souvent son attention se fixe sur les écrivains, les artistes et les philosophes, sur Platon et Montaigne, Racine et Baudelaire, Pascal et Bergson, ceux qu'unit à lui un lien de famille.

« C'est ce spectacle de famille qu'offre la Cour d'Honneur elle-même de l'École, où, de leurs loggias, les bustes de Racine, de Pascal, de Montaigne et de trente autres écrivains contemplant... »<sup>4</sup>.

De fait, la route de Giraudoux n'a jamais croisé le chemin de Damas, il n'a jamais hésité devant cette borne Michelin qui dirigeait d'un côté les Renan et les Valéry, de l'autre les Péguy et les Claudel. Pas de conversion religieuse chez lui, ni de reconversion scientifique. Du monde antique des humanités au monde moderne des réalités, il glisse sans effort et sans reniement. Sans doute les acquisitions de la curiosité ont quelque peu étoffé le très léger bagage scientifique dont l'école l'avait pourvu. Mais il ne semble pas que Giraudoux ait demandé à la Science la clé de l'univers : l'explication du monde et de ses lois lui paraissait chose trop sérieuse pour être abandonnée aux hommes de science. Et nous avons quelque peine à suivre M. David Aurel quand il fait de Giraudoux un visionnaire qui cinquante ans à l'avance aurait eu le pressentiment des récentes acquisitions de la science : greffes d'organes, prothèses artificielles et cette prothèse du cerveau qu'est l'ordinateur<sup>5</sup>.

Mais Giraudoux qui aimait les jeux de logique où la raison fait la guerre au hasard<sup>6</sup>, eût sans doute apprécié les vertus ludiques du jouet logique qu'est l'ordinateur, jouet sans hasard et sans partenaire, où le jeu élimine tout jeu, tout interstice où se glisseraient le probable, le possible

---

<sup>3</sup> Juliette traverse « une première cour carrée où le passant était épié par les soixante bustes des grands hommes qui surent le mieux observer, Lavoisier, Cuvier ou Chevreul » (*Juliette au Pays des Hommes*, p. 31).

<sup>4</sup> *Littérature*, p. 165.

<sup>5</sup> « Seul Giraudoux s'est maintenu au point où la métaphysique se transforme en visions cosmologiques, puis en véritables hypothèses scientifiques, cinquante ans environ avant que les véritables hommes de science n'arrivent à ce même point » (David Aurel. *Vie et Mort de Jean Giraudoux*, p. 10).

<sup>6</sup> En découvrant le royaume du bridge, Giraudoux prend conscience « d'avoir enfin découvert, en ce bas monde, un pays stable : celui du jeu. Les châteaux de cartes se sont révélés autrement solides, ces derniers temps, que les burgs et les citadelles. » (*Or dans la nuit, Préface à l'Essentiel du Bridge*, p. 63)

ou l'erreur, jouet intellectuel et presque immatériel inventé par Boole, Babbage ou Turing bien avant que la technologie ne le réalise.

Du moins osons-nous espérer qu'il se fût prêté de bonne grâce à ce jeu, avec cette curiosité et cette ouverture qui lui faisaient priser les jeux du stade comme ceux de l'esprit, et les jeux de société comme ceux de la technique. Lui qu'on a vu si peu effarouché au milieu des canons et des machines de guerre, si vite familier avec la machinerie du théâtre et de la mise en scène, avec les micros, les caméras et tout l'outillage de la radio et du cinéma, eût sans doute été fort à l'aise auprès des machines à calculer.

À cet instrument terrifiant qu'on présente parfois comme la machine à violer le secret des citoyens, Giraudoux aurait livré volontiers en pâture son secret d'écrivain, avec peut-être ce sourire amusé, voire ironique, qu'il opposait à la psychanalyse et qui n'a nullement paralysé l'enquête si fine et si profonde de Charles Mauron<sup>7</sup>. Tout porte même à croire qu'il eût préféré le viol auquel se livrent les machines et qui du moins ne s'adresse qu'à ce qui est public, plutôt que l'intrusion d'une critique historiciste dans sa vie, qu'à la suite de Valéry il a tenue aussi éloignée que possible de l'œuvre. Quand lui-même se fait critique, il s'abstient de lier la vie de Racine à sa production littéraire, d'expliquer Hermione par la Champmeslé et Phèdre par le Jansénisme. « Son amertume quand il est amer, dit-il de Racine, ne vient pas de ce qu'il est trompé ou boiteux, sa douceur de ce qu'il est en paix, sa vigueur de ce qu'il est herculéen, – mais de ce qu'il est écrivain »<sup>8</sup>. Puisque Racine est un homme de « métier », non « d'inspiration », un artiste conscient « dominant ses tragédies comme un potier ses vases », « les explications ou les commentaires littéraires sont les seuls qui permettent d'approcher et les seuls qu'eût volontiers écoutés le poète »<sup>9</sup>.

Ce sont sans doute les seuls, aidés ou non par les machines, qu'eût acceptés pour son compte ce précurseur du structuralisme qu'est Giraudoux<sup>10</sup>.

---

<sup>7</sup> Charles Mauron. *Le théâtre de Giraudoux. Étude psychocritique*. José Corti, 1971.

<sup>8</sup> *Littérature*, p. 31.

<sup>9</sup> *Littérature*, p. 37.

<sup>10</sup> Si Giraudoux ignore – et pour cause – l'informatique, il cite sept fois la statistique dans notre corpus, c'est-à-dire plus souvent que la moyenne des écrivains du XX<sup>e</sup> siècle. Il est vrai que la manière d'en parler est plutôt ironique quoique non vindicative. Ainsi on trouve dans *Églantine*, page 41 : « Il pouvait indiquer chaque soir, avec plus de précision que le bureau des statistiques, la proportion des étrangers, des nouveaux étrangers dans

À notre décharge et pour mériter l'indulgence des mânes de Giraudoux, nous affirmons que notre dessein ne fut nullement prémédité et qu'à origine nous prétendions pénétrer dans l'œuvre de Giraudoux autrement que par effraction et par l'informatique et sans ces chars d'assaut qui prennent les citadelles mais détruisent les jardins. Nous voulions explorer le jardin giralducien, ses zones claires et ses zones d'ombre, sans casser la moindre tige. Or notre étude présente ne ressemble en rien à celle que nous avons envisagée au départ. Notre projet initial promettait de saisir l'esprit de Giraudoux. Nous avons disposé pièges et lacets pour attraper le gibier. Et Giraudoux se laissait prendre facilement. Trop facilement. Car chaque soir nous déversions notre gibecière pleine dans le Grand Fichier qui contint bientôt enfermé tout l'esprit de Giraudoux. Enfermé mais aussi refermé : le secret n'avait fait que changer de place comme une valise diplomatique qu'on transporte sans l'ouvrir. Les phrases avaient quitté leur contexte souriant et se retrouvaient emprisonnées dans une boîte sévère mais elles refusaient de passer aux aveux et le tortionnaire n'avait d'autre recours que de reprendre sa quête le lendemain. L'esprit, malin, se déroba. Notre vue brouillée le voyait, tantôt avec la majuscule, tantôt avec la minuscule, dissimulé dans le moindre mot, dans le moindre recoin de l'œuvre. Il fallut renoncer. Depuis lors le fichier est resté clos : l'esprit de Giraudoux y serait encore s'il n'était d'essence volatile.

Comme les chimistes fixent un corps trop léger dans un précipité stable et solide, plus aisément analysable, de même avons nous pensé que le style de Giraudoux constituait un biais qui pût donner la mesure de son esprit. Déjà nous tourmentait l'idée de mesure et le désir de saisir un objet quantifiable. Les travaux de M. Guiraud, de M. Muller et de G. Herdan montraient les vertus des méthodes statistiques dans le domaine du langage, comme dans toutes les sciences humaines. Et nous nous propositions ingénument d'appliquer ces méthodes, avec l'outil informatique<sup>11</sup>, à la caractérisation d'un style. Il fallut déchanter. Non que

---

Paris... », dans *Suzanne*, p. 151 : « Ainsi, par la loi des probables et des moyennes, c'était un Anglais que m'apportait la mer. » et dans *Juliette*, p. 42 : « j'ai manqué ma vie par ma faute, avoua M. By. Il m'eût suffi peut-être, pour être heureux, de consulter les statistiques... La statistique m'a vaincu ». Bien loin que penser que la statistique ait vaincu Giraudoux, nous doutons encore qu'elle l'eût convaincu, mais il la connaissait mal et cette science n'en était d'ailleurs qu'à ses débuts.

<sup>11</sup> Nous avons rencontré quelqu'un, Michel Launay, qui avait rencontré l'ordinateur. Nous remercions le hasard d'avoir ménagé ces rencontres et Michel Launay de s'y être prêté.

les méthodes statistiques fussent décevantes et nous ne voulons pas reprendre ici les mises au point vigoureuses et rigoureuses de Pierre Guiraud<sup>12</sup> et Charles Muller<sup>13</sup>. Mais nous nous sommes aperçu que l'ordinateur était loin de posséder la compétence linguistique et stylistique que nous lui prêtions. Et la performance n'est rien quand la compétence manque. Le réseau serré des tores et des circuits était bien incapable de retenir la moindre parcelle de l'humour giraldouzien et la plus évidente des figures de style.

Nos prétentions baissèrent encore et dans un accès de pessimisme nous envisageâmes presque de réduire notre recherche à l'étude de la virgule chez Giraudoux, prenant au sérieux telle plaisanterie de P. Guiraud, qui propose l'étude de l'accent circonflexe dans le Code Civil.

## L'INFORMATIQUE LINGUISTIQUE

Nous nous arrê tâmes cependant au palier lexical. D'une part l'exemple de Muller donnait l'assurance qu'il n'était pas nécessaire de reculer plus arri ère, d'autre part la fréquentation plus assidue des centres de calcul nous avait appris qu'à ce niveau un terrain de rencontre pouvait être établi entre la machine et nous. Celle-ci savait un peu lire, et nous savions un peu compter. Elle connaissait les vingt-six lettres de l'alphabet, nous connaissions les quatre opérations et même la règle de trois. Nous nous installâmes en ménage, non sans scène. La machine a un caractère exigeant<sup>14</sup>, très pointilleux, maniaque. Elle ne supporte aucune erreur, aucun désordre, aucun à-peu-près, aucune lacune, aucun raccourci. Rien n'est jamais implicite ou évident pour elle, elle dénèche toujours les cas imprévus et malencontreux et vient systématiquement buter dans des impasses minuscules quand on lui ouvre des boulevards. Elle est

---

<sup>12</sup> « Dans un domaine où l'industrie des grammairiens s'est immensément et constamment exercée, on n'en est sur bien des points guère plus avancé que Cicéron. On a décrit plus de cent cinquante espèces de métaphores, on a examiné l'espace de pratiquement chaque mot de la langue écrite et parlée et il semble bien qu'on soit dans une impasse faute d'un système de mesure sans lequel il n'y a pas de science possible. Or cet appareil, la statistique le fournit au linguiste car le langage est un phénomène essentiellement statistique, c'est-à-dire soumis à des constantes et à des lois numériques et susceptibles, à ce titre, de définitions et d'interprétations quantitatives ». (Pierre Guiraud. *Problèmes et Méthodes de la statistique linguistique*, p. 16)

<sup>13</sup> Charles Muller. *Étude de statistique lexicale*, p. 9 à 14.

<sup>14</sup> La machine doit s'irriter fort qu'on la compare une fois de plus à une personne et qu'on lui prête des sentiments.

exigeante aussi pour les conditions de travail : elle a des sautes d'humeur par temps d'orage, elle travaille à son rythme qui n'est que celui de la lumière et ne précipite nullement la cadence même lorsque la queue des utilisateurs s'allonge et s'impatiente dans la salle d'attente. Elle a horreur de la poussière, comme ces sous-officiers sourcilleux qui passent la revue des armes en gants blancs. Elle a parfois rejeté nos bandes y décelant des impuretés infinitésimales mais rédhibitoires et l'esclave à main et à cerveau devait tout recommencer. Elle pratique à outrance la grève perlée et tour à tour l'imprimante, les dérouleurs, les lecteurs de carte, le perforateur débrayent sans préavis, bloquant la chaîne. Elle manque enfin terriblement de souplesse et ne se met jamais à la place des autres. Jamais le moindre effort pour comprendre une logique qui ne soit pas une alternative, une hésitation qui ne soit pas entre le tout et le rien.

De notre côté, nous n'étions pas sans reproche et sans faiblesse. Notre culture mathématique avait des racines étiolées et des branches mortes. Or l'informatique est une greffe qui réussit mieux sur un tronc mathématique vigoureux. La technologie d'autre part nous apparaissait proche de la magie et nous sentions nos doigts gourds et peu aptes à la prestidigitation. Notre logique surtout s'accommodait mal des angles droits et des lignes brisées. Nous avons pris l'habitude d'arrondir les angles et d'emprunter les raccourcis paresseux, parfois plus pittoresques, rarement plus courts, à la manière de La Fontaine se levant précipitamment de table deux heures avant la séance de l'Académie et s'excusant : « Ça ne fait rien, j'irai par le plus long »<sup>15</sup>. Nous avons cependant nos impatiences, nos exigences, et lorsque le résultat proposé par la machine ne correspondait pas à notre attente, nous accusions toujours la machine, comme ces couples dont la progéniture a fait une sottise et qui disent « ta fille », « ton fils ».

Au bout de quelques années cependant de vie commune, la machine s'est habituée à nous. Les scientifiques ébaubis ont parfois admiré sa complaisance, sa docilité à fournir des listes interminables, des tris sibyllins, et des réponses aux questions les plus bizarres qu'ils n'auraient pas cru qu'on pût poser sans mécontenter les circuits. Il faut avouer que les littéraires sont rares dans les centres de calcul et qu'une certaine curiosité amusée mais point trop condescendante les y entoure. Une certaine sollicitude aussi. L'accueil des techniciens aux étrangers n'est point barbare ni avare d'explications. Oserons-nous le dire enfin ?

---

<sup>15</sup> Cité par Giraudoux dans *Les Cinq Tentations de La Fontaine*, page 61.

L'agilité est venue peu à peu aussi au bout de nos doigts et de nos neurones et nous sommes devenus analyste, programmeur, opérateur, voire dépanneur. Notre tâche était facilitée par l'apparition récente sur le marché de l'informatique d'un langage puissant et universel, particulièrement adéquat aux recherches littéraires et linguistiques, le PL/1<sup>16</sup>. Ce langage qui allie la souplesse à la puissance permet donc tout à la fois le traitement des chaînes de caractères, les calculs mathématiques les plus sophistiqués et notamment le calcul booléen, la constitution et la consultation de fichiers divers, la manipulation des adresses, la gestion de la mémoire dite « pointée », la réalisation de piles, de chaînages et donc de grammaires formelles. À l'imitation des collègues des Facultés de Sciences, qui sont très généralement les programmeurs de leurs propres recherches, nous avons appris à manier directement cet outil. Au demeurant personne ne juge indigne d'apprendre à conduire une voiture et les chauffeurs sont devenus un luxe rare. L'informatique est en soi un domaine si coûteux que toute économie de personnel technique est bienvenue. À procéder soi-même, on gagne en outre du temps : car il est plus expéditif de demander directement à la machine ce que l'on veut plutôt que de l'expliquer à un programmeur professionnel. Qui peut le plus peut le moins : il est plus facile au chercheur de s'initier en quelques jours à la programmation qu'au programmeur de comprendre le but et de prévoir les résultats de telle recherche qui le dépasse. Or toute recherche est toujours un cas d'espèce : il est moins logique de demander au programmeur de s'adapter à un cas d'espèce qu'au chercheur de se conformer à une méthode générale. Outre qu'on n'a pas toujours un professionnel à son service, l'expérience montre qu'un attelage qui unit chercheur et programmeur progresse d'autant mieux qu'il met en commun plus de connaissances et que l'osmose des deux compétences est plus avancée. Et on en arrive vite à une équipe indifférenciée et comme dédoublée. Pourquoi dès lors gaspiller les efforts de deux personnes quand une seule suffit ? Ajoutons enfin qu'en programmant soi-même sa recherche, on gagne en sécurité et en fiabilité : comment s'assurer des résultats en effet si on ne sait pas contrôler leur élaboration ? Le résultat dépend de l'algorithme utilisé, qu'il est imprudent de confier à un tiers.

La programmation d'ailleurs est un exercice fort facile. Elle n'est une montagne que pour ceux qui n'ont jamais essayé. Il faudrait presque

---

<sup>16</sup> Ce sigle signifie : Programming Language Number One. Quoique ce langage en soit déjà à sa cinquième version, l'indice est resté bloqué à 1, sans doute pour marquer quelque préséance sur les autres langages.

avoir la naïveté de celui à qui on demandait s'il savait le violon et qui répondait : « Je ne sais pas, je n'ai jamais essayé ». L'apprentissage de l'informatique est si facile, qu'il suffit d'essayer. On peut même dire qu'à tous les niveaux de compétence, l'informaticien est quelqu'un qui essaie. C'est l'essai, ou la série d'essais, qui amène la réponse définitive. C'est donc plus simple que le permis de conduire. On ne peut pas tout de suite confier une voiture à un débutant, mais on peut lui confier l'ordinateur, lequel ne va jamais dans le fossé : il s'arrête avant en prévenant le conducteur qu'il y a danger et en expliquant pourquoi. C'est dans l'informatique qu'on a le meilleur exemple de maïeutique socratique. Avec l'ordinateur on accouche de la vérité, sans l'aide d'un tiers, sans le filtre sélectif, sans le miroir déformant qu'est toujours un enseignant. L'informatique est donc une autodiscipline, i.e. une discipline qui s'enseigne d'elle-même, qui se suffit à elle-même. Car elle ne suppose pas d'autre base, mathématique ou logique ou technologique. On peut ignorer la composition d'un transistor ou d'un tore magnétique, on peut ignorer les règles de la déduction correcte, on peut même ignorer tout ce qui est organisation physique, n'avoir jamais vu d'ordinateur et n'en avoir pas la moindre idée, on pourrait encore être programmeur et obtenir des résultats.

C'est que les langages évolués et particulièrement le PL/1 se sont rapprochés de l'homme en même temps qu'ils s'éloignaient de la machine. Ils se sont rapprochés du langage naturel, prenant même leur distance avec le langage algébrique. Ainsi dans le COBOL, le signe de l'addition est remplacé par la formule ADD TO. Les littéraires peuvent ainsi profiter des contraintes que les gestionnaires ont imposées aux constructeurs et de la simplification des procédures et des écritures qui en est résultée. Mais le linguiste qui voudrait approfondir l'étude de ces langages artificiels se reconnaîtra vite en paysage connu. Il se trouve que la conception de ces langages formels est empruntée pour une grande part aux théories de Chomsky et que les règles « context sensitive » y sont le plus souvent remplacées par des équivalences plus puissantes du type « context free »<sup>17</sup>. Ainsi des études théoriques purement linguistiques ont des retombées technologiques, dont, par un juste retour des choses, les linguistes peuvent demander à bénéficier les premiers. Ajoutons enfin que le langage ensembliste est actuellement une approche aisée de la programmation, dont la logique est booléenne, et particulièrement du

---

<sup>17</sup> On consultera à ce propos : *Notions sur les grammaires formelles* de Maurice Gross et André Lentin, Gauthier-Villars, 1970.

PL/1 qui permet le traitement des chaînes de bits et un calcul booléen direct, sans transposition.

Cela dit, le recours à l'ordinateur ne va pas sans de terribles contraintes. Si les premiers exercices procurent en général de l'amusement, de l'intérêt, voire de la jubilation, un contact prolongé avec l'ordinateur exige une rude santé. En six années de pratique ininterrompue nous avons amassé près d'une tonne de papier sous forme de cartes ou de listings, sans parler des bandes et des disques magnétiques. C'est peu si l'on songe aux kilomètres de rayonnages du T.L.F. C'est beaucoup pour un chercheur réduit à ses propres moyens, et traînant son matériel d'un centre de calcul à l'autre, de Nice à Marseille, de La Gaude à Paris. En six années, il aura fallu s'adapter à sept ordinateurs différents, à plusieurs langages, à plusieurs systèmes, à plusieurs constructeurs. Il aura fallu se contenter des « créneaux » dont personne ne veut, ceux de nuit, ceux de fête et particulièrement ceux de fête nocturne. Il aura fallu se débrouiller seul au milieu des clignotants, des boutons, des messages et des brochures, débloquer une imprimante qui s'obstine à noircir la même ligne, décongestionner un lecteur qui s'étouffe en mangeant trop de cartes... Il aura fallu quêter et compenser par la mendicité la modicité de nos ressources.

## ÉTABLISSEMENT DU CORPUS

Fort heureusement, une chiquenaude initiale nous avait lancé fort loin dans l'aventure avant que l'ampleur de la tâche ait été précisément mesurée. Nous avons reçu en effet de Nancy le dictionnaire des fréquences de Giraudoux, spécialement constitué pour notre recherche sur la base des seize textes de notre auteur qui avaient été retenus pour le T.L.F.<sup>18</sup> Il s'agissait d'un état imprimé qui ne pouvait pas servir directement à l'entrée des données dans la machine. Il eût été bien préférable de recevoir ce dictionnaire sous forme de bande magnétique mais la compatibilité entre matériel Bull et IBM était mal assurée et nous avons dû entreprendre la perforation de plus de 13 000 cartes. Nous en avons profité pour régler certaines difficultés qui n'avaient pas trouvé à Nancy leur solution complète :

1- L'ordre alphabétique a été restitué pour tous les mots qui comportent une voyelle accentuée et dont la place n'était pas régulière sur

---

<sup>18</sup> Nous renouvelons ici nos remerciements à M IMBS, directeur du T.L.F., et à M. GORCY, sous-directeur, pour cette contribution majeure à notre recherche.

l'état imprimé (ainsi les mots qui commencent par *é* étaient renvoyés à la fin du dictionnaire). Et cela était inévitable puisque la hiérarchie des combinaisons binaires qui symbolisent les lettres ne laissait qu'une place intermédiaire entre le 'D' et le 'F' pour le 'E' accentué ou non (dans le code à 8 bits pour 360, 11000100 = D, 11000101 = E, 11000110 = F). Nous avons préféré sauvegarder l'ordre alphabétique en sacrifiant les signes diacritiques, dont la frappe d'ailleurs était impossible vu la pauvreté du clavier de notre perforatrice. Mais à ce stade du traitement, la perte n'est pas très gênante : il suffit de distinguer les mots qui deviennent faussement homographes par certaines précisions comme PP (= participe passé) ou SUBST (= substantif) ou ADJ (= adjectif).

2- Les vrais homographes avaient été signalés par l'astérisque sur l'exemplaire de Nancy. Mais la distinction des divers emplois n'avait pas été assurée. Nous nous y sommes employé.

3- Les codes grammaticaux et sémantiques ont été ajoutés selon des principes qui seront exposés plus loin.

4- Enfin, le dictionnaire a été débarrassé de certaines impuretés (mots estropiés, noms propres mal repérés, noms composés mal coupés...).

Mais en même temps nous nous exposions à de nouvelles erreurs dont certaines risquaient d'être fort dommageables si elles concernaient les données numériques : imaginons qu'un chiffre soit décalé d'une colonne ; ce qui représentait une unité peut désigner maintenant une dizaine de milliers et le traitement de l'ensemble sera ruiné par cette méprise monstrueuse. Heureusement la parade était facile : il a suffi de porter sur la carte-fréquence le total des occurrences du mot (ce nombre était indiqué sur l'état imprimé) et de demander à la machine de faire l'addition des fréquences partielles, en comparant la somme obtenue avec celle qui lui était fournie. Lorsqu'il y avait divergence, c'est qu'une erreur s'était glissée dans la perforation des données.

Ce cadeau tombé du ciel nous créait des obligations mais aussi des contraintes. Il n'était pas possible de revenir sur les options prises à Nancy quant à la lemmatisation. Nous n'avons donc pas à justifier la norme qui a présidé à la délimitation du mot puisqu'elle nous fut imposée. Ajoutons cependant que nous l'avons trouvée bonne à l'expérience. Détacher du verbe le participe présent et le participe passé peut paraître arbitraire a priori mais le choix analytique a de gros avantages quand ce ne serait qu'il n'empêche pas la synthèse a posteriori, alors que l'inverse n'est pas vrai. Il n'était pas souhaitable non plus de

modifier – fût-ce en les améliorant – les options de Nancy. Car c'eût été s'interdire toute comparaison avec l'ensemble du T.L.F.<sup>19</sup>, ou avec les autres écrivains de ce corpus qui feront l'objet de monographies analogues à celle-ci. Dans un domaine coûteux comme celui-ci, une norme n'est bonne que si elle s'impose à tous, or trop d'arbitraire empêche qu'aucune s'impose par les seules vertus de son excellence, la discipline recommande donc d'éviter la tour de Babel et de choisir la doctrine dominante qui ne peut être que celle du T.L.F. et qui s'écarte fort peu de la tradition lexicographique et des principes que M. Muller formule dans sa thèse<sup>20</sup>.

## EXTENSION DU CORPUS

Si nous nous sommes interdits de modifier le corpus, nous n'avons pas renoncé à le compléter. Cela ne s'imposait pas pour le théâtre de Giraudoux où douze titres avaient été retenus dans le corpus de Nancy (*Siegfried*, *Amphitryon 38*, *Judith*, *Intermezzo*, *La Guerre de Troie*, *Électre*, *Cantique des Cantiques*, *Ondine*, *L'Apollon de Bellac*, *Sodome et Gomorrhe*, *La Folle de Chaillot*, et *Pour Lucrèce*). Par contre l'œuvre romanesque était fort mal représentée puisqu'elle ne comprenait que quatre textes : *Simon le Pathétique*, *Suzanne et le Pacifique*, *Siegfried et le Limousin* et *Bella*, et que leur extension chronologique était limitée. L'étude du roman giralducien supposait une assise élargie que nous avons constituée en ajoutant au corpus primitif sept textes nouveaux : *Provinciales*, *l'École des Indifférents*, *Juliette au Pays des Hommes*, *Églantine*, *Les Aventures de Jérôme Bardini*, *Combat avec l'Ange*, et *Choix des Élus*. A vrai dire, les deux premiers et les deux derniers textes de cette série ne sont représentés que sur échantillons. Pour *Provinciales* et *l'École des Indifférents*, cette réduction s'imposait du fait qu'il s'agit là de recueils de nouvelles. En ne retenant que l'une d'elles (respectivement *Sainte Estelle* et *Jacques l'Égoïste*), on conservait au récit une unité dont le recueil n'était pas pourvu au même degré. Par contre les deux derniers romans ont été réduits de manière aléatoire au cinquième de leur longueur (il a suffi d'extraire une page sur cinq). Ce faisant nous évitions d'avoir à comparer deux sous-ensembles – théâtre et roman – trop inégaux et surtout, en abaissant l'étendue de certains romans au dessous de la

<sup>19</sup> Le dictionnaire des fréquences du XXe siècle nous fut ultérieurement communiqué par le T.L.F. et les facilités de comparaison ainsi ménagées ont été mises à profit.

<sup>20</sup> *Étude de Statistique lexicale, La norme lexicologique*, p. 27.

moyenne du théâtre, on annulait l'influence de l'étendue relative – ce qui permettait l'étude des genres sans brouillage extérieur. Il s'agissait aussi d'expérimenter la méthode de l'échantillonnage, à côté de celle du dépouillement exhaustif. Enfin – faut-il l'avouer – des raisons d'économie nous ont inspiré, car les deux derniers romans sont longs et leur traitement intégral s'annonçait fort onéreux.

## LES PHASES DU TRAITEMENT

Chacun des sept textes ajoutés au corpus a fait l'objet d'une préparation complexe dont les différentes phases peuvent être ainsi résumées :

1- *Perforation du texte et vérification* (30 000 cartes perforées pour l'ensemble).

Tableau 1 : concordance

TABLEAU N° I . COORDONNES							
ABS	REL	POS	NER	PAGE	LIGN	JULIETTE AU PAYS DES HOMMES PAGE 353	
CONSEILS	2	2.6	10	12	30	8	JULIETTE* ALLAIT A L'ECOLE* NORMALE* SUPERIEURE* , SUR LES CONSEILS DU SECRETAIRE .
			15	20	53	16	JULIETTE , QUAND LE REGNE DES ANIMAUX LUI EUT ECHAPPE , AVAIT HESITE ENTRE LES DIVERS CONSEILS QUE LUI DONNAIT SON CARNET .
CONSENTEMENT	1	1.3	51	51	132	20	COMME CES DISPUTES QUE MENENT EN BAS NOTRE* DAME* ET LE SACRE* COEUR* , LE PANTHEON* ET LA GARE DE LYON* , ON VOIT D'ICI QU'ELLES SONT TRUQUEES POUR AMUSER UN PEU LES HOMMES ET QU'IL N'Y A , AU CONTRAIRE , ENTRE TOUS CES EDIFICES QU'ACCORD ET QUE CONSENTEMENT .
CONSIDERAIT	2	2.6	21	63	11	20	SI BIEN QUE TOUT CE QUE PUREMENT LES FIANCES FUT DE DIRIGER LA CONVERSATION SUR LES ECHEVISES , DONT GERARD* SE CONSIDERAIT LE FRERE PAR ORGUEIL , ELLE LA SOEUR PAR HUMILITE , ET ILS LEVERENT LES BALANCES , D'OU LA GENERATION DES ECHEVISES PRUDENTES , MAINTENANT RASSASIEE , ACHEVAIT DE DISPARAITRE SOUS LA PRESSION DES ETOURDIES , AFFAMEES DERECHER , ET VOUUES D'AILLEURS , PAR DEFINITION , A LA MORT .
			3	14	167	6	IL LA CONSIDERAIT TENDREMENT , DE SES YEUX SI DURS AUX CAILLES ET AUX GELINOTTES .
CONSIGNE	2	2.6	8	20	58	20	IL S'AGISSAIT DE GAGNER LE SUCRIER CONSIGNE DANS LE TESTAMENT DU DUC D'AUHALL* , ET DU SUCRE ETAIT PROMIS PAR LES CONCURRENTS A LEURS MONTURES .
			17	36	155	27	TOUT CELA JULIETTE* L'AVAIT BIEN CHERCHE , LE SOIR OU ELLE AVAIT VU REALISER LE VOEU CONSIGNE A LA PAGE 80 DU CARNET D - CONNAITRE UN RUSSE* , - L'EPOUSER - ETRE ABANDONNEE DE LUI , - LE REPRENDRE , - L'ABANDONNER
CONSISTAIT	1	1.3	2	15	111	10	ELLE CONSISTAIT A REGARDER CERTAINS ETRES , CERTAINS PAYS , AU VISAGE ET D'AUTRES AU CORPS .
CONSISTE	2	2.6	12	24	59	5	IL N'AVAIT JAMAIS COMMIS MEME CE MENSONGE DE FAIT QUI CONSISTE A APPELER GARDE FEU , LE PARE ETINCELLES OU LISERE LE PASSE POIL .
			9	65	175	12	TOUS LES DESASTRES DE LA PRECISITE MAL QUI CONSISTE A TRAITER LES OBJETS COMME DES HUMAINS , LES HUMAINS COMME S'ILS ETAIENT DIEUX ET VIERGES , LES DIEUX COMME DES CHATS OU DES BELETTES , MAL QUE PROVOQUE , NON PAS LA VIE DANS LES BIBLIOTHEQUE MAIS LES RELATIONS PERSONNELLES AVEC LES SAISONS , LES PETITS ANIMAUX , UN EXCESSIF PANTHEISME ET DE LA POLITESSE ENVERS LA CREATION , JULIETTE* LES ENTISSAIT SOUS SES PAS .

2- *Établissement d'un index et d'une concordance* (10 000 pages environ) grâce à des programmes originaux que nous avons publiés dans la revue C.U.M.F.I.D.<sup>21</sup>. On notera que la concordance restituée pour chaque mot le contexte de la phrase où il apparaît (voir Tableau 1) et que l'index précise non seulement les références de la forme employée (n° de la page, n° de la ligne, place du mot dans la phrase) mais aussi sa fréquence absolue dans le texte et dans chacun des chapitres, et sa fréquence relative (voir Tableau 2). Ces indications de fréquences sont perforées sur cartes, au cours du même traitement, afin de servir à l'entrée des données des programmes suivants.

Tableau 2

TABLEAU N° 2 . INDEX .		JULIETTE AU PAYS DES HOMMES										PAGE	114
PROVINCIAL	R.	12	25	16									
	ABS	(*)	1	REL	(f)	1.3		1	0	0	0	0	0
PROVINCIALE	R.	166	2	1									
	ABS	1	REL	1.3		0	0	0	0	0	0	1	0
PROVISOIREMENT	R.	103	7	2									
	ABS	1	REL	1.3		0	0	0	0	0	0	0	1
PROVOQUE	R.	74	9	38	R.109	20	37	R.175	12	36			
	ABS	3	REL	3.9		0	0	1	0	0	0	1	1
PROVOQUEES	R.	143	28	6									
	ABS	1	REL	1.3		0	0	0	0	0	1	0	0
PROVOQUER	R.	15	27	11	R. 74	9	17	R.180	22	24			
	ABS	3	REL	3.9		1	0	1	0	0	0	0	1
PRUDENTES	R.	9	25	25	R. 11	20	42						
	ABS	2	REL	2.6		2	0	0	0	0	0	0	0
PRUNELLES	R.	02	28	52	R. 00	7	47	R. 96	19	6	R. 96	27	2
	ABS	4	REL	5.2		0	0	0	4	0	0	0	0
PSEUDONYME	R.	131	25	40									
	ABS	1	REL	1.3		0	0	0	0	0	1	0	0
PSYCHE*	R.	105	17	2									
	ABS	1	REL	1.3		0	0	0	0	0	0	0	1
PSYCHOLOGIE	R.	122	11	19									
	ABS	1	REL	1.3		0	0	0	0	1	0	0	0
PSYCHOLOGUE	R.	103	19	7									
	ABS	1	REL	1.3		0	0	0	0	1	0	0	0
PSYCHOPHYSIOLOGIE	R.	122	11	22									
	ABS	1	REL	1.3		0	0	0	0	1	0	0	0
PU	R.	23	10	16	R. 29	14	0	R. 35	24	16	R. 40	6	17
	R.	39	10	57	R. 79	18	22	R. 02	28	15	R. 06	13	7
	R.	104	14	16	R.120	14	5	R.135	20	3	R.144	23	9
	R.	190	26	13	R.193	8	10	R.188	4	9			
	ABS	15	REL	17.6		0	4	1	3	2	2	2	1
PUBLICATION	R.	45	0	14	R.145	12	28						
	ABS	2	REL	2.6		0	1	0	0	0	1	0	0
PUBLIE	R.	45	0	38	R. 70	24	34	R. 91	21	18			
	ABS	3	REL	3.9		0	1	1	1	0	0	0	0
PUBLIERAIT	R.	101	26	0									
	ABS	1	REL	1.3		0	0	0	0	1	0	0	0
PUBLIQUE	R.	53	23	46									
	ABS	1	REL	1.3		0	0	1	0	0	0	0	0

a : n° page    b : n° ligne    c : place du mot dans la phrase  
d : fréquences par chapitre    e : fréquence absolue    f : relative

<sup>21</sup> *Cahiers des Utilisateurs de Machines à des Fins d'Information et de Documentation*, Juin 1970, p. 93 à 161, U.E.R. Lettres et Sciences Humaines, Nice.

3- *Lemmatisation*. Cette opération – semi-automatique – est nécessaire pour le regroupement des formes qui appartiennent au même vocable et qu'on range derrière l'infinitif, le masculin et le singulier, selon qu'il s'agit d'un verbe, d'un adjectif ou d'un substantif. Certes, cette phase du traitement peut être automatisée complètement si l'on dispose d'un dictionnaire de toutes les formes flexionnelles possibles. Mais les cas d'homographies seront alors difficiles à résoudre, ce qu'on peut faire aisément ici en recourant aux concordances. On trouvera dans C.U.M.F.I.D. n°3 page 1 à 8 notre programme de lemmatisation.

4- *Fusion*. Lorsque le fichier d'un certain texte ainsi lemmatisé a été constitué sur disque (ou sur bande) et que l'ordre alphabétique a été contrôlé, on le compare au dictionnaire des textes déjà traités et fusionnés, afin de faire apparaître les mots nouveaux. Ceux-ci reçoivent leurs codes grammatical et sémantique (opération manuelle) et le fichier du texte est modifié en conséquence. Alors seulement, la fusion peut être assurée et les fréquences du texte considéré sont portées dans le dictionnaire de l'auteur. Voir C.U.M.F.I.D. n°3, programme 13.

Quand toutes ces opérations sont menées à terme – il en coûte quelques jours à la machine et cent fois plus au chercheur – alors commence la phase noble de la recherche, celle de l'exploitation statistique.

## LA STATISTIQUE LINGUISTIQUE

La statistique linguistique a fait ses preuves et après notre plaidoyer pour l'informatique lexicale, il serait oiseux de rouvrir le dossier de la statistique linguistique et de reprendre une argumentation que Guiraud puis Muller ont imposée définitivement. Nous ne voulons ici que signaler le lien étroit qui unit la statistique à l'informatique. Les formules de la statistique sont souvent complexes et les opérations si nombreuses et embrouillées que les calculs manuels sont décourageants. L'utilisation des calculatrices de bureau – issues d'ailleurs de la technologie des ordinateurs – allège considérablement la tâche mais les erreurs peuvent se glisser dans la pose des chiffres ou dans la séquence des opérations. De plus, même douées d'une ou plusieurs mémoires, ces petites machines ne peuvent traiter les tableaux et leur emploi est donc réservé aux calculs élémentaires ou au contrôle partiel des résultats de l'ordinateur. Enfin la mise en page des résultats peut donner lieu à des erreurs de transcription alors que l'ordinateur peut se charger lui-même de l'impression claire et lisible des résultats acquis. C'est pourquoi la statistique linguistique – et

la statistique tout court – ne peut atteindre son plein développement qu'en association avec l'informatique. La statistique elle-même ne doit son essor actuel qu'aux prodigieuses ressources de l'ordinateur en matière de calcul. Il y a bien longtemps que les principes du calcul matriciel et de l'analyse factorielle ont été découverts, mais avant l'ordinateur, les applications, trop laborieuses, n'existaient guère. Aujourd'hui, les sciences économiques, la sociologie, la psychologie, la médecine et même parfois la physique et l'astrophysique ont fait de la statistique leur méthode principale d'investigation et de l'ordinateur leur outil essentiel. C'est au contact de ces disciplines que la statistique linguistique peut espérer combler son retard<sup>22</sup>. On peut se demander d'ailleurs si la notion de statistique linguistique a quelque réalité autonome. Certes la statistique est devenue très diversifiée et certains champs de la recherche particulièrement rebelles à l'analyse et à l'expérimentation ou éloignés de la loi des grands nombres nécessitent parfois un outillage spécialisé, plus sensible et plus délicat. Mais tel n'est pas le cas de la linguistique que traverse dans toute sa longueur la voie royale de la statistique classique. Pierre Guiraud encore une fois l'a bien vu : « La linguistique est la science statistique type ; les statisticiens le savent bien ; la plupart des linguistes l'ignorent encore »<sup>23</sup>. La médecine et la psychologie envient à la linguistique son aptitude à évoluer parmi les grands nombres : quand on traite une masse de trente sept millions d'individus (c'est le cas du corpus du XXe siècle dans le T.L.F.) on ne se soucie guère de l'erreur aléatoire. La sociologie, l'économie, l'astrophysique lui envient la sûreté et la précision de ses données de départ. Ceux qui se plaignent de l'incertitude dans la définition du mot n'ont pas été témoins de la désolation de l'astronome dont les mesures sont si coûteuses, si peu contrôlables et si souvent aberrantes ou du désespoir du sociologue qui dépouille son enquête et constate des erreurs et des lacunes impossibles à corriger. Toutes envient à la linguistique cette matière première si

---

<sup>22</sup> Ce retard est assez paradoxal vu que la linguistique a servi de modèle à la plupart de ces sciences et Pierre Guiraud le remarque fort justement : « On peut se demander pourquoi la linguistique si visiblement mieux douée que les autres sciences de l'homme est restée jusqu'ici en arrière ; il semble que ce soit en raison même de ses aptitudes qui lui ont permis de se définir en tant que science beaucoup plus tôt que les autres et à une époque où la statistique était encore dans l'enfance. Elle vit sur un outillage et un équipement démodés quand les tard venues bénéficient des progrès de la technique. » (*Problèmes et méthodes de la statistique linguistique*, p.23).

<sup>23</sup> Pierre Guiraud , ouvrage cité , p. 15. Cette phrase figure en exergue dans *Initiation à la statistique linguistique*, de Charles Muller.

abondante, si pure, si souple, si peu coûteuse, si disponible, une matière première que le chercheur contrôle absolument, sans intermédiaire et sans arbitraire. Or la statistique ne rend au mieux que ce qu'on lui donne. Il est très difficile de vaincre le principe de Carnot et de lutter contre la dégradation de l'énergie. Quand le désordre est inhérent aux données on s'époumonera en vain contre une entropie visqueuse, impossible à endiguer. C'est pourquoi plus encore que les données, ce sont les résultats de la statistique linguistique qui suscitent la jalousie des disciplines moins favorisées. Celles-ci se battent contre l'hypothèse nulle, qu'il leur est difficile de rejeter dans le brouillard de l'incertitude où le réel et le fantôme ont même apparence. Quand on joue sur les mots au contraire, le hasard est presque toujours éliminé du jeu. La psychologie ressemble souvent au poker, la médecine au bridge, la linguistique, elle, c'est le jeu d'échecs. Bien entendu le jeu est complexe et la stratégie infiniment savante. Mais rien n'apparaît aléatoire stricto sensu. On verra dans la suite de cette étude que le seuil de 5 % qui permet de rejeter le hasard est le plus souvent dépassé de loin. La plupart des graphiques ont également une telle évidence visuelle que le profane pourrait penser que la statistique n'est que la traduction graphique de l'évidence. Nos analyses factorielles dessinent des cercles si réguliers, en s'abstenant d'encombrer la zone centrale, que le non initié pourrait imaginer qu'il en est toujours ainsi. Dans la réalité des choses, en dehors de la linguistique, ces constellations de points s'appellent des nuages et méritent bien leur nom. Or les résultats que nous publions n'ont pas été choisis pour leur caractère spectaculaire, mais pour leur représentativité. Ceux qui sont restés dans nos cartons sont bien entendu à la disposition des chercheurs et des contrôleurs.

Bien que la statistique puisse être la meilleure école du mensonge, le contrôleur des poids et mesures (poids du sérieux, mesure de la sincérité) a cependant dans le cas présent la tâche facile. Nous n'avons pas craint en effet d'alourdir la présentation en multipliant les tableaux et en n'esquivant pas les résultats intermédiaires. Il s'agissait d'abord de faciliter le contrôle, le soutien et la correction. Dans un domaine où l'évidence fait très vite défaut et où les prises manquent, l'ascension est le plus souvent artificielle, il faut multiplier pitons et jalons, afin que la trace ne soit pas plus difficile à suivre qu'à faire. Nous avons redouté l'ivresse dangereuse des excursions en solitaire et nous nous sommes soucié de laisser derrière nous un chemin visible et praticable afin qu'on

puisse nous accompagner et au besoin nous assurer<sup>24</sup>. Quoique, au delà de cette introduction, les mots *programme*, *informatique*, *ordinateur*, aient été proscrits afin de parler le langage de l'honnête homme, nous avons soumis nos programmes eux-mêmes au contrôle des spécialistes en les publiant dans notre revue C.U.M.F.I.D. La multiplication des tableaux et graphiques correspond aussi à la volonté de fournir des éléments de comparaison aux études postérieures. Corneille par le truchement de Muller a aidé Giraudoux, et Giraudoux dont on connaît le caractère aimable peut aider quelqu'un d'autre. Peut-être même des études plus fines ou mieux inspirées peuvent-elles être menées sur Giraudoux lui-même à partir des résultats que nous publions et dont l'interprétation ne prétend ni à l'exhaustivité ni à l'exactitude.

Car il y a deux façons d'étudier le langage comme il est deux façons d'examiner la lune : ou bien l'on se sert de lunettes astronomiques et l'on bâtit des théories, grossissant les détails, éclairant les ombres, réduisant les contradictions. Ou bien on envoie quelqu'un sur la lune à qui on demande de ramasser ce qu'il voit et de bien voir ce qu'il ne ramasse pas, sans trop se préoccuper de théorie et de choix. Ou si l'astronaute choisit, c'est selon le critère de la variété, et non pas parce que telle roche est plus compatible avec telle ou telle théorie préalable. Certes entre notre entreprise et le programme Apollo il y a toute la distance qui sépare la terre de la lune. Mais parce qu'elle était coûteuse, nous avons senti que notre expédition imposait des devoirs : et notamment celui d'amasser les matériaux sans les calibrer en vertu d'hypothèses qui serviraient de crible, celui d'amasser des nombres plutôt que des mots, des faits plutôt que des conclusions, des masses plutôt que des curiosités. Sur notre montagne hérissée de chiffres – les chiffres sont des épines, non des barbelés et c'est celui qui fraye le passage qui s'y pique le plus – nous avons pris des photos. Certaines paraîtront peut-être floues, parce que l'opérateur a bougé ou parce que la brume habitait les lointains. Elles devraient toutefois, bénéficiant de la loi des grands nombres, aider à discerner les grandes lignes du paysage linguistique et littéraire, la direction des chaînes chronologiques, l'exposition des versants stylistiques, et l'aspect thématique des roches et de la végétation.

---

<sup>24</sup> Au reste il ne s'agit pas précisément ici d'une première. Le temps des pionniers est passé.