

UNE APPROCHE STATISTIQUE DES PROBLEMES QUE SOULEVE EN FRANCAIS CONTEMPORAIN LA PLACE DES ADJECTIFS QUALIFICATIFS EPITHETES (1)

ABSTRACT

This article attempts to make a study of some of the problems involved in the coexistence, in present-day French, of the order adjective-noun and of the order noun-adjective. It therefore examines the quantitative characteristics of a corpus covering 2,361 "homogeneous" and assimilated "heterogeneous" noun phrases.

A first series of tests (§ 3) leads to the conclusion that the sequence of the qualifiers (*) inside a "homogeneous" noun phrase most frequently follows the descending order of their frequencies and - this is not unrelated - the ascending order of their lengths. In addition, one finds that everything happens, when the aim is to produce "a set in the mathematical sense", as if the qualifiers most rapidly called up were those stored in the memory so as to be the most accessible ones, and as if the words thus stored were the most frequently used.

A second series of tests (§ 4) gives an idea of the perturbation of the order of nouns and adjectives by their relationships with certain grammatical categories and by their intra- and extrasyntagmatic contexts.

The article concludes with some historical considerations and a few general comments.

(*) I.e. all the components of the homogeneous NP except adverbs.

1. AVANT-PROPOS

1.1. Sans cesse reprises, jamais entièrement résolues, les questions liées au fait que l'ordre adjectif qualificatif-substantif et l'ordre substantif-adjectif qualificatif coexistent dans notre langue tiennent quelque peu du rocher de Sisyphé.

Néanmoins, on devra les étudier plus d'une fois encore, car elles ont été ordinairement abordées d'une manière trop étroite, trop rigide et, surtout, trop subjective.

1.1.1. En effet, la discussion originelle n'a presque jamais été intégrée dans un débat plus large. C'est, par exemple, au cours des dernières années seulement qu' E. SPANG-HANSEN et quelques autres linguistes (2) ont insisté sur les diverses préséances entre adjectifs qualificatifs d'un même syntagme. Et un ouvrage aussi récent que la thèse d'E. REINER prêtait encore le flanc au très grave reproche de s'être, selon l'expression de W. ZWANENBURG, "conscieusement abstenu de placer le problème dans le cadre de l'ensemble des déterminants du substantif" (3). Enfin, la perspective comparatiste reste sous-estimée à l'heure actuelle bien que, comme le rappelait M. WANDRUSKA, "die schon so oft und so gründlich untersuchten Probleme der adjektivischen

Determination im Französischen versteht man erst richtig, wenn man von der südlichen Romania ausgeht" (4).

1.1.2. Cela dit, parce qu'elle a été initialement posée en termes normatifs, la question de la place des adjectifs qualificatifs épithètes a surtout suscité la recherche de principes à valeur absolue. Pourtant, si l'on en croit H. J. CEDERGREN et D. SANKOFF, "the notion of competence must be strengthened to include representation of systematic covariation between elements of language, even when this covariation cannot be described in categorical (zero-one) terms. The representation should be in the form of abstract probabilities, which measure tendencies for rule application in various environments" (5) et on devrait apprendre à parler des lois syntaxiques comme on a appris à parler des lois phonétiques, c'est-à-dire en se rappelant sans cesse qu'elles ne s'appliquent pas nécessairement à tous les cas et qu'elles peuvent - tantôt à cause de certains contextes, tantôt à cause du système entier de la langue - être altérées profondément par des tendances concurrentes.

1.1.3. De plus, à cause sans doute de l'importance stylistique qu'a l'alternative antéposition de l'adjectif-postposition de l'adjectif, beaucoup d'études n'ont pas jugé indispensable de tabler exclusivement sur des données tangibles. Les travaux quantifiés, notamment, sont rares et - depuis ceux de J. DAMOURETTE et E. PICHON (6) jusqu'à ceux de J.-Cl. CHEVALIER (7) - ils n'offrent en majorité que très peu d'intérêt sur le plan statistique. Quant aux dépouillements plus sérieux, ils ne concernent guère le français contemporain et partent de prémisses assez vagues (ainsi l'entreprise de K. WYDLER qui repose sur une triade" [Qualifikativa] schwach differenzierter Bedeutung ... Determinativa ... Reine Qualifikativa [und Adjektiva,] die sowohl als Det. als auch als Qual., in konkreter wie in übertragener Bedeutung oder sonst in mehreren Bedeutungen,

verwendet werden" (8)).

1.2. Il y a donc là un champ de travail presque vierge où le présent article désirerait risquer une brève incursion. Nous nous en tiendrons d'ailleurs aux limites d'un corpus de taille fort réduite, mais nous essayerons d'éviter le plus possible les défauts critiqués par les §§ 1.1.1, 1.1.2 et 1.1.3. En tout cas, nous nous ferons une règle générale de n'enquêter que sur des caractères immédiatement quantifiables et que sur des formes aisément repérables.

2. MATERIAUX ANALYSES

2.1. Nature du corpus

Cette restriction nous impose de constituer notre corpus de travail à partir d'énoncés émanant d'un individu unique, car les aspects du langage qui se prêtent le mieux au calcul subissent - de personne à personne - des variations parfois importantes (9). De plus, il conviendrait de choisir des textes qui se présenteraient sous une forme écrite sans être empreints pour autant d'un cachet littéraire appuyé : de la sorte, les opérations de codage se trouveraient grandement simplifiées, mais les inconvénients engendrés par les expressions archaïques et les procédés livresques de toutes espèces ne prendraient pas une ampleur inacceptable.

En conséquence, nous pensons expédient d'utiliser la matière fournie par un ensemble de récits dus au seul Jean RAY : ils sont largement diffusés et présentent le double avantage d'appartenir à la "paralittérature" et d'enregistrer une proportion élevée (à peu près 45 %) de phrases "de conversation". Au demeurant, le choix de ces récits nous épargnera de longs et fastidieux

préparatifs, car ils ont déjà été mis sur cartes perforées pour répondre aux exigences d'une autre étude (10).

Cette dernière fournit une description détaillée de l'ensemble des textes en cause. Rappelons toutefois qu'il réunit 4 collections respectivement formées de 681 phrases des *Contes du whisky*, 675 phrases des fascicules 69 et 141 de la série *Harry Dickson*, 559 phrases du *Livre des fantômes* et 787 phrases du *Carrousel des maléfices*. Dès lors, si quelque trait s'affirme avec une netteté identique dans chacune de ces collections, il relèvera, selon toute vraisemblance, de la langue française en général et non des habitudes personnelles de l'auteur qui est mis à contribution. D'une part, en effet, la première, la troisième et la quatrième collections proviennent de recueils de contes fantastiques assez brefs tandis que la deuxième collection provient de récits policiers relativement longs, d'autre part, la première collection correspond au premier livre de Jean RAY tandis que la dernière correspond au dernier (et il y a plus de 40 ans entre ces deux ouvrages). Or des récits que ne rapprochent ni les dimensions, ni le genre sont censés avoir des coordonnées stylistiques très différentes (11) et, par ailleurs, on sait combien la technique d'un écrivain évolue du début à la fin de sa carrière (12).

2.2. Principes de codage

2.2.1. Fondements linguistiques

2.2.1.1. Nous postulons avec G. VAN HOUT qu' "Un syntagme nominal est une structure linguistique ordonnée signifiant un ensemble (au sens mathématique)" (13) et nous attribuerons à cette structure linguistique les frontières que lui ont reconnues B. POTTIER en l'étudiant comme chaîne d'unités inférieures ("Le syntagme nominal peut être défini par sa constitution séquentielle

maximale dans le discours" (14)) et M. TUȚESCU en l'étudiant comme unité de chaînes supérieures ("Nous comprenons par contextes distributionnels du GN l'ensemble des environnements discursifs - se trouvant au niveau de la phrase - dans lesquels le GN peut apparaître" (15)).

2.2.1.2. Nous opposerons, d'une part, les syntagmes nominaux homogènes (*Snho*) où "les éléments qui entourent le substantif (ou le mot fonctionnant comme tel) [le font] sans aucun intermédiaire [...] 'presque tous les petits chats bien blancs' " (16), d'autre part, les syntagmes nominaux hétérogènes (*Snhé*) où "les éléments [sont] introduits par les démarcateurs du type *et*, *de*, ou *que* [...] 'les conditions et les réalisations de l'enquête que vous menez' " (16) et les deux dénominations - notons-le au passage - seront indifféremment appliquées aux syntagmes primaires et aux syntagmes secondaires (17).

2.2.1.3. Nous considérerons que tout *Snho* est formé d'un certain nombre de petits syntagmes (*s*) contenant chacun un mot (le qualificateur *q*) qui désigne "une propriété spécifique pour [un] élément de l'ensemble défini par [le] syntagme nominal" (18) et, parfois, un ou plusieurs mots (les quantitatifs *qn*) qui se rapportent - en général par voie d'antéposition - au qualificateur ou à un autre quantitatif (19).

2.2.1.4. Nous admettrons pour finir qu'un qualificateur est un nom "relativement à un certain univers du discours ssi il y forme son extension" (20) (ainsi *bête*, *Saïda* et *ah* dans *une vieille bête*, *ma petite Saïda* et *ce ah! horrible*) et qu'un qualificateur est un "adjectif relativement à un certain univers du discours U_d lorsqu'il n'est pas [nom] relativement à U_d et lorsqu'il est [nom] relativement à un sous-ensemble non trivial de U_d " (21).

2.2.2. Méthode de classement

Ceci convenu, nous allons grouper les qualificateurs de notre corpus selon un système de classes et de sous-classes susceptibles de mettre en évidence quelques-uns des facteurs qui influent sur l'agencement de tels ou tels adjectifs par rapport à tels ou tels noms. Nous procéderons ensuite de la même manière pour les divers *Snho* de notre corpus et *Snhé* assimilés (cf. le § 2.2.2.2).

Les protocoles des groupements en question comportent une série de stipulations particulières qui sont destinées à garantir l'application des autres clauses - beaucoup plus générales - contre la variabilité et les risques de flottement.

Ces prescriptions mineures laissent à désirer dans la mesure où elles ne relèvent pas d'une logique commune et où, le plus souvent, elles pèchent par inexactitude linguistique. Toutefois, malgré leur nombre, elles n'interviendront jamais que pour trancher des cas limites dont la fréquence, sans exception ou presque, reste très basse. Aussi, dans le cadre d'une étude à caractère statistique, leurs imperfections ne tireront-elles pas à conséquence.

C'est seulement au prix de pareils défauts, d'ailleurs, que nos conventions de codage parviendront à satisfaire deux exigences primordiales : se prêter à une automation poussée et pouvoir être relayées, en ce qui concerne le détail, par de vastes répertoires préexistants (22).

Soit dit incidemment, tout mot en instance de codage se verra assigner, sauf avis contraire, le lemme et la catégorie grammaticale qui, compte tenu de son contexte propre, lui sont attribués par le *Petit Robert* ou, si ni ce dernier, ni le *Robert* en 7 volumes ne se révèlent suffisamment explicites, par la 8e

édition du *Grevisse*.

Les "parties du discours" auxquelles se réfèrent ces volumes ne résistent certes pas à un examen scrupuleux (23), mais elles ont l'avantage d'être bien connues de chacun, d'être inventoriées très complètement et d'être définies avec une précision indiscutable. Sous l'égide d'un P. GUIRAUD, d'un P. IMBS ou d'un Ch. MULLER, elles ont au reste été constamment utilisées par la linguistique quantitative française. Leur exploitation ménage donc, en l'occurrence, la possibilité de rattacher sans trop de peine les résultats de notre enquête à des dénombrements d'une étendue beaucoup plus importante.

Précisons encore que, ci-dessous, les exemples français, néerlandais et anglais qui seront imprimés en italique auront été respectivement tirés de notre corpus, de la *Nouvelle grammaire-index de la langue néerlandaise* (24) et de la *Grammaire descriptive de l'anglais contemporain* (25). Quant aux exemples qui seront imprimés en caractères romains et qui seront mis entre guillemets, ils auront été soit forgés pour les besoins de la cause, soit empruntés à une source décrite par une note particulière. Dans tous ces énoncés types, les mots qui servent à la stricte illustration de l'unité linguistique concernée et les mots qui ne servent pas à celle-ci seront distingués d'une manière ou d'une autre : les derniers apparaîtront le plus souvent entre parenthèses s'ils proviennent du texte original et entre crochets s'ils lui ont été ajoutés.

Enfin, les symboles de deux classes ou sous-classes d'unités linguistiques se verront quelquefois associés par \cap , par + ou bien par | qui, en informatique, exprime la concaténation. Une configuration de forme $A\cap B$ désignera les unités de A qui appartiennent également à B .

Une configuration de forme $A + B$ désignera les unités de A qui ont une unité de B comme complément. Une configuration de forme $A | B$ indiquera que les unités de B suivent d'une façon immédiate les unités de A .

Accessoirement, | pourra marquer aussi un découpage de la phrase que l'usage orthographique n'admet pas (par exemple : *le | quel*).

2.2.2.1. Classement des qualificateurs.

Nous classerons les noms en noms propres (*NP*) et noms communs (*NC*).

Dans la pratique, il s'agira respectivement des noms dont la lettre initiale est une majuscule et des noms dont la lettre initiale est une minuscule : le critère s'avère des plus commode pour l'analyse par ordinateur et, si le *Robert* ne le retient pas (26), du moins, le § 236 du *Grevisse* l'adopte-t-il.

Nous diviserons également les adjectifs en deux catégories : d'une part, les non-qualificatifs, c'est-à-dire l'ensemble des adjectifs exclamatifs, indéfinis, interrogatifs, numériques, possessifs et relatifs ainsi que l'ensemble des "articles" contractés, définis, indéfinis et partitifs, d'autre part, les qualificatifs au sens large (*QSL*).

Ces derniers seront affectés de l'indice "*1er antéposé*" ou "*1er postposé*", "*2e antéposé*" ou "*2e postposé*", etc. lorsque, par rapport à un nom, ils occuperont la même place que *cher* ou *humide* dans *cher homme* et *son étoupe humide*, que *pauvre* ou *malodorant* dans *le pauvre cher homme* et *son étoupe humide et malodorante*, etc.

Au demeurant, ils seront rangés en deux grandes sous-catégories : les qualificatifs

au sens strict (*QSS*) et les qualificatifs d'origine verbale, lesquels, à leur tour, seront subdivisés en qualificatifs à finale "ent" ou "ant" (*QNT*) et qualificatifs-participes passés (*QPP*).

La frontière entre les adjectifs d'origine verbale et les formes verbales qui se trouvent à leur origine réclame, tant elle est floue, un jalonnement très abondant.

Pour être enregistré comme *QNT*, un mot susceptible d'être aussi regardé comme participe présent aura à satisfaire aux critères fournis par l'orthographe (27) ou, quand celle-ci ne sera pas discriminante, aux critères énoncés par le § 772 du *Grevisse*.

De plus, nous coderons comme *QNT* quelques *OSL* à finale "ent" qui ne proviennent pas d'un verbe (*lent*, par exemple) : cette démarche, en effet, permet des raccourcis de programmation très appréciables et n'entraîne, malgré son caractère aberrant, que d'infimes séquelles.

Pour être enregistré comme *QPP*, un mot susceptible d'être aussi regardé comme participe passé aura à répondre à deux séries d'exigences : primo, ne constituer ni, avec un auxiliaire, un syntagme copule, ni, avec un nom ou un pronom, une proposition circonstancielle indépendante, secundo, n'être ni précisé par une locution adverbiale de temps ou un adverbe de temps (tel que *fraîchement* dans *mon arme fraîchement aiguisée*), ni prolongé par un attribut (tel que *femme* dans *l'exactitude faite femme*), par un complément d'agent, par un complément d'attribution, par un complément circonstanciel de lieu qu'introduirait une des locutions prépositives ou prépositions à *l'abri de*, à *l'intérieur*

de, en haut de, au long de, contre, dans, de, derrière, par, parmi, sous, sur ou vers, par un complément circonstanciel de manière autre qu'une comparaison (tel que à une vitesse inaccoutumée dans convoi lancé à une vitesse inaccoutumée) ou encore par une proposition infinitive de finalité (telle que de ne pas se conduire selon la norme géométrique dans courbe absolument décidée de ne pas se conduire selon la norme géométrique) (28).

De plus, nous coderons toujours comme *QPP* les mots qui devront être amputés ou augmentés d'un préfixe négatif avant d'être identifiés avec le participe passé d'un verbe repris au *Petit Robert* (*inconnue*, par exemple, dans *une dimension encore inconnue*).

Sera, par contre, répertorié comme *QSS* le participe passé de tout verbe dont les autres modes et temps sont obsolètes (c'est-à-dire de tout verbe dont l'infinitif est attesté dans le *Godefroy* ou le *Huguet*, mais non dans le *Petit Robert*). Cette disposition sera évidemment caduque pour les quelques passages où un état de langue ancien se trouve recréé (pastiches, etc.).

2.2.2.2. Classement des *Snho* et des *Snhé* assimilés.

"La caractéristique essentielle du démarcateur de coordination [étant] de laisser la construction de l'énoncé au niveau acquis au moment où il apparaît"(29), le petit nombre des *Snhé* qui ont pour seul démarcateur une conjonction de coordination insérée entre *s* seront, ci-dessous, provisoirement assimilés à des *Snho*.

2.2.2.2.1. Nous opposerons tout d'abord aux autres *Snho* - qui seront appelés "non-bases" ($\sim B$) - les *Snho* que complète une unité linguistique introduite par un démarcateur de détermination ou bien par un pronom relatif (16) et nous dirons d'eux qu'ils constituent des "bases" (*B*).

Néanmoins, lorsqu'une unité linguistique introduite par un démarcateur de détermination pourra être regardée aussi bien comme le complément d'une proposition que comme le complément d'un *Snho*, nous tiendrons toujours ce dernier pour une $\sim B$ (ainsi *un immense effort* dans la phrase *Blomme fit un immense effort pour parler*).

Quant aux expressions du genre *drôles de petites bêtes*, dont le mot initial présente un caractère assez ambigu, elles seront traitées en *Snho* où *drôles* est un *2e antéposé* et non en *Snhé* où *drôles* serait un substantif et formerait donc une *B* (cf. sur ce point - *dura lex, sed lex* - le '*Robert*, s.v. "drôle", II et le § 303, R.2 du *Grevisse*).

Nous classerons les *B* en *Ba* et en *Bn* suivant qu'un adjectif ou qu'un nom y servira de support à l'unité linguistique qui vient les compléter.

Toutefois, lorsque cette unité linguistique sera introduite par un démarcateur exprimant une comparaison, on éprouvera souvent quelque difficulté à décider si c'est à une *Ba* ou à une *Bn* que l'on a affaire. En l'occurrence, il nous a semblé expédient de considérer que la *B* sera du premier type quand elle comprendra un adjectif susceptible de s'unir au syntagme comparant par l'entremise de l'auxiliaire "être" (par ex. *des nabots hauts (comme trois pommes)[sont hautes]*) et que la *B* sera du second type quand elle comprendra un nom susceptible de s'unir au syntagme comparant par l'entremise de l'auxiliaire

“avoir” (par ex. *les pieds palmés (comme les canards) [ont les pieds]*).

Par ailleurs, une classe mixte ($Ba \cap Bn$) a été prévue pour regrouper les *Snho* qui satisferont en même temps à la définition des *Ba* et à la définition des *Bn* (p.ex. *(il y a) un tel silence (qui tombe sur la ville) (que la lune vagabondant sur les toits semble insolente et tapageuse)*).

Au sein des $Bn + \sim P$ (*Bn* et $Ba \cap Bn$ qu’une unité linguistique autre qu’une proposition ou qu’un syntagme verbal vient compléter par le truchement d’un démarcateur comme *à, autour, contre, de, en, envers, par, pour, etc.*), nous distinguerons deux sous-catégories : d’une part, les bases $Bn + c \sim P$ et $Ba \cap Bn + c \sim P$ dont dépend une séquence subordonnée courte, c’est-à-dire ne présentant pas de *B*, ne comprenant pas de *QSL* et incluant un seul nom ou une seule paire titre de civilité-nom de personne, d’autre part, les bases $Bn + l \sim P$ et $Ba \cap Bn + l \sim P$ dont dépend une séquence subordonnée longue, c’est-à-dire ne remplissant pas les 3 conditions qui viennent d’être formulées. Ainsi, seront $Bn + c \sim P$: *ta bonne main (de jadis), la douloureuse semence (de mes larmes)* et *la hautaine demeure (de la douairière de la Closerie)*. En revanche, seront $Bn + l \sim P$: *les rouges convoitises (de ceux de Whitechapel), ce merveilleux Midi (de votre belle patrie)* et *les aigres suées (de leurs maris, petits employés de la City ou boutiquiers de Battersea)*. Quant aux *Bn* et $Ba \cap Bn$ qu’une proposition ou qu’un simple syntagme verbal vient compléter par le truchement d’un démarcateur comme *à, de, dont, lequel, où, que, qui, quoi, etc.*, ils seront répertoriés sous les étiquettes $Bn + P$ et $Ba \cap Bn + P$.

Au reste, pour éviter la création de classes mixtes dont les effectifs seraient extrêmement réduits, les *Snho* qui cumulent les caractères des $Bn + c \sim P$ et

des $Bn + I \sim P$ (comme *le clair glacé* dans *le clair (de lune) glacé (des grosses perles électriques)*) seront toujours comptés au nombre des $Bn + I \sim P$ ou bien des $Ba \cap Bn + I \sim P$ et les *Snho* qui cumulent les caractères des $Bn + \sim P$ et des $Bn + P$ (comme *quelques rares instruments* dans *quelques rares instruments de médecine qui paraissaient avoir été achetés d'occasion*) seront toujours comptés au nombre des $Bn + P$ ou bien des $Ba \cap Bn + P$.

2.2.2.2.2. Cela noté, le tri des *Snho* ne s'effectuera pas seulement en fonction de séquences spécifiques qu'ils contribuent à former, mais encore en fonction d'éléments spécifiques qui interviennent dans leur propre formation. Cette partition complémentaire s'exprimera par un double jeu d'indices : des chiffres d'abord, notifiant la nature du mot initial du *Snho*; des lettres ensuite, symbolisant l'ordonnance, dans le *Snho*, des mots qui appartiennent aux catégories grammaticales du nom, du *QSL* ainsi que du quantificatif du *QSL*.

Le chiffre *I* signalera tout *Snho* qui commence par un adjectif exclamatif (p. ex. *quel affreux bruit*) ou bien qui, débutant par un autre qualificateur, constitue le premier élément d'une proposition exclamative marquée du signe de ponctuation adéquat (p. ex. *trois horribles choses !*) et formée d'un seul syntagme nominal (ce n'est donc pas le cas dans un procès comme *la haute finance après la noblesse !*). Accessoirement, le même chiffre romain signalera aussi tout *Snho* primaire et toute *B* qui se trouvera en apostrophe (p. ex. *cher monsieur* et *vieux colporteur (de faux axiomes et de théorèmes caducs)*).

Le chiffre *II* signalera tout *Snho* qui commence par un adjectif interrogatif (p. ex. *quelle étrange mouche (vous pique ?)*) ou bien qui, débutant par un autre qualificateur, constitue le premier élément d'une proposition interrogative

marquée du signe de ponctuation adéquat et formée d'un seul syntagme nominal (p. ex. *sale guetteur (d'étoiles ?), de la fausse monnaie ?, etc.*).

Le chiffre *III*, de son côté, signalera tout *Snho* qui n'entre pas dans l'une des deux catégories ci-dessus. Il sera suivi d'un chiffre arabe qui précisera si tel ou tel *Snho* commence par un adjectif possessif (indice 1), par un adjectif démonstratif, indéfini, numéral ou relatif (indice 2 : p. ex. *ce brave homme, toutes ses diaboliques métamorphoses, seuls quelques pauvres diables, on ne sait quelle raison précise* (30), etc.), par un article contracté ou défini (indice 3) ou bien par un adjectif qualificatif, un article indéfini ou partitif (31) ou même un nom (indice 4 : p. ex. *(à) grosses gouttes, une vieille bête, d'autres petits états, (de) personnes nombreuses*).

Par ailleurs, nous représenterons les *Snho* au moyen des lettres *an* (p. ex. *la vilaine aventure*), *na* (p. ex. *l'air marin*) ou bien *ana* (p. ex. *les beaux yeux gris*) selon qu'ils offriront un *QSL* avant leur nom et aucun après, un *QSL* après leur nom et aucun avant ou bien un *QSL* avant et un *QSL* après leur nom (à condition que les *QSL* en cause ne soient ni précédés d'un quantitatif, ni concaténés ou coordonnés à un autre *QSL*, ni affectés d'un complément).

Nous représenterons les *Snho* au moyen des lettres *dan* (p. ex. *le très grand monde*), *nda* (p. ex. *l'humanité tout entière*) ou bien *anda* (p. ex. *(en) vieil ivoire finement sculpté*) selon qu'ils offriront un *QSL* précédé d'un quantitatif avant leur nom et pas de *QSL* après, un *QSL* précédé d'un quantitatif après leur nom et pas de *QSL* avant ou bien un *QSL* précédé d'un quantitatif après leur nom et un *QSL* avant (à condition que les *QSL* en cause ne soient ni

concaténés ou coordonnés à un autre *QSL*, ni affectés d'un complément).

Nous représenterons les *Snho* au moyen des lettres *aan* (p. ex. *la mystérieuse petite lumière*), *naa* (p. ex. *des chairs molles et décomposées*), *aana* (p. ex. *un bon et gros homme jovial*) ou bien *anaa* (p. ex. *un jeune homme très pâle et très beau*) selon qu'ils offriront plusieurs *QSL* avant leur nom et pas de *QSL* après, plusieurs *QSL* après leur nom et pas de *QSL* avant, plusieurs *QSL* avant leur nom et un *QSL* après ou bien plusieurs *QSL* après leur nom et un *QSL* avant (à condition que les *QSL* en cause ne soient pas affectés d'un complément).

Enfin, nous représenterons les *Snho* au moyen des lettres *acn* (p. ex. *une telle aversion (que je la tuai)*), *nac* (p. ex. *l'ombre vide (de toute présence)*) ou bien *anac* (p. ex. *un petit théâtre voisin (de Drury Lane)*) selon qu'ils offriront au moins un *QSL* affecté d'un complément avant leur nom et pas de *QSL* après, au moins un *QSL* affecté d'un complément après leur nom et pas de *QSL* avant ou bien au moins un *QSL* affecté d'un complément après leur nom et au moins un *QSL* avant (32).

Pour en avoir terminé avec l'énoncé de toutes ces conventions, il nous faut encore souligner trois faiblesses sérieuses de notre codage : primo, en raison de l'assimilation aux *Snho* de certains *Snhé*, l'ordinateur classera sous une même étiquette des syntagmes différant aussi fortement par la structure que - prenons l'exemple de la désignation *aan* - *le pauvre cher homme, le terrible et mystérieux animal* et *l'atroce, l'indéfinissable contact*, secundo, il réunira en un *Snho* le *QSL* et - exclusivement - le nom le plus proche de celui-ci dans des syntagmes du genre *des larmes et du sang répandus* (notre corpus n'en

comprend que 6) et, tertio, il ne tiendra pas compte du nom propre, mais seulement du titre de civilité dans l'enregistrement de syntagmes tels que (*la place chaude*) du bon Mr. Pilgrim.

3. OBSERVATIONS (33)

3.1. Quelques résultats et quelques suppositions.

Notre corpus comprend 2361 *Snho* et *Snhé* assimilés qui, offrant un ou plusieurs *QSL*, nous intéressent ici au premier chef.

Les qualificateurs autres que les adjectifs non qualificatifs s'y rangent, d'après la place occupée à l'intérieur des syntagmes, en 6 catégories : *2es antéposés*, *1ers antéposés*, *noms*, *1ers postposés*, *2es postposés* et *3es postposés* (34). A ces catégories, correspondent, dans l'ordre, les 6 colonnes primaires du Tableau I.

Chacune de ces colonnes se compose, soit de 3 colonnes secondaires qui fournissent des chiffres relatifs aux *NP*, aux *NC* et à l'ensemble *NO* de ces 2 sous-catégories, soit de 4 colonnes secondaires qui fournissent des chiffres relatifs aux *QSS*, aux *QNT*, aux *QPP* et à l'ensemble *QSL* de ces 3 sous-catégories.

Les diverses colonnes sont segmentées en 5 bandes primaires. Celles-ci correspondent, dans l'ordre, aux collections de textes procédant des *Contes du whisky* (CW), des fascicules 69 et 141 de la série *Harry Dickson* (HD), du *Livre des fantômes* (LF), du *Carrousel des maléfices* (CM) et de l'ensemble RAY de ces parties du corpus.

Chacune des bandes primaires se compose de 5 bandes secondaires identifiées

par les abréviations NMB, OCC, LET, O/N et L/N.

Dans le cadre de la collection qu'elle concerne, une bande NMB permet de connaître le nombre de qualificateurs enregistrés pour chaque (sous-) catégorie. Symbolisons par l le nombre de lemmes différents enregistrés dans le cadre d'une collection donnée et pour une (sous-) catégorie donnée et symbolisons par o_i le nombre d'occurrences qui relèvent du i ème des lemmes dans le cadre de cette collection et pour cette (sous-) catégorie. Un total apparaissant en bande NMB équivaudra donc à $\sum_{i=1}^l o_i$ (cf. n. 35).

Une bande OCC donne un indice de fréquence. Dans le cadre de la collection qu'elle concerne, elle signale pour chaque (sous-) catégorie la somme des produits obtenus en multipliant deux facteurs : d'une part, le nombre d'occurrences qui relèvent du i ème des lemmes différents enregistrés dans le cadre de cette collection et pour une (sous-) catégorie donnée, d'autre part, le nombre d'occurrences qui relèvent du même lemme dans la collection RAY et avec n'importe quelle place à l'intérieur de n'importe quel syntagme (36). Symbolisons par O_i ce second nombre. Un total apparaissant en bande OCC équivaudra donc à $\sum_{i=1}^l (o_i \cdot O_i)$.

Une bande LET donne un indice de longueur. Dans le cadre de la collection qu'elle concerne, elle signale pour chaque (sous-) catégorie la somme des produits obtenus en multipliant deux facteurs : d'une part, le nombre d'occurrences qui relèvent du i ème des lemmes différents enregistrés dans le cadre de cette collection et pour une (sous-) catégorie donnée, d'autre part, le nombre de lettres et traits d'union - 16 au maximum (37) - qu'offre le même lemme au *Petit Robert*. Symbolisons par G_i ce nombre d'unités graphiques. Un total apparaissant

en bande LET équivaldra donc à $\sum_{i=1}^I (o_i \cdot G_i)$.

Dans le cadre de la collection qu'elle concerne, une bande O/N signale pour chaque (sous-) catégorie le quotient de l'indice OCC par le nombre NMB.

Et, dans le cadre de la collection qu'elle concerne, une bande L/N signale pour chaque (sous-) catégorie le quotient de l'indice LET par le nombre NMB.

3.1.1. Caractéristiques globales du recensement.

3.1.1.1. Examinons d'abord les chiffres qui, aux quatrièmes bandes secondaires du Tableau I, intéressent les *NO* et les *QSL*.

En ce qui concerne le corpus entier, la valeur O/N des *2es antéposés* (23, 372) est, remarquons-nous, MOINS élevée que celle des *1ers antéposés* (34, 862), laquelle est plus élevée que celle des *noms* (13, 468), laquelle est plus élevée que celle des *1ers postposés* (6, 358), laquelle est plus élevée que celle des *2es postposés* (5, 889), laquelle, enfin, est plus élevée que celle des *3es postposés* (2, 272).

Nous observons, par ailleurs, que la tendance a une forte stabilité : en ce qui concerne les diverses parties du corpus (CW; HD; LF et CM), les diverses valeurs O/N des *2es antéposés* (16, 75; 22, 625; 22,090 et 27,937) sont MOINS élevées que leurs homologues des *1ers antéposés* (37, 591; 34, 283; 29, 155 et 38,062), lesquelles sont plus élevées que leurs homologues des *noms* (13,275; 14,579; 11,988 et 14,359), lesquelles sont plus élevées que leurs homologues des *1ers postposés* (7,166; 5,368; 5,894 et 6,705), lesquelles sont respectivement plus élevée, MOINS élevée, plus élevée et plus élevée que leurs homologues des

2es postposés (7,035; 7,25; 4,111 et 5,333), lesquelles, enfin, sont plus élevées que leurs homologues des *3es postposés* (2,25; 2,5; 2 et 2,5).

Considérons que ces 5 séries de 4 comparaisons forment 20 épreuves indépendantes et que, pour chacune d'entre elles, un échec (rendu aisément repérable ci-dessus par un mot en capitale) ou bien un succès est enregistré selon que, dans telle ou telle partie du corpus, la valeur O/C d'une catégorie se révèle inférieure ou bien supérieure à celle de la catégorie qui la suit immédiatement à l'intérieur des *Snho* et *Snhé* assimilés.

Si l'issue de chaque épreuve avait la même probabilité d'être défavorable que d'être favorable, les chances d'obtenir, en 20 essais, 5 échecs ou moins et 15 succès ou plus seraient égales à $(1/2)^{20} (C_{20}^{15} + C_{20}^{16} + C_{20}^{17} + C_{20}^{18} + C_{20}^{19} + C_{20}^{20})$

(cf. n.38). Autrement dit, la situation constatée ici et les situations comptant davantage de réussites n'auraient, à elles toutes, que 1,9553 chance sur 100 de se présenter. Ce que nous avons appelé un échec - l'infériorité, dans telle ou telle partie du corpus, de la valeur O/C d'une catégorie par rapport à celle de la catégorie suivante - est donc beaucoup moins normal que ce que nous avons appelé un succès : la supériorité, dans telle ou telle partie du corpus, de la valeur O/C d'une catégorie par rapport à celle de la catégorie suivante.

3.1.1.2: Examinons maintenant les chiffres qui, aux cinquièmes bandes secondaires du Tableau I, intéressent les *NO* et *QSL*.

En ce qui concerne le corpus entier, la valeur L/N des *2es antéposés* (5,953) est moins élevée que celle des *1ers antéposés* (5,995), laquelle est moins élevée

que celle des *noms* (6, 290), laquelle est moins élevée que celle des *1ers postposés* (6, 824), laquelle est moins élevée que celle des *2es postposés* (7, 128), laquelle, enfin, est moins élevée que celle des *3es postposés* (7, 181).

Cette fois encore, la tendance a une forte stabilité : en ce qui concerne les diverses parties du corpus (CW; HD; LF et CM), les diverses valeurs L/N des *2es antéposés* (5,75; 5,75; 5,636 et 6,375) sont respectivement moins élevée, moins élevée, moins élevée et PLUS élevée que leurs homologues des *1ers antéposés* (5,890; 5,967; 6,131 et 5,996), lesquelles sont moins élevées que leurs homologues des *noms* (6,069; 6,413; 6,386 et 6,356), lesquelles sont moins élevées que leurs homologues des *1ers postposés* (6,545; 7,244; 6,922 et 6,702), lesquelles sont respectivement PLUS élevée, moins élevée, moins élevée et moins élevée que leurs homologues des *2es postposés* (6,321; 7,625; 7,222 et 7,4), lesquelles, enfin, sont respectivement moins élevée, moins élevée, moins élevée et PLUS élevée que leurs homologues des *3es postposés* (6,5 ; 9,5; 8,333 et 4,5).

Ainsi que nous l'avons fait tout à l'heure, nous considérerons que ces 20 comparaisons forment 20 épreuves. Mais, à présent, un échec (rendu aisément repérable ci-dessus par un mot en capitale) sera enregistré chaque fois que, dans telle ou telle partie du corpus, la valeur L/N d'une catégorie se révélera supérieure à celle de la catégorie qui suit immédiatement (un succès sera enregistré dans le cas inverse).

Dès lors, si, pour toute épreuve, le résultat défavorable offrait autant de probabilité que le résultat favorable, un score aussi bon ou meilleur que le score qui a été obtenu (3 échecs et 17 succès) n'aurait que 0,1215 chance sur 100 d'être atteint (38). Par conséquent, la supériorité, dans telle ou telle partie du corpus, de la valeur L/N d'une catégorie par rapport à celle de la catégorie suivante est

beaucoup moins normale que son infériorité.

3.1.2. En guise d'hypothèse

Comme les valeurs O/N et L/N donnent, l'une, le nombre moyen des occurrences enregistrées dans le corpus entier par les membres de chaque catégorie de *NO* et *QSL* qui appartiennent à telle ou telle partie du corpus ou au corpus entier, l'autre, le nombre moyen des lettres et traits d'union enregistrés au *Petit Robert* par les membres de chaque catégorie de *NO* et *QSL* qui appartiennent à telle ou telle partie du corpus ou au corpus entier, on peut donc, si l'on néglige les 5 irrégularités notées par le § 3.1.1.1 et les 3 irrégularités notées par le § 3.1.1.2 et si l'on extrapole du corpus à l'ensemble des phrases françaises, formuler deux propositions concurrentes, mais non contradictoires : primo, **que les noms et qualificateurs au sens large se succèdent, à l'intérieur des syntagmes nominaux homogènes et syntagmes nominaux hétérogènes assimilés, selon l'ordre décroissant des fréquences**, secundo, **que les noms et qualificateurs au sens large se succèdent, à l'intérieur des syntagmes nominaux homogènes et syntagmes nominaux hétérogènes assimilés, selon l'ordre croissant des longueurs**.

3.2. Quelques corroborations et quelques postulats

Ces deux conjectures -que nous appellerons dorénavant l'hypothèse fréquence et l'hypothèse longueur- réclament un fondement plus solide qu'une simple confrontation de moyennes : les valeurs O/N et les valeurs L/N ne donnent aucun renseignement sur la dispersion des séries statistiques visées et, partant, ne les caractérisent pas d'une manière très satisfaisante (46).

C'est pourquoi nous laisserons momentanément de côté le Tableau I et nous nous reporterons à un inventaire d'une autre espèce

Les 2361 *Snho* et *Snhé* assimilés de RAY offrent 2668 paires de qualificateurs dont le premier élément appartient à l'une de nos 6 catégories de *NO* et *OSL* et dont le second élément appartient à la catégorie immédiatement consécutive. Ces paires se répartissent en 5 groupes : *2es antéposés-1ers antéposés*, *1ers antéposés-noms*, *noms-1ers postposés*, *1ers postposés-2es postposés* et *2es postposés-3es postposés* (39). Les 5 colonnes du Tableau II correspondent, dans l'ordre, aux groupes en cause.

Ces colonnes sont segmentées, tout comme les colonnes du Tableau I, en 5 bandes primaires qui correspondent, dans l'ordre, à CW, HD, LF, CM et RAY.

Chacune des bandes primaires se compose de 3 bandes secondaires identifiées par les abréviations HF, NBR et HL.

Une bande HF fournit des chiffres relatifs à l'hypothèse fréquence.

Une bande NBR, dans le cadre de la collection qu'elle concerne, permet de connaître le nombre de paires de qualificateurs enregistrées pour chaque groupe.

Une bande HL fournit des chiffres relatifs à l'hypothèse longueur.

Les bandes secondaires HF et HL se composent, l'une et l'autre, de 3 bandes tertiaires identifiées par les abréviations C, N et I.

Dans le cadre de la collection qu'elle concerne, une bande C signale pour chaque groupe le pourcentage de paires de qualificateurs dont les particularités confirment la conjecture testée, c'est-à-dire dont le premier qualificateur a un nombre O plus élevé que le second (s'il s'agit de l'hypothèse fréquence) ou dont le premier qualificateur a un nombre G moins élevé que le second (s'il s'agit de l'hypothèse longueur).

Dans le cadre de la collection qu'elle concerne, une bande N signale pour chaque groupe le pourcentage de paires de qualificateurs dont les particularités ne confirment, ni n'infirmement la conjecture testée, c'est-à-dire dont le premier qualificateur a le même nombre O que le second (s'il s'agit de l'hypothèse fréquence) ou dont le premier qualificateur a le même nombre G que le second (s'il s'agit de l'hypothèse longueur).

Et, dans le cadre de la collection qu'elle concerne, une bande I signale pour chaque groupe le pourcentage de paires de qualificateurs dont les particularités infirment la conjecture testée, c'est-à-dire dont le premier qualificateur a un nombre O moins élevé que le second (s'il s'agit de l'hypothèse fréquence) ou dont le premier qualificateur a un nombre G plus élevé que le second (s'il s'agit de l'hypothèse longueur).

3.2.1. Caractéristiques particulières du recensement

3.2.1.1. Sur la base du Tableau II, nous chercherons d'abord à mesurer l'ampleur de la différence entre le nombre de fois où l'hypothèse fréquence n'a été ni confirmée, ni infirmée et le nombre de fois où l'hypothèse longueur n'a été ni confirmée, ni infirmée.

Nous recourrons, ce faisant, à la méthode bien connue qui a été mise au point par Pearson. Il n'est sans doute pas inutile de détailler les modalités de son emploi dans la présente évaluation puisque celle-ci inaugure toute une série d'estimations où un rôle similaire sera joué par la même technique.

Considérons que les 2668 cas où l'hypothèse fréquence a été testée et les 2668 cas où l'hypothèse longueur a été testée constituent les deux sous-ensembles d'une vaste population de 5336 membres (40).

Les cas où une hypothèse n'a été ni confirmée, ni infirmée s'y trouvent au nombre de 629 (cf. n. 41).

Ils devraient se trouver, si la population en cause était parfaitement homogène malgré sa double origine, à peu près au nombre de 314 dans le premier sous-ensemble et à peu près au nombre de 314 dans le second sous-ensemble.

Or, ces deux nombres calculés - $629/(2668/5336) = 314,5$ - s'écartent sensiblement des nombres recensés dans la réalité : 223, d'une part, 406, d'autre part.

Nous avons constitué deux petits répertoires qui ont chacun n colonnes correspondant aux n sous-ensembles de la population et k bandes correspondant aux k sous-ensembles de la population. Les nk cases du Tableau III fournissent les nombres c qui ont été calculés pour les nk sous-ensembles de sous-ensembles. Les nk cases du Tableau IV fournissent les nombres r recensés dans la réalité pour les nk sous-ensembles de sous-ensembles.

Etant égale à $\sum_{i=1}^n \sum_{j=1}^k (r_{ij} - c_{ij})^2 / c_{ij}$, la quantité dite χ^2 s'obtiendra ici

par la suite d'opérations $[(223 - 314,5)^2 + (406 - 314,5)^2]/314,5$ et vaudra 53,241.

Or Pearson a montré que la loi de distribution de χ^2 est définie par $y = [e^{-\chi^2/2} (\chi^2)^{(\nu/2)-1}] / [2^{\nu/2} (\nu/2 - 1)!]$ où e est le nombre transcendant - à peu près 2,7182818 - qui résulte de la somme $\sum_{n=0}^{\infty} 1/n!$ et où ν est le nombre entier - appelé "nombre de degrés de liberté" - qui résulte du produit $(n - 1)(k - 1)$ sauf avec $n = 1$ (on a alors $\nu = k - 1$) ou avec $k = 1$ (on a alors $\nu = n - 1$) (cf. n. 42).

En vertu de ce fait, une table (43) a pu être élaborée qui permet, dès que ν est déterminé, de trouver immédiatement quelle est la probabilité P d'un χ^2 égal ou bien supérieur à celui qu'on a obtenu.

Ici, $\nu = 1$ (avec $n = 2$ et $k = 1$) et cette probabilité P - qui traduit la probabilité d'un écart égal ou bien supérieur à celui qui sépare les nombres c des nombres r - n'est même pas de 1 sur 1000.

Dès lors, nous devons conclure que la différence offerte par nos deux hypothèses au point de vue du nombre des paires de qualificateurs qui ne les confirment, ni ne les infirment est significative :

By convention, we regard the 0.05 level of probability corresponding to a calculated risk of 5 in 100 of being wrong, as a serviceable border-line between significance and non-significance [...] What cannot be regarded as a random sample by the token of the .05 (2σ) level of probability, may still be one if we chose the .003 (3σ) level as the border line between 'significant' and 'not significant'. In language statistics, the 3σ level would seem to be more appropriate (44).

Cette différence est à peine ^{marginale} significative dans CW; HD; LF et CM où elle

se solde par des χ^2 respectivement égaux à 13,296; 15,791; 3,751 et 24,674 (P < 0,001; 0,001; 0,1 et 0,001).

Nous aurions tort, d'ailleurs, de nous en étonner, car l'hypothèse fréquence est testée par la comparaison de deux valeurs du caractère "nombre d'occurrences dans le corpus", lequel peut prendre beaucoup de valeurs, tandis que l'hypothèse longueur est testée par la comparaison de deux valeurs du caractère "nombre de lettres et traits d'union au *Petit Robert*", lequel peut seulement prendre 16 valeurs (37).

Or, sauf quand la distribution est tout à fait anormale, les chances que deux mots consécutifs offrent la même valeur d'un caractère donné sont moins fortes avec un nombre de valeurs élevé qu'avec un nombre de valeurs relativement modique.

Soit dit au passage, si l'hypothèse qui suppose aux *1ers antéposés* une fréquence plus élevée qu'aux *noms* et qui suppose aux *noms* une fréquence plus élevée qu'aux *1ers postposés* était juste, elle devrait n'être ni confirmée, ni infirmée pour le groupe *1ers antéposés-noms* beaucoup moins souvent qu'elle n'est ni confirmée, ni infirmée pour le groupe *noms-1ers postposés* et, de fait, nous enregistrons là une différence qui s'avère significative ($\chi^2 = 39,580$ (P < 0,001)).

Pareillement, si l'hypothèse qui suppose aux *1ers antéposés* une longueur plus faible qu'aux *noms* et qui suppose aux *noms* une longueur plus faible qu'aux *1ers postposés* était juste, elle devrait n'être ni confirmée, ni infirmée pour le groupe *1ers antéposés-noms* beaucoup plus souvent qu'elle n'est ni confirmée, ni infirmée pour le groupe *noms-1ers postposés* et, de fait, nous enregistrons là aussi une différence (mais - on s'y attendait quelque peu - elle ne s'avère pas

significative ($\chi^2 = 0,492$ ($P < 0,6$)).

Au reste, la conclusion à tirer de la disparité qui, avant cette petite digression, a été notée entre nos deux conjectures semble tout à fait évidente : lorsque nous voudrions mettre directement en balance les résultats obtenus avec l'hypothèse fréquence et les résultats obtenus avec l'hypothèse longueur, nous aurons l'obligation de ne jamais y inclure le nombre des fois où ni une confirmation, ni une infirmation n'est acquise.

3.2.1.2. Ce principe étant posé, le Tableau II va nous servir - tantôt pour la totalité des groupes, tantôt pour chacun des groupes et tantôt dans la totalité RAY des collections, tantôt dans chacune des collections CW; HD; LF et CM - à comparer le nombre des cas où notre première conjecture (notre seconde conjecture) est confirmée au nombre des cas où notre première conjecture (notre seconde conjecture) est infirmée.

Quand la méthode de Pearson sera appliquée aux comparaisons en cause, ν sera égal à 1 : effectivement, il s'agira partout de $n = 1$ sorte d'hypothèses et de $k = 2$ sortes de résultats.

3.2.1.2.1. Pour la totalité des groupes, RAY enregistre significativement plus ($\chi^2 = 363,701$ ($P < 0,001$)) de paires confirmant l'hypothèse fréquence (1694) que de paires infirmant celle-ci (751) et significativement plus ($\chi^2 = 113,190$ ($P < 0,001$)) de paires confirmant l'hypothèse longueur (1384) que de paires infirmant celle-ci (878).

3.2.1.2.2. Pour la totalité des groupes, CW; HD; LF et CM enregistrent significativement plus (χ^2 respectifs = 94,744; 83,250; 70,222 et 120,000 (P < 0,001)) de paires confirmant l'hypothèse fréquence (nombres respectifs : 490; 349; 419 et 436) que de paires infirmant celle-ci (nombres respectifs : 229; 146; 209 et 167) et significativement plus (χ^2 respectifs = 33,184; 25,498; 18,222 et 399,006 (P < 0,001)) de paires confirmant l'hypothèse longueur (nombres respectifs : 409; 278; 355 et 342) que de paires infirmant celle-ci (nombres respectifs : 260; 171; 250 et 197).

3.2.1.2.3. Pour chacun des groupes *1ers antéposés-noms*, *noms-1ers postposés* et *1ers postposés-2es postposés*, RAY enregistre significativement plus (χ^2 = 265,268; 114,464 et 4,846 (P < 0,001; 0,001 et 0,05)) de paires confirmant l'hypothèse fréquence (807; 798 et 56) que de paires infirmant celle-ci (272; 424 et 35) et significativement plus (χ^2 = 32,338; 74,062 et 5,148 (P < 0,001; 0,001 et 0,05)) de paires confirmant l'hypothèse longueur (561; 735 et 58) que de paires infirmant celle-ci (386; 440 et 36).

Pour chacun des autres groupes, il est illicite d'utiliser la méthode de Pearson en ce qui concerne la totalité RAY des collections (45). Notons cependant que, pour chacun des groupes *2es antéposés-1ers antéposés* et *2es postposés-3es postposés*, RAY enregistre plus de paires confirmant l'hypothèse fréquence (24 et 9) que de paires infirmant celle-ci (19 et 1) (cf. n. 46) et plus de paires confirmant l'hypothèse longueur (23 et 7) que de paires infirmant celle-ci (15 et 1).

3.2.1.2.4. Enfin, pour le groupe *1ers antéposés-noms*, CW; HD; LF et CM enregistrent significativement plus (χ^2 respectifs = 76,756; 51,251; 54,524 et 84,508 ($P < 0,001$)) de paires confirmant l'hypothèse fréquence (nombres respectifs : 234; 155; 203 et 215) que de paires infirmant celle-ci (nombres respectifs : 79; 52; 79 et 62) et, à une exception près, significativement plus (χ^2 respectifs = 12,300; 7,534; 2,394 et 13,094 ($P < 0,001$; 0,01; 0,2 et 0,001)) de paires confirmant l'hypothèse longueur (nombres respectifs : 171; 104; 143 et 143) que de paires infirmant celle-ci (nombres respectifs : 112; 68; 118 et 88).

Pour le groupe *noms-1ers postposés*, CW; HD; LF et CM enregistrent significativement plus (χ^2 respectifs = 25,498; 39,386; 18,202 et 35,792 ($P < 0,001$)) de paires confirmant l'hypothèse fréquence (nombres respectifs : 233; 180; 192 et 193) que de paires infirmant celle-ci (nombres respectifs : 136; 79; 117 et 92) et, à une exception près, significativement plus (χ^2 respectifs = 12,300; 7,534; 2,394 et 13,094 ($P < 0,001$; 0,01; 0,2 et 0,001)) de paires confirmant l'hypothèse longueur (nombres respectifs : 221; 156; 186 et 172) que de paires infirmant celle-ci (nombres respectifs : 134; 92; 119 et 95).

Pour chacun des autres groupes, il est illicite d'utiliser la méthode de Pearson en ce qui concerne chacune des collections CW; HD; LF et CM.

3.2.1.3. Afin d'apprécier par contraste la solidité de l'hypothèse fréquence et la solidité de l'hypothèse longueur, nous allons - toujours à partir du Tableau II - opposer les scores que la première conjecture et les scores que la seconde conjecture ont obtenus en trois domaines distincts : celui des χ^2 exprimant la prédominance des succès sur les échecs, celui du volume des succès et des échecs et, en dernier lieu, celui du lien entre la proportion de succès et d'échecs et les divers groupes.

3.2.1.3.1. Les χ^2 traduisant la prédominance des succès sur les échecs se révèlent beaucoup plus élevés avec l'hypothèse fréquence qu'avec l'hypothèse longueur.

La chose - non seulement dans la totalité RAY des collections, mais encore dans chacune des collections CW; HD; LF et CM - se vérifie à la fois pour la totalité des groupes (cf. les §§ 3.2.1.2.1 et 3.2.1.2.2) et pour chacun des deux groupes les plus importants : *1ers antéposés-noms* et *noms-1ers postposés* (cf. les §§ 3.2.1.2.3 et 3.2.1.2.4).

Elle se vérifie donc pour chacun des endroits où la méthode de Pearson a pu être utilisée, à une exception près : le groupe *1ers postposés-2es postposés* où la totalité RAY des collections enregistre un χ^2 moindre avec l'hypothèse fréquence qu'avec l'hypothèse longueur (cf. le § 3.2.1.2.3 et, subsidiairement, la fin du § 3.2.1.1).

3.2.1.3.2. Nous allons encore comparer - tantôt pour la totalité des groupes, tantôt pour chacun des groupes et tantôt dans la totalité RAY des collections, tantôt dans chacune des collections CW; HD; LF et CM - le nombre des cas où notre première conjecture est confirmée (est infirmée) au nombre des cas où notre seconde conjecture est confirmée (est infirmée).

Quand la méthode de Pearson sera appliquée aux comparaisons en cause, ν sera égal à 1 : effectivement, il s'agira partout de $n = 2$ sortes d'hypothèses et de $k = 1$ sorte de résultats.

On verra ci-dessous que le sens du rapport des succès de l'hypothèse fréquence aux succès de l'hypothèse longueur et le sens du rapport des échecs de l'hypothèse fréquence aux échecs de l'hypothèse longueur constituent deux phénomènes constants et symétriques (47) et qu'ils étaient annoncés d'une manière assez précise (cf. n. 48 et n. 49) par le phénomène brièvement évoqué dans le § 3.2.1.3.1.

3.2.1.3.2.1. En effet, pour la totalité des groupes, RAY enregistre significativement plus ($\chi^2 = 31,221$ ($P < 0,001$)) de paires confirmant l'hypothèse fréquence (1694) que de paires confirmant l'hypothèse longueur (1384) et significativement moins ($\chi^2 = 9,901$ ($P < 0,01$)) de paires infirmant la première (751) que de paires infirmant la seconde (878).

3.2.1.3.2.2. Pour la totalité des groupes, CW; HD; LF et CM enregistrent significativement plus (χ^2 respectifs = 7,298; 8,039; 5,291 et 11,357 ($P < 0,01$; 0,01; 0,05 et 0,001)) de paires confirmant l'hypothèse fréquence (nombres respectifs : 490; 349; 419 et 436) que de paires confirmant l'hypothèse longueur (nombres respectifs : 409; 278; 355 et 342) et moins - mais ce n'est pas significatif (χ^2 respectifs = 1,965; 1,971; 3,662 et 2,472 ($P < 0,2$; 0,2; 0,1 et 0,2)) - de paires infirmant la première (nombres respectifs : 229; 146; 209 et 167) que de paires infirmant la seconde (nombres respectifs : 260; 171; 250 et 197).

3.2.1.3.2.3. Pour le groupe *1ers antéposés-noms*, RAY enregistre significativement plus ($\chi^2 = 44,236$ ($P < 0,001$)) de paires confirmant l'hypothèse fréquence (807)

que de paires confirmant l'hypothèse longueur (561) et significativement moins ($\chi^2 = 19,750$ ($P < 0,001$)) de paires infirmant l'hypothèse fréquence (272) que de paires infirmant l'hypothèse longueur (386).

Pour chacun des groupes *noms-1ers postposés* et *1ers postposés-2es postposés*, RAY enregistre plus et MOINS (48) - mais ce n'est pas significatif ($\chi^2 = 2,589$ et $0,035$ ($P < 0,2$ et $0,9$) - de paires confirmant l'hypothèse fréquence (798 et 56) que de paires confirmant l'hypothèse longueur (735 et 58) et moins - mais ce n'est pas significatif ($\chi^2 = 0,296$ et $0,014$ ($P < 0,7$ et $0,95$)) - de paires infirmant l'hypothèse fréquence (424 et 35) que de paires infirmant l'hypothèse longueur (440 et 36).

Pour chacun des autres groupes, il est illicite d'utiliser la méthode de Pearson en ce qui concerne la totalité RAY des collections.

3.2.1.3.2.4. Enfin, pour le groupe *1ers antéposés-noms*, CW; HD; LF et CM enregistrent significativement plus (χ^2 respectifs = $9,8$; $10,042$; $10,404$ et $14,480$ ($P < 0,01$; $0,01$ et $0,001$)) de paires confirmant l'hypothèse fréquence (nombres respectifs : 234; 155; 203 et 215) que de paires confirmant l'hypothèse longueur (nombres respectifs : 171; 104; 143 et 143) et, à une exception près, significativement moins (χ^2 respectifs = $5,701$; $2,133$; $7,720$ et $4,506$ ($P < 0,02$; $0,2$; $0,01$ et $0,05$)) de paires infirmant l'hypothèse fréquence (nombres respectifs : 79; 52; 79 et 62) que de paires infirmant l'hypothèse longueur (nombres respectifs : 112; 68; 118 et 88)).

Pour le groupe *noms-1ers postposés*, CW; HD; LF et CM enregistrent plus - mais ce n'est pas significatif (χ^2 respectifs = $0,317$; $1,714$; $0,095$ et $1,208$ ($P < 0,7$. $0,2$;

0,8 et 0,3)) - de paires confirmant l'hypothèse fréquence (nombres respectifs : 233; 180; 192 et 193) que de paires confirmant l'hypothèse longueur (nombres respectifs : 221; 156; 186 et 172) et PLUS (cf. n. 49), moins, moins et moins - mais ce n'est pas significatif ($\chi^2 = 0,014; 0,988; 0,016$ et $0,048$ ($P < 0,95; 0,5; 0,9$ et $0,9$)) - de paires infirmant l'hypothèse fréquence (nombres respectifs : 136; 79; 117 et 92) que de paires infirmant l'hypothèse longueur (nombres respectifs : 134; 92; 119 et 95).

Pour chacun des autres groupes, il est illicite d'utiliser la méthode de Pearson en ce qui concerne chacune des collections CW; HD; LF et CM.

3.2.1.3.3. Nous allons à présent comparer entre eux - tantôt dans la totalité RAY des collections amputée des paires du plus petit groupe (45), tantôt dans chacune des collections CW; HD; LF et CM amputées des paires des deux plus petits groupes (45) - les nk nombres des cas où notre première conjecture (notre seconde conjecture) a obtenu l'une des k appréciations pour l'un des n groupes.

Quand la méthode de Pearson sera appliquée aux comparaisons en cause, ν sera tantôt égal à 3, tantôt égal à 2 : effectivement, il s'agira tantôt de $n = 4$, tantôt de $n = 3$ sortes de paires et, partout, de $k = 2$ sortes de résultats.

3.2.1.3.3.1. Pour chacun des groupes *2es antéposés-1ers antéposés*, *1ers antéposés-noms*, *noms-1ers postposés* et *1ers postposés-2es postposés*, RAY enregistre, d'une part, moins, plus, moins et moins ($\sum_{i=1}^{n=4} (r_{i1} - c_{i1})^2 / c_{i1} = 1,113 + 4,876 + 2,680 + 0,771$, soit 9,440) de paires confirmant l'hypothèse fréquence que ne l'impliquerait l'absence de lien entre la proportion de succès et d'échecs et les divers groupes,

d'autre part, plus, moins, plus et plus ($\sum_{i=1}^{n=4} (r_{i2} - c_{i2})^2 / c_{i2} = 2,501 + 10,952 + 6,026 + 1,734$, soit 21,213) de paires infirmant la même hypothèse que ne l'impliquerait la même absence de lien. Au reste, ces différences sont significatives : $\chi^2 = 9,440 + 21,213$, soit 30,653 ($P < 0,001$).

Pour chacun des groupes *2es antéposés-1ers antéposés*, *1ers antéposés-noms*, *noms-1ers postposés* et *1ers postposés-2es postposés*, RAY enregistre, d'une part, moins, moins, plus et plus ($\sum_{i=1}^{n=4} (r_{i1} - c_{i1})^2 / c_{i1} = 0,001 + 0,531 + 0,411 + 0,005$, soit 0,948) de paires infirmant l'hypothèse longueur que ne l'impliquerait l'absence de lien entre la proportion de succès et d'échecs et les divers groupes, d'autre part, plus, plus, moins et moins ($\sum_{i=1}^{n=4} (r_{i2} - c_{i2})^2 / c_{i2} = 0,003 + 0,835 + 0,644 + 0,008$, soit 1,490) de paires infirmant la même hypothèse que ne l'impliquerait la même absence de lien. Toutefois, ces différences ne sont pas significatives : $\chi^2 = 0,948 + 1,490$, soit 2,438 ($P < 0,5$).

3.2.1.3.3.2. Pour chacun des groupes *1ers antéposés-noms*, *noms-1ers postposés* et *1ers postposés-2es postposés*, CW; HD; LF et CM enregistrent, d'une part, plus, moins et moins (Σ respectives = 1,816 + 1,522 + 0,000, soit 3,338; 0,408 + 0,097 + 0,913, soit 1,418; 1,197 + 0,950 + 0,062, soit 2,209 et 1,153 + 0,770 + 0,403,

soit 2,326) de paires confirmant l'hypothèse fréquence que ne l'impliquerait l'absence de lien entre la proportion de succès et d'échecs et les divers groupes, d'autre part, moins, plus et plus (Σ respectives = 3,940 + 3,304 + 0,001, soit 7,245; 1,005 + 0,240 + 2,255; soit 3,500; 2,393 + 1,903 + 0,125, soit 4,421 et 2,985 + 1,996 + 1,049, soit 6,030) de paires infirmant la même hypothèse que ne l'impliquerait la même absence de lien. Au reste, ces différences sont, à une exception près, significatives : χ^2 respectifs = 3,338 + 7,245, soit 10,583; 1,418 + 3,500, soit 4,918; 2,209 + 4,421, soit 6,630 et 2,326 + 6,030, soit 8,356 ($P < 0,01$; 0,1; 0,05 et 0,02).

Pour chacun des groupes *1ers antéposés-noms*, *noms-1ers postposés* et *1ers postposés-2es postposés*, CW; HD; LF et CM enregistrent, d'une part, moins, plus et moins ($\Sigma = 0,024 + 0,070 + 0,263$, soit 0,357); moins, plus et moins ($\Sigma = 0,040 + 0,060 + 0,069$, soit 0,169); moins, plus et plus ($\Sigma = 0,544 + 0,383 + 0,047$, soit 0,974) et moins, plus et plus ($\Sigma = 0,121 + 0,019 + 0,356$, soit 0,496) de paires confirmant l'hypothèse longueur que ne l'impliquerait l'absence de lien entre la proportion de succès et d'échecs et les divers groupes, d'autre part, plus, moins et plus ($\Sigma = 0,038 + 0,110 + 0,416$, soit 0,564); plus, moins et plus ($\Sigma = 0,066 + 0,097 + 0,113$, soit 0,276); plus, moins et moins ($\Sigma = 0,761 + 0,535 + 0,065$, soit 1,361) et plus, moins et moins ($\Sigma = 0,214 + 0,033 + 0,623$, soit 0,870) de paires infirmant la même hypothèse que ne l'impliquerait la même absence de lien. Toutefois, ces différences ne sont pas significatives : χ^2 respectifs = 0,357 + 0,564, soit 0,921; 0,169 + 0,276, soit 0,445; 0,974 + 1,361, soit 2,335 et 0,496 + 0,870, soit 1,366 ($P < 0,7$; 0,85; 0,5 et 0,6).

3.2.1.3.3.3. Deux remarques s'imposent à l'esprit lorsqu'on examine les chiffres des §§ 3.2.1.3.3.1 et 3.2.1.3.3.2 et que, dans la totalité RAY des collections ou dans chacune des collections CW; HD; LF et CM, on rapproche systématiquement -

d'hypothèse à hypothèse et de groupe à groupe - les quantités égales à $(r_{ij} - c_{ij})^2/c_{ij}$ qui sont homologues : primo, la proportion de succès et d'échecs est davantage liée aux divers groupes avec l'hypothèse fréquence qu'avec l'hypothèse longueur et, secundo, les succès de l'hypothèse fréquence sont surtout liés au groupe *1ers antéposés-noms* et ses échecs, au groupe *noms-1ers postposés* tandis que les succès de l'hypothèse longueur sont surtout liés au groupe *noms-1ers postposés* et ses échecs, au groupe *1ers antéposés-noms*.

Ces deux dissimilitudes entre l'hypothèse fréquence et l'hypothèse longueur proviennent essentiellement de dissimilitudes entre la distribution du caractère que l'une et la distribution du caractère que l'autre avance comme principe explicatif.

En effet, la série statistique des *NO* et *QSL* dont le lemme compte 1, 2 ..., n occurrences et la série statistique des *NO* et *QSL* dont le lemme compte 1, 2 ..., 16 lettres et traits d'union ne s'avèrent similaires ni par l'étendue ($n > 16$, cf. le §3.2.1.1), ni - mais entrer dans les détails serait trop long - par la configuration de l'histogramme (50).

Or, le volume des succès pour un groupe donné (le volume des échecs pour un groupe donné) dépend partiellement de la probabilité que le premier élément d'une paire de ce groupe soit suivi d'un second élément dont le nombre d'occurrences confirme (dont le nombre d'occurrences infirme) l'hypothèse fréquence ou dont le nombre de lettres et traits d'union confirme (dont le nombre de lettres et traits d'union infirme) l'hypothèse longueur. Et les dissimilitudes entre la distribution du caractère "nombre d'occurrences" et la distribution du caractère "nombre de lettres et traits d'union" font, primo, que la probabilité en cause joue un rôle plus important avec l'hypothèse fréquence qu'avec l'hypothèse longueur, secundo, que la probabilité en

cause est surtout, avec l'une, favorable pour les *1ers antéposés-noms* et défavorable pour les *noms-1ers postposés* et, avec l'autre, défavorable pour les *1ers antéposés-noms* et favorable pour les *noms-1ers postposés*.

3.2.1.3.4. En définitive, quoique ni l'hypothèse fréquence, ni l'hypothèse longueur ne paraissent irrecevables (cf. le § 3.2.1.2), la première doit sans nul doute être préférée à la seconde, car la prédominance de ses succès sur ses échecs est beaucoup plus marquée (cf. le § 3.2.1.3.1) et le volume de ses succès est beaucoup plus grand tandis que le volume de ses échecs est beaucoup plus petit (cf. le § 3.2.1.3.2).

Néanmoins - on ne le perdra pas de vue - le fait que le nombre des lettres et traits d'union, comme le nombre des occurrences, est calculé non sur le mot, mais sur le lemme désavantage tant soit peu l'hypothèse longueur. De surcroît, le système des tests utilisés plus haut désavantage légèrement, lui aussi, cette conjecture (cf. le § 3.2.1.1) : en effet, si les cas où il n'y a ni confirmation, ni infirmation atteignaient avec l'hypothèse fréquence un total supérieur au total réel et si l'ensemble des cas qui engendrent la différence entre ces deux totaux se substituaient à un ensemble de succès et d'échecs dont la proportion égalerait la proportion globale des succès réels et des échecs réels, le χ^2 exprimant la prédominance des succès sur les échecs serait inférieur au χ^2 réel.

Enfin - on ne le perdra pas de vue non plus - l'hypothèse fréquence est moins satisfaisante que l'hypothèse longueur pour ce qui en est de l'absence de lien entre la proportion de succès et d'échecs et les divers groupes (cf. le § 3.2.1.3.3).

3.2.2. En guise de postulat

Une fois terminée la récolte des principaux renseignements que le Tableau II était

susceptible de fournir à l'analyse, quelle portée convient-il donc d'assigner à l'hypothèse fréquence et à l'hypothèse longueur ?

3.2.2.1. Le fait que, sur le plan de la justesse, ces hypothèses donnent, l'une et l'autre, pleine satisfaction a sans doute de quoi surprendre tout d'abord. En vérité, il n'y a pourtant là rien que de très normal, car fréquence et longueur se trouvent être des variables fort étroitement liées (51).

Et le fait que notre première conjecture semble plus pertinente que notre seconde conjecture est également très naturel, car - dans le domaine des mots - la dimension est régie par l'usage plus que l'usage par la dimension :

The result was that word length seemed to be the determining factor for the frequency of use of words, the function varying according to language and, to some extent, also according to the writer : the shorter the word, the more frequent its occurrence. This is in accordance with the artificial codes as we find them, e. g. for telegraphic communication, for shorthand writing, for certain post office communications. But strictly speaking, the position with these codes is somewhat different. It is not the length of the symbol which determines its probability of occurrence but, conversely, its probability of occurrence has been responsible for its length. A brief consideration will show that in language it is precisely the same, if we consider the distribution of words according to length as a product of time (52).

Compatibles et interdépendantes, nos hypothèses ne s'appliquent d'ailleurs pas seulement au langage de Jean RAY, mais bien à l'ensemble du français contemporain puisque (cf. la fin du §2.1) tout ce qui fait leur double validité dans le corpus entier réapparaît - à peu de chose près - dans chacune des diverses parties du corpus. Rien, dès lors, n'interdit à ces deux conjectures de servir de point de départ à quelque postulat linguistique.

3.2.2.2. Au reste, les qualificateurs qui ont été négligés jusqu'ici - les adjectifs non qualificatifs - sont-ils impliqués dans ces relations de fréquence et de longueur ?

Comme ils se succèdent selon des règles très strictes à l'intérieur des *Snho* (53) et occupent généralement une même position vis-à-vis des autres qualificateurs, l'examen du problème n'est plus confiné aux limites d'un corpus dont l'exigüité provenait, en grande partie, des servitudes inhérentes à la localisation des *NO* et *QSL* au sein des multiples espèces de syntagmes : tous les chiffres nécessaires seront fournis par le célèbre *Dictionnaire des fréquences* (54).

Si l'on s'en rapporte à ce dénombrement qui est le plus vaste du domaine français, nos deux hypothèses, appliquées à la façon dont s'ordonnent entre eux les adjectifs non qualificatifs, s'avèrent être de mise. Néanmoins, elles sont loin d'être avalisées partout (ainsi, dans "les quatorze premiers ..." où "quatorze", ayant une fréquence de 0,0000220 et une longueur de 8 lettres, est à la fois moins employé et moins bref que les lemmes de "les" et de "premiers"). Effectivement, certains adjectifs non qualificatifs constituent une classe fort homogène par le sens et les conditions d'emploi. Or, la préséance entre les mots d'une classe de ce genre et d'une classe du même genre paraît dépendre exclusivement des membres de celle-ci et des membres de celle-là qui, parce qu'ils sont les plus utilisés, ont le plus de chances de former paire ensemble. A cet égard, l'exemple qui précède, avec ses deux classes de numéraux, est sans doute très suggestif. Comme le notait J.H. GREENBERG :

Combinning [the successively lower frequency of higher numbers] with the cardinal/ordinal contrast, we would predict that there is a similar hierarchy in ordinals, but that each ordinal is less frequent than the corresponding cardinal. This receives striking confirmation for English from the Lorge Magazine Count [...] Very similar results were obtained for Spanish, French and German (55).

Appliquées à la façon dont s'ordonnent les adjectifs non qualificatifs et les autres qualificateurs, nos deux hypothèses s'avèrent à nouveau, selon le témoignage du *Dictionnaire des fréquences*, entièrement fondées (56) : effectivement, les adjectifs non qualificatifs, en fonction de l'étroitesse de leur paradigme et de l'emploi constant qui en résulte pour eux ainsi qu'en fonction de leur brièveté, apparaissent presque tous comme des "mots-outils" auprès des *NO* et *QSL*.

3.2.2.3. Le rapprochement des remarques que réunit le § 3.2.1 au sujet des qualificateurs à place variable et des remarques que signale le § 3.2.2 au sujet des qualificateurs à place fixe milite en faveur de l'extension de nos deux propositions : **il semble qu'en français contemporain, les qualificateurs d'un même syntagme nominal homogène se succèdent le plus souvent selon l'ordre décroissant de leurs fréquences et - ceci n'étant pas sans lien avec cela - selon l'ordre croissant de leurs longueurs.**

Nous pourrions écrire, en d'autres termes, que les qualificateurs d'un syntagme nominal homogène ont des potentiels d'ouverture (57) directement proportionnels (mais non d'une manière linéaire) à leurs fréquences et - ceci n'étant pas sans lien avec cela - inversement proportionnels (mais pas d'une manière linéaire) à leurs longueurs.

On notera que cette généralisation, avec sa gradation souple, intègre aisément la théorie de H. WEINRICH et sa dichotomie féconde, mais par trop rigide :

Apparemment, l'adjectif, de par sa place variable, se range dans l'une ou dans l'autre des deux classes de monèmes. L'adjectif antéposé fait fonction de morphème, tandis que l'adjectif postposé fait fonction de lexème [...] l'information contenue dans un morphème est très faible, tandis que l'information contenue dans un lexème comme

poète est considérablement plus forte [...] les lexèmes dépassent les morphèmes en nombre, mais [...] les morphèmes dépassent les lexèmes en fréquence (58).

Et on notera que cette généralisation intègre de même certaines observations d'ordre rythmique qui ont été faites à maintes reprises. Ainsi, dans la *Grammaire Larousse du français contemporain* où nous lisons :

- 1^o L'adjectif monosyllabique est placé devant le substantif polysyllabique : *le vieux français* (cf. *le français moderne*), *un gros travail*, *de faux policiers*, etc.;
- 2^o L'adjectif polysyllabique est placé après le substantif monosyllabique : *un cas curieux*; *un film épouvantable*.
[...]
- 3^o L'adjectif monosyllabique a tendance à se placer après un substantif lui aussi monosyllabique : *un lit dur*, *la vie chère* etc.

D'une façon générale, on a tendance à placer en tête du groupe l'élément le plus court (59).

Cela dit - nous aimerions le souligner au terme de la présente approche - **tout se passe, lorsqu'il s'agit d'évoquer "un ensemble au sens mathématique" à l'aide d'un syntagme nominal homogène, comme si les qualificatifs le plus rapidement mobilisés étaient les qualificatifs stockés dans la mémoire de façon à être les plus accessibles et comme si les mots stockés dans la mémoire de façon à être les plus accessibles étaient les mots le plus fréquemment utilisés.**

3.2.2.4. Pour être en droit de voir un lien entre les agencements internes des *Snho* et les différences qu'offrent leurs qualificatifs au point de vue de la fréquence, il serait nécessaire, objectera-t-on assurément, de fournir la preuve que les générations de locuteurs qui ont participé à la fixation de ces agencements-là ont, tant soit peu, perçu ces différences-là.

Or, une telle perception existe.

Et, de surcroît, il n'est pas interdit de penser qu'elle se situe dans une certaine mesure sur le plan de la conscience :

D. H. Howes (1954) a demandé à des sujets (10 à 14 étudiants par groupe) de classer des séries de mots (15 à 25) selon la fréquence avec laquelle ils estimaient les utiliser, et selon la fréquence avec laquelle ils estimaient que ces mots étaient utilisés par le groupe social auquel ils appartenaient. On constate que l'estimation des fréquences pour le groupe calculée à partir des fréquences d'utilisation individuelle est en bon accord avec les tables de fréquences (les corrélations, ρ de Spearman, varient de 0,57 à 0,87). On peut se demander toutefois ce que signifient ces fréquences (60).

[Fraisse, Noizet et Flament (1963)] ont cherché à utiliser une variable de portée plus psychologique que la fréquence tirée des tables, mais qui lui serait liée : la *familiarité* du mot pour le sujet. Deux types d'estimation ont été utilisées : soit la cotation directe sur une échelle en 5 points, soit une technique de comparaisons par paires. Ces deux techniques donnent des estimations de la familiarité qui se trouvent en étroite corrélation. On constate d'autre part une forte corrélation entre les estimations de familiarité et les fréquences tabulées (0,90), bien qu'il y ait certaines distorsions (61).

Mais il est hors de doute, en tout cas, que cette perception joue pleinement au niveau de l'inconscient :

Howes et Solomon (1951) (et Solomon et Howes, 1951) ont montré que le seuil de perception tachistoscopique est une fonction de la fréquence des mots dans la langue [...] La corrélation entre fréquence et seuil est de -0,73 : quand la fréquence décroît, les seuils s'élèvent (62).

Solomon et Postman (1952) ont repris l'expérience avec des mots inconnus des sujets (mots turcs) qui sont présentés aux sujets un certain nombre de fois : de 10 fois pour un groupe à 1 fois pour d'autres. Ils retrouvent, avec ces mots inconnus et une fréquence contrôlée de lecture de ces mots, la même relation (63).

Rosenzweig et Postman (1957) ont montré que la même relation existe entre l'identification des mots masqués par un bruit et leur fréquence [...] Les corrélations obtenues entre le taux d'intelligibilité et le logarithme de la fréquence étaient de -0,78 pour les mots anglais et de -0,65 pour les mots français (64).

Enfin, on voudra bien s'en souvenir, le lien en cause a dû se former d'une manière progressive et insensible : d'une part, chaque génération s'est vue d'autant plus incitée à l'établir que, en héritage, elle le recevait déjà çà et là établi et que (cf. § 3.2.2.2) les membres les plus usuels d'une classe de mots de sens similaire et de modalités d'emploi similaires offrent assez souvent une fréquence similaire, d'autre part, les générations ont été nombreuses dans l'histoire de la langue française !

3.2.2.5. Il sera sans doute encore objecté que, à cause de sa signification pratique et de son principe quasi mécanique, le lien fréquences-arrangement devrait affecter en français les constituants de séquences autres que les *Snho*.

Pour répondre sérieusement à la critique, on serait forcé de procéder à une foule d'enquêtes malaisées et incertaines, mais - si l'on s'en tient à une consultation expéditive du *Dictionnaire des fréquences* - on peut effectivement voir ce lien dans les séquences d'un pronom fonctionnel et d'un adjectif non qualificatif telles que "nous autres" ou "eux deux" (65), les séquences de pronoms conjoints telles que "le lui" ou "t'y" (66), les séquences d'éléments d'un verbe à temps composé et d'un adverbe telles que "a mal mangé" ou "est tombé lourdement", les séquences d'adverbes associés par des démarcateurs de coordination telles que "bien et mal" ou "lentement, mais sûrement", etc.

De même, il sera vraisemblablement objecté que le lien fréquences-arrangement

devrait affecter les qualificateurs des *Snho* en des langues autres que le français.

Admettre le bien-fondé de la critique est assurément imprudent et entreprendre d'y répondre relève davantage de la spéculation que de l'étude scientifique.

Néanmoins, à condition de faire grand cas de la constatation que les membres les plus usuels d'une classe de mots unis par le sens et par les modalités d'emploi sont aussi des mots unis par la fréquence (cf. la fin du § 3.2.2.4), on peut retrouver ce lien en des langues comme l'italien et l'espagnol dont la structure ressemble beaucoup - et pour cause - à la structure de notre langue : de fait, leurs classes de qualificateurs, qui adoptent dans la constitution du syntagme un arrangement analogue à l'arrangement adopté en français (67), enregistrent des fréquences (68) analogues, mutatis mutandis, aux fréquences enregistrées en français.

Si l'on ne tient pas compte du nom, on peut de plus percevoir le lien discuté en des langues comme le néerlandais ou l'anglais qui ont des traits beaucoup moins proches des traits de notre langue : leurs diverses classes d'adjectifs non qualificatifs et de *QSL* se succèdent à l'intérieur du syntagme selon un ordre approximativement comparable à celui du français (69) tandis que, au point de vue de la fréquence, ils présentent des contrastes approximativement comparables à ceux du français (70). Quant au problème posé ici par le nom, il est sans doute à considérer en fonction de la prédilection qu'ont certains parlars pour les macro-unités linguistiques comprises entre les limites d'un ouvreure à terme (71). Les langues germaniques semblent effectivement avoir traité le *Snho* comme l'une de ces unités et converti deux qualificateurs privilégiés - l'article et le nom - en bonnes initiale et finale d'un même ouvreure. Au reste, les problèmes additionnels que pose la place de quelques qualificateurs particuliers ne semblent pas

non plus constituer de difficultés irréductibles. En abordant, par exemple, la célèbre question de "l'article postposé" du suédois, il convient de faire la part des choses : le suffixe *-en*, *-et* ou *-na* a un rôle si effacé qu'il se voit doublé par un article complémentaire - antéposé celui-là - dès qu'un adjectif qualificatif accompagne le nom ("*den varma sommaren, det varma badet, de varma somrarna*"(72)).

Cela dit, on peut toujours deviner le lien discuté en des langues qui, apparentées à la nôtre, ne la rejoignent pourtant aujourd'hui que sur un nombre relativement restreint de points. Mais on est souvent obligé de faire preuve de beaucoup de bonne volonté et on est plus d'une fois amené à conclure que les noms et les adjectifs s'agencent selon des lignes de force toutes différentes (en russe notamment où la construction de l'attribut sans copule et l'emploi d'un jeu de déclinaisons très complexe et très vivace exercent une contrainte non négligeable sur l'ordonnance des qualificateurs).

Il sera peut-être objecté, enfin, que le lien fréquences-arrangement devrait affecter les constituants de séquences autres que les *Snho* en quelques langues autres que le français.

Or, la chose - que les alinéas précédents suggéraient d'ailleurs à leur manière - n'est pas sans rappeler une thèse chère à G. K. ZIPF :

To represent 'I shall go' or '*ich werde gehen*' symbolically, we may use $A b^1 b^2$ [...] we have said that the predicate is inserted between b^1 and b^2 [...] Since the inserted predicate in an independent clause may and frequently does consist of two or more words, the question now arises as to which should come first in the sequence. To this the grammarians answer, 'the least important word comes first' [...] The reader may feel that we are scarcely more definite when, instead of saying that the least important element comes first in the predicate, we say that the least

independent element comes first. But we are really far more definite. We know that the least independent elements are also the relatively more frequent [...] and by the law of abbreviation, that the more frequent are the shorter [...] Automatically come the pronominal objects, the noun objects or the adverbs, and then the phrases (73).

3.3. Quelques à-côtés et quelques ajustements

Maintenant, il ne faudrait pas que l'ensemble de ces réflexions le fasse oublier, notre double conjecture est loin de valoir en toute circonstance.

En effet, on a vu (cf. le § 3.2.1.2.1) que, dans la totalité RAY des collections, 28,148 % des paires infirment l'hypothèse fréquence et 39,908 % des paires infirment l'hypothèse longueur. Ce taux de carence est considérable même si l'on a égard à la relative grossièreté de nos tests. De surcroît, il ne varie pas beaucoup de l'une à l'autre des collections CW; HD; LF et CM : il y est respectivement de 29,060; 27,087; 30,442 et 25,535 % avec la première conjecture et de 32,994; 31,725; 36,930 et 30,122 % avec la seconde conjecture (cf. le § 3.2.1.2.2). L'importance et la constance des échecs en cause donnent à croire que des forces contraires aux forces privilégiées par nos deux hypothèses ont - sur une échelle moindre assurément, mais avec une régularité égale (47) - bel et bien joué.

Une de ces tendances parasites - la propension à figer en lexie des mots cooccurrents (74) - s'est déjà reflétée dans les chiffres des Tableaux I et II.

Lors de l'examen des caractéristiques globales de notre corpus, il s'est avéré (cf. le § 3.1.1.1), d'abord, que la valeur O/N enregistrée par une catégorie donnée était plus élevée que la valeur O/N enregistrée par la catégorie immédiatement consécutive, ensuite, que les valeurs L/N correspondantes se trouvaient entre elles

dans le rapport inverse. En ce qui concerne le corpus entier, une seule exception se manifestait : la valeur O/N des *2es antéposés* était inférieure à celle des *1ers antéposés* et, en ce qui concerne les diverses parties du corpus, 8 exceptions apparaissaient dont 4 - la moitié - étaient identiques à la précédente et dont 1 était fort similaire : dans CW, la valeur L/N des *2es antéposés* se révélait, comparativement à celle des *1ers antéposés*, elle aussi anormale. Au demeurant, lors de l'examen des caractéristiques particulières du corpus, il s'est encore avéré (cf. le § 3.2.1.2.3) que, pour le groupe *2es antéposés-1ers antéposés*, la totalité RAY des collections enregistrait, comme pour tous les autres groupes, plus de paires confirmant nos hypothèses que de paires infirmant celles-ci, mais que, ailleurs, le nombre des échecs était beaucoup moins important. Et l'unique groupe pour lequel la totalité RAY des collections enregistrait plus de paires infirmant notre conjecture préférentielle que de paires infirmant notre conjecture de réserve était précisément le groupe *2es antéposés-1ers antéposés*.

Or, la propension à figer en lexies les paires de qualificateurs qui entrent dans un groupe donné est, primo, conditionnée par la fréquence des deux catégories de qualificateurs que réunit ce groupe, secundo, conditionnée par la fréquence des membres de ces deux catégories. Une telle propension se manifeste donc surtout pour les groupes *1ers antéposés-noms* et *noms-1ers postposés* puisque leur formation ne suppose pas la formation d'un groupe préalable. Mais, si le groupe *noms-1ers postposés* satisfait un peu mieux à la première condition que le groupe *1ers antéposés-noms*, celui-ci l'emporte - et de loin - sur celui-là quant à la seconde condition.

En conséquence, on a envie de rapprocher, d'une part, les anomalies notées au sujet du groupe *2es antéposés-1ers antéposés* (dont maintes paires, pour que nos

hypothèses ne fussent pas aussi souvent démenties, auraient dû voir leurs deux éléments intervertis) et, d'autre part, l'obstacle que la propension à figer en lexies des mots cooccurrents met à la dissociation des deux éléments de maintes paires du groupe *1ers antéposés-noms* (75).

Une autre tendance parasite qu'il ne nous a pas été possible d'ignorer est celle, bien connue, de l'analogie.

Afin de concilier le fait que les cardinaux précédaient toujours les ordinaux et le fait que - irrégulièrement dans la perspective adoptée par nos hypothèses - les cardinaux correspondant à des nombres élevés avaient une fréquence inférieure et une longueur supérieure à la fréquence et à la longueur des ordinaux correspondant à des nombres bas, nous avons considéré (cf. le § 3.2.2.2), en raison de la très grande homogénéité de ces classes de mots, que les cardinaux correspondant à des nombres relativement élevés (qui sont peu courants) avaient cédé à l'attraction exercée par les cardinaux correspondant à des nombres peu élevés (qui sont relativement courants).

La tendance analogique est sans doute le facteur perturbateur qu'on reverra le plus souvent à l'oeuvre. On devra même, apparemment, en tenir compte chaque fois qu'on rencontrera une classe de mots ayant la particularité de toujours précéder ou de toujours suivre une autre classe de mots et d'être caractérisée, de surcroît, par une très grande homogénéité.

Ainsi, conviendra-t-il, selon toute vraisemblance, d'expliquer la postposition constante (76) des adjectifs "de nationalité" en prenant pour point de départ l'attraction exercée sur la poignée d'entre eux qu'on utilise assez souvent par

l'énorme masse de tous ceux dont on ne fait presque jamais usage.

De même, la postposition quasi obligatoire (77) des adjectifs "de couleur" sera-t-elle à interpréter en fonction de l'habitude que nous donne la langue française de commencer par le mot déterminé et de finir par le mot déterminant tout couple de substantifs qui constitue une unité lexicale sans démarcateur interne. D'abord, en effet, une vaste série de substantifs tels que "abricot", "acajou", "amadou", etc. forme avec des adjectifs "de couleur" un paradigme cohérent (78). Ensuite, de multiples mots jouent aussi bien un rôle de substantifs que d'adjectifs "de couleur" : cela dépend uniquement du fait qu'ils sont ou qu'ils ne sont pas qualifiés par un adjectif (comme dans "bleu foncé"), accompagnés d'un nom (comme dans "bleu horizon" et "bleu de nuit"), inclus en une forme composée (comme dans "bleu-noir") ou coordonnés à un nom (comme dans "bleu et or").

Cela dit, c'est à l'aide de données quantitatives que nous avons apprécié la vigueur des forces privilégiées par l'une et par l'autre de nos deux hypothèses. Pour pouvoir saisir de la même façon l'intensité de la tendance analogique qui vient d'être évoquée et l'intensité de quelques tendances parasites d'ampleur sensiblement moindre, nous devons refaire appel à nos anciens inventaires et nous aurons à en consulter de nouveaux. Nous demanderons à ces répertoires de nous indiquer dans quelle mesure l'ordre des *QSL* et des *NO* est lié à leur parenté avec telle ou telle classe de mots (Tableaux I, II et V) ainsi qu'à leurs contextes intrasyntagmatique et extrasyntagmatique (Tableau VI).

Christian DELCOURT

(la suite de cet article paraîtra dans un prochain numéro de *Revue*).

NOTES

- (1) La majorité des codages et la totalité des calculs que réclamait cette étude ont été accomplis par l'ordinateur IBM 360/20 du Laboratoire d'Analyse Statistique des Langues Anciennes de l'Université de Liège. Nous souhaiterions dire ici notre gratitude profonde au créateur et au directeur de ce centre de recherches hautement spécialisé, M. le Professeur L. DELATTE, qui s'est toujours penché sur nos travaux avec la plus grande bienveillance et qui les a enrichis de ses conseils féconds et de ses remarques stimulantes. Nous aimerions également exprimer notre reconnaissance très vive à Melle S. GOVAERTS, à M. J. DENOOZ et aux autres membres de l'équipe du L.A.S.L.A. qui nous ont fait profiter de leur grande expérience et qui nous ont guidé dans la résolution de maints problèmes délicats. Nous nous en voudrions enfin de ne pas rendre hommage à M. le Professeur Et. EVRARD dont les cours nous ont révélé le vaste champ de la linguistique mathématique et dont les encouragements nous ont incité à persévérer dans cette discipline.
- (2) E. SPANG-HANSEN, "Le cas de deux ou plusieurs adjectifs épithètes postposés", *Revue romane*, 2, 1967, 1, pp. 61-68. Cf. aussi notre § 3.2.2.5.
- (3) W. ZWANENBURG, "Reiner, Erwin. La place de l'adjectif épithète en français", *Neophilologus*, 54, 1970, pp. 191-192, p. 191.
- (4) M. WANDRUSKA, "Gedanken zu einer Kritik der romanischen Sprachen",

in K. BALDINGER, éd., *Festschrift Walther von Wartburg zum 80. Geburtstag*, I, Tübingen, M. Niemeyer, 1968, pp. 3-20, p. 15.

- (5) H.J. CEDERGREN et D. SANKOFF, "Variable Rules : Performance as a Statistical Reflection of Competence", *Language*, 50, 1974, 2, pp. 333-355, p. 352.
- (6) J. DAMOURETTE et E. PICHON, *Des mots à la pensée. Essai de grammaire de la langue française*, II, Paris, J.L.L. d'Artrey, 1932, § 528.
- (7) J.-Cl. CHEVALIER et al., *Grammaire Larousse du français contemporain*, Paris, Larousse, 1964, § 312.
- (8) K. WYDLER, *Zur Stellung des attributiven Adjektivs vom Latein bis zum Neufranzösischen*, Berne, A. Francke, 1956, pp. 100-101.
- (9) Cf., par exemple, Ch. MULLER, *Initiation à la statistique linguistique*, Paris, Larousse, 1968, § II. 7 : "l'étendue du vocabulaire est fonction aussi du style : elle est déterminée, en partie du moins, par le lexique du locuteur dans la situation stylistique où il est placé".
- (10) Chr. DELCOURT, "Peut-on prédire la longueur d'une phrase française lorsqu'on connaît la nature de son premier mot ?", *Revue de l'Organisation internationale pour l'étude des langues anciennes par ordinateur*, 1973, 4, pp. 29-60, § 2. Cette publication faisait entrer 2672 phrases en ligne de compte. Nous les avons augmentées ici de 2 blocs de 15 phrases chacun qui, à l'époque, avaient été omis par inadvertance.

- (11) Cf. B. BRAINERD, "An Exploratory Study of Pronouns and Articles as Indices of Genre in English", *Language and Style*, 5, 1972, pp.239-259; B. BRAINERD, "On the Distinction between a Novel and a Romance", *Computers and the Humanities*, 7, 1973, 5, pp. 259-270; H. KUČERA and W. N. FRANCIS, *Computational Analysis of Present-Day American English*, Providence, R. I., 1967; etc.
- (12) Cf., entre autres, G. HERDAN, *The Advanced Theory of Language as Choice and Chance*, Berlin, Springer, 1966, p. 230.
- (13) G. VAN HOUT, *Franc Math. Essai pédagogique sur les structures grammaticales du français moderne*, I, Paris, Didier, 1973, § 1.5.5. Nous dirons de même que "Le signifiant d'une relation, marquant le temps, l'aspect, la modalité est un syntagme verbal" (G. VAN HOUT, *op. cit.*, II, § VII.3.1).
- (14) B. POTTIER, "Note sur le syntagme nominal français", in J. RENSON et M. TYSENS, éd., *Mélanges de linguistique romane et de philologie médiévale offerts à Maurice Delbouille*, I, Gembloux, Duculot, 1964, pp. 517-520, § 1. Nous nous en rapporterons exclusivement aux observations de B. POTTIER, *Introduction à l'étude des structures grammaticales fondamentales*, 2e éd., Nancy, Publications de la Faculté des Lettres et Sciences humaines de Nancy, 1964, §§ 8, 9 et 10 (cf., à titre indicatif, M. TUTEȘCU, *Le groupe nominal et la nominalisation en français moderne*, Paris-Bucarest, Klincksieck-Société roumaine de linguistique romane, 1972, §§ 4.4.1 et 4.4.2).

- (15) M. TUȚESCU, *op. cit.*, § 1. Nous adopterons entièrement les vues qu' expose cet ouvrage dans son chapitre sur les "Contextes distributionnels du groupe nominal". Toutefois, à l'instar de B. POTTIER, *op. cit.* et de M. SALKOFF, *Une grammaire en chaîne du français*, Paris, Dunod, 1973, § 5.4.2 qui, eux aussi, avaient des obligations de nature informative, nous considérerons qu'un infinitif ne forme jamais un syntagme nominal complet, mais seulement une partie de syntagme nominal ou alors un syntagme verbal. Cet a priori nous interdira de faire fond sur le § 1.1.6 de M. TUȚESCU, car la structure # Nom Conjonction Proposition # est absolument inusitée. Il n'en reste pas moins que, pour reprendre un commentaire de G. VAN HOUT, *op. cit.*, II, § VII.5.1, "l'infinitif possède un statut proche de celui du syntagme nominal. En effet, l'infinitif et le syntagme nominal désignent chacun un ensemble. Mais, l'infinitif entre dans un système morphologique particulier, distinct de celui du syntagme nominal, et cela selon la partition temporelle de l'ensemble".
- (16) B. POTTIER, *op. cit.*, § 8.1. Une exception sera quand même faite. B. POTTIER écrit que "La construction nominale hétérogène est signalée par des démarcateurs, dont les principaux sont les suivants : (1) Coordination : *et, ou, ni* (le *mais* lié); (2) Détermination : *de*; (3) Complémentation : *que (qui, dont, où, lequel...), -ant*" (§ 10.1). Ce dernier démarcateur (auquel on pourrait ajouter les finales des participes passés) a peu de points communs avec les autres. Pour des raisons de pure commodité, nous supposerons que le mot où il intervient n'est pas le premier d'une partie de *Snhé*, mais bien le premier d'un syntagme verbal (cf. notre § 2.2.2.1).

- (17) On se reportera, pour ce distinguo, à M. TUȚESCU, *op. cit.*, § 0.2.6 :
 "Nous comprenons par le GN p r i m a i r e le groupe nominal constituant immédiat de la phrase, alors que le GN s e c o n d a i r e est conçu comme le constituant d'un autre groupe (nominal, prédicatif, prépositionnel, adverbial)".
- (18) Ces formules proviennent de G. VAN HOUT, *op. cit.*, I (§§ II.2.3., II.3.2 (20) et II.3.3), mais elles ont été TRONQUEES ET ALTEREES afin de répondre aux besoins spécifiques de la présente étude. G. VAN HOUT, par (21) exemple, opposait très nettement qualificateurs et quantificateurs (§ II.2.3). Nous, nous annexons globalement les seconds aux premiers : la démarche nous facilitera, en effet, la proposition ultérieure d'un modèle du *Snho* valable pour divers états de la langue française. Cela dit, rappelons que "ssi" est l'abréviation consacrée pour "si et seulement si" et qu'on appelle "sous-ensemble non trivial de l'ensemble E" tout sous-ensemble d'E à l'exclusion d'E lui-même et de l'ensemble vide \emptyset .
- (19) Cf. B. POTTIER, *op. cit.*, § 9.2 : "La *quantification* est exprimée par les quantitatifs (à tort appelés "adverbes"), qui indiquent le degré de l'adjectivation".
- (22) Cf. Ch. MULLER, *op.cit.*, § II.5 : "Aussi est-il sage, à ce stade, d'adopter une norme pratique, sans grandes prétentions scientifiques, ce qui conduit à la rattacher autant que possible à un ouvrage de référence".
- (23) A ce sujet, on consultera, par exemple, H. MITTERAND, "Grammaire française : observations sur les pré-déterminants du nom", *Etudes de*

linguistique appliquée, 2, 1963, pp. 126-134.

- (24) J.-J. TAVERNIER, *Nouvelle grammaire-index de la langue néerlandaise*, 7e éd., Bruxelles, A. De Boeck, 1955.
- (25) R.W. ZANDVOORT, *Grammaire descriptive de l'anglais contemporain*, Lyon-Paris, IAC, 1949.
- (26) A juste titre, semble-t-il d'ailleurs. Cf., entre autres, G. VAN HOUT, *op. cit.*, I, § VII.7.5.
- (27) Voir à ce propos M. GREVISSE, *Le bon usage. Grammaire française avec des remarques sur la langue française d'aujourd'hui*, 8e éd., Gembloux, J. Duculot, 1964, §§ 769 et 771.
- (28) Cette liste fait seulement état des repères qui serviront à distinguer *QPP* et participes passés dans les bornes de notre corpus. Pour un corpus différent, elle serait sans doute différente.
- (29) B. POTTIER, *op. cit.*, § 10.4.
- (30) Cf. M. SALKOFF, *op. cit.*, § 4.1.1.1.
- (31) Cf. M. GREVISSE, *op. cit.*, § 326 : "L'article partitif n'est autre chose, pour le sens, qu'un article indéfini".

- (32) Si nous n'avons pas évoqué certaines combinaisons possibles (un *QSL* précédé d'un quantitatif avant le nom et un *QSL* après (*dana*), plusieurs *QSL* avant le nom et plusieurs *QSL* après (*aanaa*), au moins un *QSL* affecté d'un complément avant le nom et au moins un *QSL* après (*acna*), au moins un *QSL* affecté d'un complément avant le nom et au moins un *QSL* affecté d'un complément après (*acnac*)), c'est seulement parce que notre corpus n'en contient aucun exemple. Par ailleurs, il n'a pas été prévu de désignation spécifique pour les *Snho* et *Snhé* assimilés qui n'offrent pas un *QSL* au moins.
- (33) Aucune des données numériques dont fait état cet article n'a été arrondie.
- (34) Nous n'avons pas tenu compte du seul *4e postposé* que comprenait notre corpus. Son indice O/N était de 12 et son indice L/N, de 4.
- (35) Rappelons que

$$\sum_{i=1}^n x_i = x_1 + x_2 + x_3 \dots + x_{n-1} + x_n$$

- (36) Quoique nous prenions aussi en considération les syntagmes non nominaux, nos relevés d'occurrences auraient une signification trop incertaine sur une base moins large que RAY.

(37) Pour ne pas encombrer la mémoire de la machine, nous avons adopté un mode d'input qui rogne à droite les querques lemmes faisant plus de 16 lettres et traits d'union.

(38) Il s'agit d'une forme particulière - avec $p = 1/2 = q$ -

de la somme $C_n^k p^k q^{n-k} + C_n^{k+1} p^{k+1} q^{n-k-1} + \dots + C_n^{n-1} p^{n-1} q + C_n^n p^n q^0$ qui établit la probabilité d'un minimum de k succès sur n épreuves (voir, à ce propos, A. MONJALLON, *Introduction à la méthode statistique*, 6e éd., Paris, Librairie Vuibert, 1969, §§ 78 et 79 ainsi que Ch. MULLER, *op. cit.*, § 1.4). Rappelons que $C_n^k = n! / [k! (n-k)!]$ et que $n! = 1 \times 2 \times 3 \dots \times n-1 \times n$.

(39) Nous n'avons pas tenu compte de la seule paire *3e postposé-4e postposé* que comprenait notre corpus. Elle infirmait l'hypothèse fréquence et ne confirmait, ni n'infirmit l'hypothèse longueur (cf. n. 34).

(40) Lorsqu'on veut obtenir le total des paires enregistrées dans une collection donnée et pour un ou plusieurs groupes donnés, il suffit d'additionner les nombres NBR qui apparaissent au Tableau II dans cette collection-là et pour ce ou ces groupes-là.

- (41) Lorsqu'on veut obtenir le total des paires qui confirment, qui ne confirment, ni n'infirment ou qui infirment une conjecture donnée dans une collection donnée et pour un ou plusieurs groupes donnés, il faut, selon le cas, traduire en nombres réels les pourcentages C, N ou I qui apparaissent au Tableau II avec cette hypothèse-là, dans cette collection-là et pour ce ou ces groupes-là et il faut additionner les nombres ainsi établis.
- (42) Cf. A. MONJALLON, *op. cit.*, § 112, Ch. MULLER, *op. cit.*, § 1.15 ainsi que J.-J. PINTY et Cl. GAULTIER, *Dictionnaire pratique de mathématiques et statistiques en sciences humaines*, Paris, Editions Universitaires, 1971, s.v. "Distribution de χ^2 (Ki deux)".
- (43) Elle est reprise dans la plupart des traités de statistique et, notamment, dans G. HERDAN, *Language as Choice and Chance*, Groningen, P. Noordhoff, 1956, p. 341, A. MONJALLON, *op. cit.*, p. 273 et Ch. MULLER, *op. cit.*, p. 241.
- (44) G. HERDAN, *op. cit.*, pp. 91-95. La lettre σ qui est utilisée à cet endroit désigne l'écart-type d'une population gaussienne. Pour plus de renseignements, voir G. HERDAN, *op. cit.*, pp. 312-315.
- (45) Cf. A. MONJALLON, *op. cit.*, § 112 : "L'effectif total N doit être assez grand, 50 au moins. L'effectif de chaque classe ne doit pas être inférieur à 5, mais bien plus grand si possible".
- (46) Résultat tout différent de ce que laissait attendre notre § 3.1.1.1.

- (47) Abstraction étant faite du nombre des fois où ni une confirmation, ni une infirmation n'est acquise, la chose ne va *nullement* de soi.
- (48) Quoique exceptionnel, le sens de ce rapport était annoncé par les modalités du phénomène évoqué dans notre § 3.2.1.3.1.
- (49) Le sens de ce rapport est exceptionnel et n'était pas annoncé par les modalités du phénomène évoqué dans notre § 3.2.1.3.1. La supériorité numérique des échecs de l'hypothèse fréquence sur les échecs de l'hypothèse longueur n'a toutefois pas beaucoup d'importance ici ($\chi^2 = 0,014$ ($P < 0,95$)). Elle vient, au reste, illustrer à sa façon ce que disait notre n. 47.
- (50) Cf. A. MONJALLON, *op. cit.*, § 19 : "Une classe donnée a pour image un rectangle dont la base est le segment de l'axe Ox compris entre les frontières de cette classe et dont la hauteur est mesurée par l'effectif correspondant sur l'axe Oy. Un tel rectangle ayant été construit pour chaque classe de la série statistique, nous nous trouvons en présence d'un graphique formé de rectangles adjacents, de même largeur et dont les hauteurs représentent les divers effectifs des classes. On appelle alors *histogramme* de la série la ligne bordant l'ensemble des rectangles".
- (51) G.K. ZIPF, *The Psycho-Biology of Language*, 2e éd., Cambridge (Massachusetts), The M.I.T. Press, 1965, ch. II, 1ère partie. Pour le français, voir entre autres P. GUIRAUD, *Problèmes et méthodes de la statistique linguistique*, Dordrecht, D. Reidel Publishing Company, 1959, ch. VII.

- (52) G. HERDAN, *The Advanced Theory of Language as Choice and Chance*, Berlin, Springer, 1966, p. 295.
- (53) Cf. à ce sujet M. SALKOFF, *op. cit.*, § 4.1.1.
- (54) P. IMBS et coll., *Dictionnaire des fréquences, I, Table alphabétique*, 4 vol. et *IV, Table de répartition des homographes*, 1 vol., Paris, Didier, 1971.
- (55) J.H. GREENBERG, *Language Universals*, La Haye-Paris, Mouton, 1966, pp. 42-43.
- (56) On ne verra pas un argument contre la première de nos conjectures dans des syntagmes tels que "un mien cousin" si, du moins, on veut bien tenir compte de la fréquence qu'avaient les adjectifs "mien", "tien", etc. à l'époque où les lexies en cause appartenaient aux lexies vivaces de la langue française. Soit dit par parenthèse, des syntagmes tels que "un mio cugino" n'ont rien d'obsolescent et l'hypothèse fréquence qui est peut-être aussi applicable à l'italien (cf. notre § 3.2.2.5) s'en accommode parfaitement.
- (57) Par "potentiel d'ouverture" d'un mot, nous entendons ici le nombre moyen de lemmes séparant la place que ce mot occupe dans une unité linguistique et la fin de cette dernière. Nous avons décidé ailleurs (Chr. DELCOURT, *loc. cit.*, § 5) de dire que divers éléments du discours qui enregistrent, dans des circonstances analogues, un même potentiel d'ouverture constituent des ouvreurs d'un même type, mais l'acception du vocable "ouvreur" est tout autre, par exemple, chez P. CHARAUDEAU (*Etudes de linguistique appliquée*, 11, 1973, pp. 5-70) qui, du reste, ne l'utilise nullement comme M. COYAUD

(*Linguistique et documentation*, Paris, Larousse, 1972, § 3.1.).

- (58) H. WEINRICH, "La place de l'adjectif en français", *Vox romanica*, 25, 1966, pp. 82-89. B. SCIARONE s'en est violemment pris à l'usage que cet article fait du mot "information" : "cela n'a rien à voir avec la signification du mot" ("L'adjectif en français moderne", *Zeitschrift für Romanische Philologie*, 83, 1967, pp. 583-598, p. 588). Toutefois, comme il l'écrivait lui-même "plus un signe est fréquent, plus l'occurrence de ce signe est probable et moins il donne d'information" et, comme il le rappelait en citant P. GUIRAUD, "il existe des relations numériques stables entre l'extension du sens des mots et leur fréquence" (*Ibidem*; cf. G.K. ZIPF, *Human Behavior & The Principle Of Least Effort*, réimpr., New York, Hafner Publishing Company, 1972, pp. 30-31 : "a law-of-meaning distribution exists according to which the m average number of meanings per word of a thousand words (when ranked in the order of decreasing frequency) will equal the square root of the average frequency of the words' occurrence (or will decrease according to the square root of the rank)"). Dès lors ...
- (59) J.-Cl. CHEVALIER et al., *op. cit.*, § 314.
- (60) P. FRAISSE, J. PIAGET et coll., *Traité de psychologie expérimentale*, VIII, 2e éd., Paris, PUF, 1972, p. 65 (cf. D. H. HOWES, "On the interpretation of word frequency as a variable affecting speed of recognition", *Journal of experimental Psychology*, 48, 1954, pp. 106-112).
- (61) *Ibidem* (cf. P. FRAISSE, G. NOIZET et C. FLAMENT, "Fréquence et familiarité du vocabulaire", in J. de AJURIAGUERRA et al., *Problèmes de psycholinguistique*, Paris, PUF, 1963, pp. 157-167).

- (62) P.FRAISSE, J. PIAGET et coll., *op. cit.*, p. 66 (cf. D. H. HOWES et R. L. SOLOMON, "Visual duration threshold as a function of word probability", *Journal of experimental Psychology*, 41, 1951, pp. 400-410 ainsi que R. L. SOLOMON et D. H. HOWES, "Word frequency, personal values, and visual duration threshold", *Psychological Review*, 58, 1951, pp. 256-270).
- (63) *Ibidem* (cf. R. L. SOLOMON et L. POSTMAN, "Frequency of usage as a determinant of recognition thresholds for words", *Journal of experimental Psychology*, 43, 1952, pp. 195-201).
- (64) P. FRAISSE, J. PIAGET et coll., *op. cit.*, pp. 66 et 67 (cf. M. R. ROSEN-ZWEIG et L. POSTMAN, "Intelligibility as a function of frequency of usage", *Journal of experimental Psychology*, 54, 1957, pp. 412-422).
- (65) Cf. G. VAN HOUT, *op. cit.*, I, § VI, 8.1. : "Nous appelons pronoms fonctionnels, les pronoms dont la forme varie selon la fonction". Au reste, la préposition est relativement récente entre les nominaux autonomes (cf. G. VAN HOUT, *op. cit.*, I, § VI, 3.1 : "ce sont des syntagmes nominaux non disponibles") et les QSL. Autrefois, on disait "rien mauvais", "quelque chose grande", etc. (voir à ce propos J.A. QUILLACQ, *La langue et la syntaxe de Bossuet*, réimpr., Genève, Slatkine Reprints, 1968, pp. 85 et 731) et le lien fréquences-arrangement ne paraît pas étranger à ces séquences originelles non plus. On notera au passage que les langues germaniques postposent régulièrement le QSL qui accompagne un nominal autonome (*anything else, something big, ..., iets lekkers, weinig moois, ...*) et ce, en dépit du fait qu'elles antéposent presque toujours l'adjectif qui accompagne un nom.

(66) On voudra bien consulter à ce sujet G. VAN HOUT, *op. cit.*, II, § XIV, 3.

(67) Les exemples qu'allègue M. WANDRUSKA, "Die Integration des Adjektivs oder bonnet blanc et blanc bonnet", *Zeitschrift für französische Sprache und Literatur*. 75, 1965, 1, pp. 145-163 sont suggestifs à cet égard. Il y a quand même, entre le français et l'italien ou l'espagnol, des différences aussi importantes que celle-ci :

Im Französischen hat sich die im Partizip enthaltene verbale Dynamik in den meisten Fällen der engeren Fügung *a-S* widersetzt, und wenn *sacré, maudit, prétendu, distingué* usw. dem Substantiv vorausgehen können, so hat gerade in dieser Verwendung das Adjektiv jedes verbale Element verloren, das nur noch in einigen formelhaften Wendungen lebt, wie *le détesté rival, le regretté collègue* [...]. In der südlichen Romania hat dagegen das Partizip die gleiche Beweglichkeit wie das Adjektiv und kann mühelos auch dem Substantiv vorausgehen (p. 160).

Cf. également M. WANDRUSKA, "Gedanken zu einer Kritik der romanischen Sprachen", in K. BALDINGER, éd., *op. cit.*, pp. 3-20.

(68) Cf. les chiffres publiés par Alph. JUILLAND et V. TRAVERSA, *Frequency Dictionary of Italian Words*, La Haye-Paris, Mouton, 1973 ainsi que par Alph. JUILLAND et E. CHANG-RODRIGUEZ, *Frequency Dictionary of Spanish Words*, La Haye-Paris, Mouton, 1964.

(69) Cf. les comparaisons de J. WEISSHAUPT, "Place de l'adjectif en néerlandais et en français", in E. NIEUWBORG, éd., *Mélanges J. L. Pauwels*, Louvain, Publications Universitaires de Louvain, 1970, pp. 109-119 : "En français les adjectifs à signification absolue semblent également préférer les places extrêmes" (p. 115), "En français, il y a une symétrie par rapport au néerlandais" (p. 117), "il me semble qu'une certaine tendance peut apparaître

aussi bien en français qu'en néerlandais" (p. 118), etc. Et la description de l'anglais que donne Z. VENDLER, *Adjectives and Nominalizations*, La Haye-Paris, Mouton, 1968, 2e partie n'est pas sans rappeler, entre autres, celle du français que donne R. C. KELLEY, "The Order of Preposed Adjectives in French", *The French Review*, 43, 1970, 5, pp. 783-794. Quant aux similitudes du néerlandais et de l'anglais, elles apparaissent clairement dans J. VAN ROEY, "On the Phrase 'Adjective(s) + Noun' in Modern English and Dutch", in El. NIEUWBORG, éd., *op. cit.*, pp. 61-79.

- (70) Cf. J. LINSCHOTEN, *De la Court's Frekwentietelling van nederlandse woorden*, Utrecht, Psychologisch Laboratorium Rijksuniversiteit Utrecht, [1963] ainsi que E. L. THORNDIKE et I. LORGE, *The Teacher's Word Book of 30,000 Words*, New York, Bureau of Publications, Teacher's College, Columbia University, 1944.
- (71) Si l'on esquissait une partition des ouvreurs tels qu'ils ont été définis plus haut (cf. notre n. 57), on serait amené sans doute (cf. Chr. DELCOURT, *loc. cit.*, n.46) à distinguer - outre les mots et locutions sans potentiel propre ("remplisseurs de pause", incises, etc.) - ce que, faute de mieux, on pourrait appeler, primo, des ouvreurs à relance ("(une nourriture) non seulement (riche,) mais encore (très variée)"), secundo, des ouvreurs à terme ("ces(trois personnes-) ci" ou , à la rigueur, "nous (ven l)ons"), tertio, des ouvreurs à suite quasi obligée (la plupart des mots-outils en français) et, quarto, les ouvreurs à suite facultative (la plupart des mots pleins en français). Cela dit, à l'intérieur de deux systèmes linguistiques différents, une espèce d'ouvreurs déterminée n'a ni le même emploi (les ouvreurs à suite facultative, par exemple, opèrent "sur une moins longue distance" dans les langues à déclinaisons pauvres que dans les langues à déclinaisons riches), ni la même importance : ainsi les ouvreurs qui nous intéressent principalement ici, les ouvreurs

à terme, sont-ils très nombreux en néerlandais (que l'on songe aux adverbes pronominaux, aux paires verbe-particule séparable, mot de subordination-verbe, auxiliaire-participe passé, verbe principal à l'indicatif-verbe subordonné à l'infinitif, etc.), plus occasionnels en anglais (on y rencontre cependant toute une série de tournures du genre *whom (did you give the letter) to*) et rares - sauf aux niveaux de l'orthographe, de la morphologie (cf. K. BALDINGER, "Post- und Prädeterminierung im Französischen", in K. BALDINGER, éd., *op. cit.*, pp. 87-106) et de la prosodie (cf. notamment B. MALMBERG, *Les domaines de la phonétique*, Paris, P.U.F., 1971) - en chacune des langues de la Romania occidentale.

(72) Cf. Om. BJÖRKHAGEN, *Modern Swedish Grammar*, 8e éd., Norstedts-Stockholm, Svenska Bokförlaget, 1956, § 79.

(73) G.K. ZIPF, *The Psycho-Biology of Language*, 2e éd., Cambridge (Massachusetts), The M.I.T. Press, 1965, ch. V, § II. 4.

(74) B. POTTIER, *op. cit.*, §§ 1.1 et 1.2 :

le *coefficient de cohérence* entre les éléments d'une suite de mots peut varier de 0 à 100. Les emplois répétés de certaines associations finissent par s'inscrire dans l'inventaire lexical des sujets parlants, aux côtés des mots simples [...] le coefficient de cohérence peut être estimé grâce à des critères fonctionnels et, pour les cas plus délicats, en se fondant sur des relevés statistiques venant confirmer les "sentiments linguistiques".

(75) Sur cette base, on pourrait émettre des suppositions qui seraient entièrement différentes des nôtres et rejoindraient notamment la conception formulée par J. W. OLLER, *Coding Information in Natural Languages*,

La Haye-Paris, Mouton, 1971, § 11.2.1. :

The Nsg *woman* is modified by the Adj *young* to form the unit *young woman* which in turn is modified by the unit *very extraordinarily beautiful*. The latter unit consists of an Adj *beautiful* which is modified by the unit *very extraordinarily* which consists of an Adv *extraordinarily* modified by an Int *very*. The unit formed then by the modification of *young woman* by *very extraordinarily beautiful* is further modified by the Det *the*.

Mais ce serait accorder beaucoup de poids à un phénomène qui, avant tout, est local et résiduel - nos hypothèses offrent quand même plus de réussites que d'échecs pour le groupe *2es antéposés-1ers antéposés* - et qui, sans doute, tient principalement à des imperfections propres à l'enregistrement de données linguistiques : saura-t-on jamais déterminer où finit la lexie et où commence le mot ?

- (76) A dire vrai, l'emploi que nous faisons du terme "constant" est abusif, car on trouve chez les poètes quelques tournures comme "Hollandaise tulipe rouge" (M. CLEMEUR, *Nous n'irons plus à Palalda*, Paris, Ed. de la Cité, 1959, p. 32) ou comme "La flamande clameur" (J. MOULIN, *Rue Chair et Pain*, Paris, P. Seghers, 1961, p.7) et on en découvre même, de temps à autre, chez les prosateurs : M. BODEUX, par exemple, a intitulé un de ses ouvrages *Liégeoise Idylle* (Bruxelles, Vromant et C^o, 1905) et de "flamandes hyperdulies" surgissent à la page 147 de C. HANLET, *Les écrivains belges contemporains de langue française 1800-1946*, I, Liège, H. Dessain, 1946. Toutefois, il ne s'agit évidemment là que d'un petit lot de curiosités plaisantes.

- (77) Les infractions - qui, ici également, sont surtout commises par les textes poétiques - s'avèrent bien moins exceptionnelles pour les adjectifs "de couleur" que pour les adjectifs "de nationalité". La survie dans le français moderne de lexies du genre : "blanc-bec", "blanc-estoc", "blanc-manger", "blanc-manteau", "blanc-seing", "rouge-gorge", "rouge-queue" ou même "verljus" n'est pas étrangère à la chose. Mais l'élément prépondérant réside sans nul doute ailleurs : la fréquence des adjectifs "de couleur" est beaucoup plus grande, effectivement, que la fréquence des adjectifs "de nationalité".
- (78) Le paradigme inclut 88 substantifs au moins (cf. Ad. THOMAS, *Dictionnaire des difficultés de la langue française*, Paris, Larousse, 1956, s.v. "couleur"). D'autre part, il est si homogène et les cas d'antéposition de l'adjectif "de couleur" sont - toutes proportions gardées - si répandus que quelques-uns de ces substantifs se retrouvent à l'occasion devant le substantif qu'ils déterminent (cf. par exemple H. CORNELUS, *Mer en Terre*, Bruxelles, André de Rache, 1973, p. 35 où il est question d'un "orange leurre").

TABLEAU I.

	2es antéposés				1ers antéposés				noms			1ers postposés				2es postposés				3es postposés							
	QSS	QNT	QPP	QSL	QSS	QNT	QPP	QSL	NP	NC	NO	QSS	QNT	QPP	QSL	QSS	QNT	QPP	QSL	QSS	QNT	QPP	QSL	QSS	QNT	QPP	QSL
CW	NMB	8	-	-	8	312	14	2	328	11	690	701	322	92	66	420	21	3	4	28	2	1	1	4	2	1	4
	OCC	134	-	-	134	12256	60	14	12330	38	9268	9306	2793	110	157	3010	171	5	21	197	6	2	1	9	6	2	1
	LET	46	-	-	46	1807	115	10	1932	73	4182	4255	2084	249	416	2749	126	30	21	177	12	8	6	2	1	2	2
	O/N	16,75	-	-	16,75	39,28	4,28	7	37,59	3,45	13,43	13,27	8,51	3,43	2,37	7,16	8,14	1,66	5,25	7,03	3	2	1	2,2	1	2	1
	L/N	5,75	-	-	5,75	5,79	8,21	5	5,89	6,63	6,06	6,06	6,47	7,78	6,30	6,54	6	10	5,25	6,32	6	8	6	6,5	8	6	6,5
HD	NMB	7	-	1	8	208	7	-	215	13	456	469	223	19	48	290	16	2	6	24	2	-	-	1	-	-	1
	OCC	174	-	7	181	7314	57	-	7371	98	6740	6838	1406	58	93	1557	143	4	27	174	5	-	-	5	-	-	5
	LET	41	-	5	46	1228	55	-	1283	79	2929	3008	1625	155	321	2101	120	22	41	183	19	-	-	19	-	-	19
	O/N	24,85	-	7	22,62	35,16	8,14	-	34,28	7,53	14,46	14,57	6,30	3,05	1,93	5,36	8,93	2	2,5	7,25	2,5	-	-	2,5	-	-	2,5
	L/N	5,85	-	5	5,75	5,90	7,85	-	5,96	6,07	6,28	6,41	7,28	8,15	6,68	7,24	7,5	11	6,83	7,62	9,5	-	-	9,5	-	-	9,5
LF	NMB	11	-	-	11	265	30	1	296	9	601	610	276	22	52	350	17	4	6	27	2	-	1	3	-	-	3
	OCC	243	-	-	243	8481	147	2	8630	37	7276	7313	1847	77	139	2063	91	10	10	111	5	-	1	6	-	-	6
	LET	62	-	-	62	1559	250	6	1815	71	3825	3896	1910	157	356	2423	130	29	36	195	17	-	8	25	-	-	25
	O/N	22,09	-	-	22,09	32,00	4,9	2	29,15	4,11	12,10	11,98	7,46	3,5	2,67	5,89	5,35	2,5	1,66	4,11	2,5	-	1	2	-	-	2
	L/N	5,63	-	-	5,63	5,88	8,33	6	6,13	7,88	6,36	6,38	6,89	7,13	6,84	6,92	7,64	7,25	6	7,22	8,5	-	8	8,33	-	-	8,33
CM	NMB	14	2	-	16	270	14	3	287	12	569	581	248	33	38	319	21	5	4	30	2	-	-	2	-	-	2
	OCC	428	19	-	447	10954	49	21	10924	30	8313	8343	1940	87	112	2139	150	6	4	160	5	-	-	5	-	-	5
	LET	89	13	-	102	1593	113	15	1721	105	3588	3693	1643	257	238	2138	149	49	24	222	9	-	-	9	-	-	9
	O/N	30,57	9,5	-	27,93	40,2	3,5	7	38,06	2,5	14,60	14,35	7,82	2,63	2,94	6,70	7,14	1,2	1	5,33	2,5	-	-	2,5	-	-	2,5
	L/N	6,35	6,5	-	6,37	5,9	8,07	5	5,99	8,75	6,30	6,35	6,62	7,78	6,26	6,70	7,09	9,8	8	7,4	4,5	-	-	4,5	-	-	4,5
RAY	NMB	40	2	1	43	1055	65	6	1126	45	2316	2361	1069	106	204	1379	75	14	20	109	8	1	2	11	2	2	11
	OCC	979	19	7	1003	38905	313	37	39255	203	31597	31800	7936	352	501	8769	555	25	62	642	21	2	2	25	2	2	25
	LET	238	13	5	256	6187	533	31	6751	328	14524	14852	7262	818	1331	9411	525	130	122	777	57	8	14	79	8	14	79
	O/N	24,47	9,5	7	23,37	36,87	4,81	6,16	34,86	4,51	13,64	13,46	7,42	3,13	2,45	6,35	7,4	1,78	3,1	5,88	2,62	2	1	2,27	2	1	2,27
	L/N	5,95	6,5	5	5,95	5,86	8,2	5,16	5,99	7,28	6,27	6,29	6,79	7,71	6,52	6,82	7	9,28	6,1	7,12	7,12	8	7	7,18	8	7	7,18

Ce tableau concerne exclusivement les qualificatifs au sens large et noms issus des syntagmes nominaux homogènes et syntagmes nominaux hétérogènes assimilés qui, dans notre corpus, comptent un ou plusieurs qualificatifs au sens large. Ils ont d'abord été catégorisés selon leur place à l'intérieur des syntagmes. Ensuite, les uns (QSL) ont été sous-catégorisés en qualificatifs au sens strict (QSS), qualificatifs à finale "ant" ou "ant" (QNT) et qualificatifs-participes passés (QPP). Les autres (NO) ont été sous-catégorisés en noms communs (NC) et noms propres (NP) (cf. notre § 2.2.2.1).

Le tableau signale pour les parties du corpus qui procèdent des *Contes du whisky* (CW), de la série *Harry Dickson* (HD), du *Livre des fantômes* (LF) et du *Carrousel des maléfices* (CM) et pour le corpus tout entier (RAY) les chiffres qu'offrent les diverses (sous-)catégories de QSL et NO aux points de vue du nombre de membres (NMB) et des indices global de fréquence (OCC), global de longueur (LET), moyen de fréquence (O/N), et moyen de longueur (L/N) (cf. notre § 3.1).

TABLEAU II

			2es antéposés-1ers antéposés	1ers antéposés-noms	noms-1ers postposés	1ers postposés-2es postposés	2es postposés-3es postposés
CW	HF	C	37, 5	71, 34	55, 47	60, 71	75
		N	-	4, 57	12, 14	10, 71	-
		I	62, 5	24, 08	32, 38	28, 57	25
		NBR	8	328	420	28	4
	HL	C	37, 5	52, 13	52, 61	39, 28	75
		N	12, 5	13, 71	15, 47	25	25
I		50	34, 14	31, 90	35, 71	-	
HD	HF	C	25	72, 09	62, 06	41, 66	100
		N	-	3, 72	10, 68	20, 83	-
		I	75	24, 18	27, 24	37, 5	-
		NBR	8	215	290	24	2
	HL	C	50	48, 37	53, 79	50	100
		N	25	20	14, 48	12, 5	-
I		25	31, 62	31, 72	37, 5	-	
LF	HF	C	63, 63	68, 58	54, 85	55, 55	66, 66
		N	-	4, 72	11, 71	11, 11	33, 33
		I	36, 36	26, 68	33, 42	33, 33	-
		NBR	11	296	350	27	3
	HL	C	72, 72	48, 31	53, 14	59, 25	66, 66
		N	-	11, 82	12, 85	3, 70	33, 33
I		27, 27	39, 86	34	37, 03	-	
CM	HF	C	75	74, 91	60, 50	46, 66	100
		N	-	3, 48	10, 65	23, 33	-
		I	25	21, 60	28, 84	30	-
		NBR	16	287	319	30	2
	HL	C	50	49, 82	53, 91	63, 33	-
		N	12, 5	19, 51	16, 30	13, 33	50
I		37, 5	30, 66	29, 78	23, 33	50	
RAY	HF	C	55, 81	71, 66	57, 86	51, 37	81, 81
		N	-	4, 17	11, 38	16, 51	9, 09
		I	44, 18	24, 15	30, 74	32, 11	9, 09
		NBR	43	1126	1379	109	11
	HL	C	53, 48	49, 82	53, 29	53, 21	63, 63
		N	11, 62	15, 89	14, 79	13, 76	27, 27
I		34, 88	34, 28	31, 90	33, 02	9, 09	

Ce tableau concerne les paires de qualificateurs dont le premier élément appartient à l'une des six catégories de *QSL* et *NO* du Tableau I et dont le second élément appartient à la catégorie immédiatement consécutive. Elles ont été groupées selon leur place à l'intérieur des syntagmes.

Le tableau signale pour CW, HD, LF et CM et pour RAY les chiffres qu'offrent les divers groupes de paires de *QSL* et *NO* aux points de vue, primo, du pourcentage de confirmations (C), de non-confirmations-non-infirmations (N) et d'infirmations (I) de l'hypothèse fréquence (HF), secundo, du nombre des membres (NBR) et, tertio, du pourcentage de confirmations (C), de non-confirmations-non-infirmations (N) et d'infirmations (I) de l'hypothèse longueur (HL) (cf. notre § 3.2).

Tableau III	HFT	HLT	Tableau IV	HFT	HLT
c	314,5	314,5	r	223	406

Ces tableaux concernent les paires de *QSL* et *NO* qui ne confirment, ni n'infirmement l'une ou l'autre de nos deux conjectures.

Le tableau III (le tableau IV) signale le nombre calculé *c* de ces paires (le nombre recensé *r* de ces paires) que la population formée par tous les cas où l'hypothèse fréquence a été testée - HFT - et par tous les cas où l'hypothèse longueur a été testée - HLT - devrait compter si elle était parfaitement homogène (compte dans la réalité).

Cf. notre § 3.2.1.1.

UNE APPROCHE DES PROBLEMES QUE SOULEVE EN FRANCAIS
CONTEMPORAIN LA PLACE DES ADJECTIFS QUALIFICATIFS EPITHETES.

Erratum.

Les deux tableaux qui apparaissent à la page 117 de cet article auraient dû avoir la forme:

	HFT	HLT		HFT	HLT
c	314,5	314,5	r	223	406
	2353,5	2353,5		2445	2262

Il convient dès lors de corriger la première ligne de la page 72 en écrivant:

" $[(223 - 314,5)^2 + (406 - 314,5)^2]/314,5] + [(2445 - 2353,5)^2 + (2262 - 2353,5)^2]/2353,5]$ et vaudra 60,355."

Nous prions le lecteur de bien vouloir nous pardonner cette erreur.

Un rectificatif qui portera sur le test de Pearson sera inclus dans la seconde partie de notre article.

Christian DELCOURT.