

LE LIVRE D'ISAIE ET L'ANALYSE CRITIQUE DES SOURCES TEXTUELLES

RECHERCHES SUR DES METHODES D'ANALYSES STATISTIQUES DESCRIPTIVES  
(L'ANALYSE DES CORRESPONDANCES ET LA CLASSIFICATION AUTOMATIQUE)

par G.E. WEIL  
A. SALEM  
M. SERFATY

A un moment de l'histoire des sciences où le chercheur contemporain voit sans cesse se renouveler les outils de la recherche et les méthodes de travail, le philologue, l'historien des textes comme le linguiste se devaient de tenter, eux aussi, dans leur domaine, la remise en question de leurs méthodes et des outils de travail dont ils disposaient.

Dans le souci d'aborder le plus objectivement possible l'analyse critique des textes anciens et d'échapper à la seule méthode conjecturale, les savants et les chercheurs contemporains ont tenté de doter l'analyse philologique de moyens nouveaux dont les critères objectifs permettraient d'orienter leurs recherches vers des résultats moins discutables que ceux qui étaient obtenus jusqu'alors. Le souci qu'ils avaient de passer d'une analyse qualitative à une analyse quantitative pouvait leur laisser espérer une meilleure approche des problèmes qu'ils rencontraient dans leurs recherches. Ceux qui ont montré ce souci de renouvellement avaient mesuré, avec le temps, le poids que pouvaient faire peser sur leur jugement face à l'analyse des textes : les critères historiques, esthétiques, quelquefois théologiques, ainsi que les critères d'école si subjectifs du fait de leur nature que même le recours aux *realia* ne pouvait les rendre objectifs. Le recours aux *realia*, lorsque ceux-ci sont seulement utilisés pour appuyer par l'analyse une théorie prescrite qu'aucun moyen scientifique ne permet dans l'instant de vérifier, ne peut rendre objectivement démonstratif ce que la con-

jecture subjective a décidé de démontrer. On peut, sans déroger, se retourner vers les excellents maîtres que furent, à la fin du XIXe siècle, en matière d'études historiques, Langlois et Seignobos et rappeler ici leur définition toujours si actuelle de l'*hypercritique* qui fut et demeure un des travers des écoles de critique textuelle les plus modernes (1).

"... C'est l'excès de critique qui aboutit, aussi bien que l'ignorance la plus grossière, à des méprises. C'est l'application des procédés de la critique à des cas qui n'en sont pas justiciables. L'hypercritique est à la critique ce que la finasserie est à la finesse. Certaines gens flairent des rébus partout, même là où il n'y en a pas. Ils subtilisent sur des textes clairs au point de les rendre douteux, sous prétexte de les purger d'altérations imaginaires. Ils distinguent des traces de truquage dans des documents authentiques. Etat d'esprit singulier ! à force de se méfier de l'instinct de crédulité, on se prend à tout soupçonner.

Il est à remarquer que plus la critique des textes et des sources réalise des progrès positifs, plus le péril d'hypercritique augmente. En effet, lorsque la critique de toutes les sources historiques aura été correctement opérée, le bon sens commandera de s'arrêter. Mais on ne se résignera pas : on raffinera, comme on raffine déjà sur les textes les mieux établis, et ceux qui raffineront tomberont fatalement dans l'hypercritique. "Le propre des études historiques et de leurs auxiliaires, les sciences philologiques, dit E. Renan, est, aussitôt qu'elles ont atteint leur perfection relative, de commencer à se démolir(2)". L'hypercritique en est la cause..."

Pour échapper à ce dilemme qu'annonçaient prophétiquement Langlois et Seignobos, et qui est une tentation constante, les chercheurs modernes ont voulu se donner des outils de travail nouveaux. Ils se sont tournés alors vers les mathématiques, ou mieux encore vers le langage mathématique, pour savoir

s'ils pourraient puiser dans leur arsenal les instruments de mesure qu'ils espéraient utiliser et qui leur permettraient enfin une approche objective de l'analyse textuelle, sans risque de surenchère. Il est très vite apparu qu'ils pourraient trouver dans les méthodes statistiques certains des instruments auxquels ils aspiraient, dans la mesure où les statisticiens eux-mêmes recherchaient, au même moment, des ensembles de plus en plus vastes à explorer. Le développement de la science informatique, l'usage de plus en plus courant des ordinateurs pour le traitement des données textuelles et la rencontre des philologues et des statisticiens devaient créer l'engouement que la linguistique moderne connaît pour la quantification et, plus encore, pour une certaine forme de la classification (3).

## II

Ce qui semblait une découverte moderne des procédés de la quantification, avec la mise en oeuvre sur des quantités représentant les aspects qualitatifs des données textuelles d'un jeu de plus en plus compliqué de formules statistiques, ne représente en fait qu'une reformulation moderne des premiers procédés d'analyse textuelle utilisés par les scribes de l'Antiquité. La longue période qui, à travers l'histoire des méthodes de critique et d'hypercritique philologique, a conduit les chercheurs modernes à reposer les problèmes de l'analyse des textes, semble leur avoir fait oublier les méthodes quantitatives et naïvement arithmétiques qu'avaient utilisés les scribes de la Grèce antique pour décompter les occurrences dans les scholies des textes d'Homère, tout autant que les scribes-grammairiens du Judaïsme et de l'Islam pour établir les textes fondamentaux de leurs Ecritures Sacrées (4), textes qu'ils avaient la charge de transmettre dans leurs formes intangibles en vertu des consensus théologiques qui les motivaient. Les commentaires quantitatifs qui accompagnent la Bible

hébraïque comme le texte du Qoran offraient à ceux qui les utilisaient des garanties certaines tant pour l'authentification des textes eux-mêmes que pour leur normalisation grammaticale à un moment où, dans l'histoire de la grammaire, les prémices de l'analyse philologique s'exprimaient d'abord à partir d'exemples comparés analogiquement avant d'être peu à peu généralisés à travers des catalogues quantifiés d'expressions normalisées. Les maîtres de la Tradition juive de Babylonie furent très vraisemblablement les initiateurs de l'apparat quantitatif du texte de la Bible lorsqu'ils prirent conscience de la nécessité d'établir une paraphrase canonique du Pentateuque en araméen, qui était devenu la langue vernaculaire du Judaïsme de l'exil babylonien. Ils établirent le texte de la paraphrase biblique à partir des catalogues quantifiés qu'ils durent constituer. En nous reportant aux fragments que nous avons retrouvés, nous constatons qu'ils avaient rapporté toutes les équivalences sémantiques et stylistiques des expressions hébraïques de base dans le rendu araméen qui allait être considéré comme canonique, et ceci dans les limites de l'exégèse traditionnelle (5).

Les scribes grammairiens juifs qui prirent à charge, indépendamment de la paraphrase, l'établissement et la transmission des textes sacrés dans leur forme hébraïque ne semblent jamais avoir soulevé aucun des problèmes concernant les sources, apparentes à la simple lecture, dont pouvaient résulter les textes qu'ils transmettaient. La place fondamentale et normative que pouvait occuper le texte biblique dans le consensus religieux juif de leur époque et qui sera la même, durant des siècles, que celui accepté dans sa fidélité au texte par l'ensemble du monde chrétien et du monde musulman, semblait empêcher toute autre interpellation du Texte qui ne fut pas conforme à son rôle d'ouvrage unique de référence (6).

Cependant, avant les heures glorieuses des grandes batailles de la critique ou de l'hypercritique textuelle allemande, les maîtres de la judaïcité la plus orthodoxe avaient ressenti, sans les exprimer en termes de critique historique, les problèmes inhérents à la nature des textes qu'ils étudiaient. Dans le cas précis du livre d'Isaïe qui fait l'objet de notre étude méthodologique, un passage du Talmud relatif à l'établissement du canon biblique, à la fin du 1er siècle, laisse entrevoir la perplexité des maîtres juifs devant la partition typologique du livre attribué au prophète Isaïe fils d'Amos (7). La discussion qui nous est relatée avec une assez grande franchise et qui porte sur la place de ce livre dans l'ordre chronologique qu'il doit occuper parmi les ouvrages prophétiques reçus dans la seconde partie du canon biblique hébreu, montre le partage des avis des canonistes juifs quant à la typologie de l'oeuvre du Prophète. D'une part, se fondant sur les épigraphes des premières prophéties, ils situent l'ouvrage comme l'oeuvre d'un des premiers prophètes du royaume de Judée, précédant le siège et la destruction de Jérusalem, et d'autre part, se fondant sur les consolations promises à Israël et l'annonce du bonheur messianique des dernières prophéties, ils laissent entrevoir le sentiment qu'ils éprouvent devant des prophéties de type exilique ou même post-exilique. De ce premier et très ancien témoignage d'analyse littéraire, tiré du Talmud, on peut dire qu'en dehors du livre de Jérémie, auquel ils attribuent la rédaction de toute son oeuvre, les canonistes juifs ont affirmé la mise en forme postérieure à l'existence des auteurs pour tous les autres livres prophétiques, dont celui d'Isaïe. Sans pour autant affirmer la pluralité d'auteurs qu'ils sous-entendent pour certaines parties de quelques ouvrages, ils attribuent le collationnement, la mise en forme et l'"édition" de toutes les oeuvres prophétiques à des cercles de savants qui ont nom *Académie d'Ezéchias* et *Grande Synagogue* (7).

Les problèmes posés aux canonistes de Palestine par l'oeuvre d'Isaïe n'avaient finalement jamais trouvé de solution. Le livre d'Isaïe occupe dans les manuscrits hébreux de la Bible l'une ou l'autre des places proposées par les premiers canonistes, selon les traditions régionales reçues par les écoles de scribes qui copiaient les manuscrits bibliques. Dès le XI<sup>e</sup> siècle, le problème de la bipartition du livre d'Isaïe est soulevé à nouveau par deux maîtres célèbres du Judaïsme de la période islamo-espagnole, Moïse b. Samuel ibn Chiquitilla de Cordoue et Abraham b. Meir ibn Ezra de Tudèle, qui n'hésitent pas, dans leurs commentaires, à opposer les deux parties de l'oeuvre, chapitres 1 à 39 et chapitres 40 à 66, quant à l'époque où doivent se réaliser les consolations ainsi que sur la finalité de l'annonce messianique pour le peuple d'Israël et pour Jérusalem. L'annonce des consolations et de l'avènement messianique font l'objet des chapitres de la seconde partie de l'oeuvre, à l'opposé des prophéties sur la Judée et sur les Nations d'alentours qui font l'objet de la première partie de l'ouvrage (8). Abraham ibn Ezra ira même jusqu'à évoquer en toutes lettres dans son commentaire à Is. 49,7 le *"mystère de la seconde moitié du livre"*, dont il a fait état dès le début de son commentaire au chapitre 40, qui constitue le premier chapitre de la seconde partie, dans un langage qu'il choisit toujours volontairement elliptique.

Naturellement, un ouvrage comme le livre d'Isaïe, présentant des contrastes typologiques aussi marqués, deviendra bientôt le terrain d'élection où s'affronteront tous les auteurs qui ont marqué l'histoire de la critique et de l'hypercritique biblique. Comme chacun de ces auteurs développe des arguments dignes d'intérêt et comme tous sont des savants sérieux et dignes de foi, la permanente surenchère des auteurs les uns par rapport aux autres conduira l'honnête bibliste à se demander où se trouve la vérité et quel sérieux peut être accordé à des méthodes d'analyse qui conduisent à des résultats si opposés.

	IsaTe	IsaTe		Proto-IsaTe	IsaTe		Trito-IsaTe	Découpages Extrêmes
	Chp. 1 à 66	Proto-IsaTe Chp. 1-39	Deutéro-IsaTe Chp. 40-66		Deutéro-IsaTe	Trito-IsaTe		
Ier Siècle	Tradition judéo-chrétienne la plus large							
XIe Siècle		Moïse b. Samuel Ibn Gikatilla						
XIe Siècle		Abraham b. Meïr Ibn Ezra						
1850	Einrich August Hahn							
1850	Rudolph Stier							
1850				I <sub>1</sub> IsaTe 1-12 I <sub>2</sub> Anonym. 12-23	Ernst Meier II <sub>1</sub> Anonym. 24-27, 34-36, 38 II <sub>2</sub> Is. 28-33,37	III Is. 40-66		
1855	S.D. Luzzatto							
1865		Franz Delitzsch IsaTe Prophète IsaTan						
1892				I <sub>1</sub> IsaTe 1-12 I <sub>2</sub> IsaTe 13-23 I <sub>3</sub> IsaTe 24-34 I <sub>4</sub> Rois 36-39	Bernhard Duhm II <sub>1</sub> Dt-Is. 40-55	III Trito-Is. 56-66		
1909							I Is. 1-36 II Is. 40-48 III Is. 40-55 IV Is. 56-57 V Is. 58-66 VI Is. 63-66	Budde
1926		Eduard König						
1928		Charles C. Torrey						
1933					Karl Elliger	III <sub>1</sub> Trito 40-55 III <sub>2</sub> Trito 56-66		
1936		Mayer Lambert (Inédit) I IsaTe II <sub>1</sub> Anonyme 40-52,13 II <sub>2</sub> Anonyme 52,13-66						
1941-43	E.J. KISSANO							
1951	O.T. Allis							

D'autres avant nous ont déjà retracé l'histoire de la critique du livre d'Isaïe (9), nous ne la reprendrons pas ici, si ce n'est pour donner en un bref tableau les découpages obtenus dans le mouvement de la méthode conjecturale par quelques-uns des principaux auteurs qui ont marqué, d'une part, les écoles défendant l'unicité d'auteur du livre d'Isaïe et, d'autre part, les savants qui ont le plus marqué les écoles allant vers une pluralité d'auteurs de plus en plus large. Sans prendre encore parti dans le concert quelque peu tonitruant qui oppose les tenants de l'une ou l'autre des écoles, notre groupe de recherche reste confondu d'admiration devant la science des savants qui peuvent user de toutes les couleurs que leur offre la palette de l'arc-en-ciel pour déclarer authentique ou inauthentique tel ou tel chapitre et pour certains tel ou tel verset ou tel ou tel élément de verset qu'ils teintent aussitôt, dans leurs éditions, aux couleurs des ensembles auxquels ils peuvent les rattacher, ce qui offre au lecteur abasourdi une oeuvre biblique caviardée aux couleurs du moment, celles-ci restant naturellement interchangeables selon les écoles, le lieu et l'instant.

Nous apporterons, avant d'aborder l'exposé des méthodes nouvelles que nous avons voulu mettre en oeuvre pour résoudre objectivement les problèmes posés, un dernier témoignage historique d'une confortable antiquité. Parmi les manuscrits bibliques découverts en 1947 à Qumran dans le désert de Judée près de la mer Morte, les deux manuscrits d'Isaïe sur cuir tanné, que l'on peut dater d'environ un siècle avant notre ère, sont venus apporter le témoignage de l'antique compilation du livre d'Isaïe dans la forme que nous lui connaissons aujourd'hui, et sur lesquels les tenants des écoles critiques et hypercritiques se sont tant acharnés. Le manuscrit A, le plus complet, et qui marque le plus grand nombre de variantes orthographiques avec le texte actuel, comme le manuscrit B, le plus proche du texte massorétique, contiennent tous les deux tous les chapitres du livre du 1er au 66ème. Aucune trace de séparation entre



le chapitre 39 et le chapitre 40 n'est apparente dans aucune des deux recensions. Dans le manuscrit A, le chapitre 40 commence à la suite du précédent sur la dernière ligne d'une même colonne. Seule, une séparation de trois lignes marque, contre toute attente, une distance entre le chapitre 33 et celui qui le précède. Une étude de l'orthographe du texte permet de penser que cette séparation fortuite est due à un incident de copie, peut-être au changement du manuscrit de base servant à la copie de l'oeuvre en cours. La conclusion la plus importante pour notre étude apportée par ce dernier témoignage textuel authentique et ancien est que, dès le IIe siècle avant notre ère, le livre d'Isaïe existait dans la forme exacte que nous lui connaissons aujourd'hui.

### III

En dépit de tout ce qui a été écrit sur le livre d'Isaïe, les problèmes demeurent posés. Il était donc particulièrement important que notre groupe de recherche tentât lui aussi sur cet ouvrage, à l'unité d'auteur si controversée, ses premières expériences en vue de définir un outil mathématique destiné à de plus larges applications en matière de critique textuelle, et ceci en dépit du fait que d'autres avant lui avaient tenté sur ce même livre leurs propres et très récentes expériences quantifiées (10).

Notre tentative se place dans une direction de recherche très précise qui ne veut utiliser comme éléments de l'analyse que les unités significatives du discours dégagées de toute signification. Seules, seront prises en compte dans notre approche de l'analyse textuelle objective, la nature grammaticale des éléments de la phrase et leur distribution contextuelle. L'utilisation des méthodes statistiques permettant de vérifier de façon très fine les inégalités de distribution, notre hypothèse de travail sera de vérifier si ces inégalités permettent de déceler,

avec un certain degré de certitude, l'unité ou la pluralité d'auteur ou, simplement, des différences typologiques.

Après avoir étudié avec beaucoup de sérieux les méthodes d'analyses quantitatives ou statistiques qui avaient été utilisées d'une manière très générale en matière d'analyse textuelle dans les dernières années, il nous est apparu que l'on pouvait légitimement se poser le problème de la valeur des méthodes utilisées. Travaillant sur des éléments essentiellement dépendants, ce qui est indiscutablement la situation dans une analyse textuelle (11), quelle valeur peut-on accorder aux résultats exprimés à partir des tests statistiques les plus couramment employés en analyse textuelle dont l'emploi dans les sciences exactes s'applique à des objets radicalement différents. En effet, l'emploi canonique de ces tests est de mesurer la relation qui existe entre l'écart observé d'un ensemble d'événements par rapport au calcul théorique de cet ensemble, considéré comme composé de parties radicalement indépendantes et distribuées au hasard. Devant la constatation d'un fait qui semble faire sortir les applications étudiées, dans le domaine des analyses textuelles, *"hors des limites de l'épure"*, il nous est apparu très important de définir d'abord les domaines où les méthodes de statistique descriptive qui étaient à notre disposition, dans l'arsenal méthodologique très large dont dispose cette discipline, pouvaient être utilisées en matière d'analyse textuelle sans sortir de leur domaine de validité.

De fait, en analyse textuelle on se trouve placé face à deux situations bien distinctes.

Dans le premier cas, on a affaire à une information faite d'un très grand nombre d'unités discrètes, et dans ce cas l'on peut penser à utiliser l'outil statistique qui est par nature destiné à l'étude des relations entre les phénomènes me-

surés sur des vastes ensembles, à condition d'avoir choisi en pleine connaissance de cause l'outil que l'on veut employer pour le but précis que l'on veut atteindre. En analyse textuelle, la recherche de l'outil et la définition du but que l'on peut s'assigner ne représentent pas les moindres difficultés de la méthode.

Dans le second cas, on se trouve devant des petits nombres et la question se pose de savoir si l'outil statistique peut encore être employé en analyse philologique sans risquer de fausser l'analyse (12). Il semble, cependant, qu'aucun outil statistique ne pourra mettre en évidence les raisons des aramaïsmes qui apparaissent dans tel ou tel livre de la Bible, ni la raison de la présence du nom de Cyrus (-VI<sup>e</sup> siècle) dans le livre d'Isaïe, ouvrage attribué à un auteur ayant vécu au VIII<sup>e</sup> siècle avant notre ère. Tout ce qui peut être discriminant dans les faits représentés par des petits nombres, nous semble relever non de la méthode statistique, mais du langage arithmétique, donc du raisonnement philologique sur lequel nous avons élevé un doute causé par la crainte de l'usage de la seule méthode conjecturale. Pour que la méthode philologique pure puisse être considérée comme une méthode scientifique démonstrative, il faudrait que soient définis, avec une grande rigueur, et demeurent inchangés durant toute l'application, les procédés ou les lois qui peuvent conduire à valider une recherche de forme hypothético-déductive. Au contraire, seul, ce qui relève de l'analyse de données ou de tableaux mettant en jeu des grands nombres peut sembler être l'objet d'une étude statistique.

Devant cette situation, et sachant bien que la science n'avance qu'à très petits pas, notre tentative a été de définir les limites des méthodes statistiques que nous allions mettre en oeuvre pour nos travaux. Déçus ou peu assurés par les résultats qui avaient été obtenus jusque-là en analyse textuelle, grâce aux méthodes d'analyses statistiques dites paramétriques, nous avons tenté, sur les

amicales instigations de notre collègue J.-P. Benzécri, d'utiliser deux procédés d'analyse des données qu'il avait élaborés, mis au point et vulgarisés grâce à ses travaux et à ceux de ses collaborateurs (13). Le choix que nous avons fait de ces méthodes était principalement fondé sur le fait que celles-ci nous permettaient d'appliquer dès l'abord un principe méthodologique que nous nous étions imposé : éviter d'introduire par quelque biais que ce fût les découpages proposés par la critique biblique, découpages dont nous voulions apprécier la validité en tant qu'hypothèse de départ. Les deux méthodes d'analyse des données, qui nous étaient proposées, permettaient de travailler sur des tableaux de contingence dans lesquels était conservé le découpage en chapitres auquel nous sommes habitués. Si la division capitulaire est due à Etienne Langton, futur archevêque de Canterbury qui, aux environs de 1205, la prescrivit sur le texte de la Vulgate latine révisée à Paris, la division capitulaire fut par la suite utilisée sans problème par les auteurs juifs dans les premières concordances hébraïques manuscrites et, enfin, introduite dans les éditions de la Bible hébraïque. Dans le cas du livre d'Isaïe, le découpage capitulaire se confond assez bien avec la division traditionnelle hébraïque et, de ce fait, il offre une partition du texte qui n'est ni plus mauvaise ni meilleure qu'une autre; cette partition a surtout l'avantage de ne conditionner en rien la présentation des éléments du texte, en conservant, en vue de l'analyse, les données dans la forme où le lecteur et le chercheur peuvent les retrouver dans le texte. Il est évident qu'un découpage automatique, obtenu sur des données présentées selon la division capitulaire, ne peut être totalement comparé à un découpage obtenu sur des données échappant à cette division. On peut, par hypothèse, poser le principe que la qualité d'une suite d'analyses statistiques bien conduites, sur des ensembles obtenus par des découpages objectivement arbitraires, permettra d'atteindre un degré de finesse suffisant pour dessiner les contours d'ensembles totalement cohérents.

Dans une première approche, en partant du comptage de mots dans les chapitres du livre, nous avons tenté de mettre en évidence les similitudes ou les grandes oppositions qui peuvent exister entre les différents éléments du découpage textuel, du point de vue de la distribution du vocabulaire, à partir des 89 mots les plus fréquents qui y sont employés (14).

Dans le cadre de cette étude, le mot est défini comme une unité graphique représentée par une suite concaténée de signes séparés des autres unités graphiques par des symboles privilégiés tels que le blanc. Ces unités graphiques ont été ramenées à leur forme radicale pour les verbes et au lemme de base pour les noms et les particules qui sont les trois éléments qui constituent le discours hébreu dans une première approximation. Chaque fois où, dans cet article, il sera question de mots ou d'occurrences de mots dans un chapitre, l'expression utilisée sera à entendre dans les limites de cette définition.

Dans une seconde approche, nous avons tenté de réaliser, par touches successives, un découpage en ensembles relativement homogènes réunissant des petits groupes de chapitres consécutifs. Le problème posé par l'attribution définitive de ces ensembles homogènes, sur le plan de la distribution du vocabulaire, à un ou à plusieurs auteurs distincts, ne pourra être tranché de manière définitive dans le cadre de cette étude préliminaire. La seule démarche possible pour que le statisticien puisse contribuer à résoudre ce problème, réside dans une étude parallèle, à l'aide des mêmes méthodes d'analyse, du degré d'hétérogénéité compatible avec une unicité d'auteur mesurable sur des oeuvres mieux cernées, par ailleurs, qui pourraient même être extérieures à la littérature biblique.

## IV

### METHODOLOGIE DE L'ANALYSE FACTORIELLE DES CORRESPONDANCES

#### 1. LE TABLEAU DES DONNEES

La démarche suivie a consisté à analyser un tableau  $P(I, J)$  à deux entrées réalisé à partir du comptage des occurrences des 89 mots les plus fréquents dans les 66 chapitres d'Isaïe. A l'intersection de la colonne  $J$  et de la ligne  $I$  de ce tableau, on trouvera le nombre de fois où le mot numéro  $i$  apparaît dans le chapitre  $j$ .

Les chapitres d'Isaïe, qui sont au nombre de 66, sont de longueurs assez variables. Le chapitre 37, qui est le plus long chapitre, contient 566 mots; le chapitre 12, qui est le plus court, en contient 62. La majorité des chapitres oscille entre 150 et 400 mots, ce qui donne un rapport d'environ un à trois. D'autres partitions de l'oeuvre auraient été tout aussi possibles pour notre analyse : tranches de  $n \times 100$  mots, tranches de  $n \times 100$  versets, etc. La partition en chapitres présentait l'avantage décisif que les comptages de mots étaient plus faciles à réaliser à partir de données déjà saisies sur un support magnétique.

Nous avons retenu les mots *les plus fréquents* dans l'ensemble du texte pour constituer les lignes de notre tableau, les 66 colonnes représentant les 66 chapitres où se répartissent les occurrences de ces mots. Le choix de l'ensemble des *mots les plus fréquents* comme élément de l'analyse est une démarche courante en analyse des données qui peut être exprimée dans le domaine qui nous préoccupe de la manière suivante :

Lorsque l'on compare des textes ou des parties de texte en vue de rechercher s'ils ont pu être écrits par le même auteur, le fait qu'un mot rare apparaisse ou n'apparaisse pas dans l'un ou l'autre des textes ou partie de texte est moins important qu'une différence marquée de la fréquence d'emploi d'un mot très abondant dans ces textes ou parties de texte.

Nous tenterons, au cours de ces analyses, de faire apparaître des différences assez appréciables dans le tissu du discours pour permettre de mettre en évidence des ensembles homogènes en opposition les uns par rapport aux autres. C'est pourquoi il a été jugé utile d'écarter de l'analyse, dans un premier temps, les 11 noms propres qui apparaissaient dans la liste des 100 mots les plus fréquemment employés dans le livre d'Isaïe que nous avons obtenue par un comptage automatique. Il en résulte un tableau de 66 chapitres x 89 mots; ce tableau contient 6941 mots ou unités graphiques sur les 16931 que contient le livre d'Isaïe dans sa totalité. Ceci correspond à un *pourcentage de lecture global* de 6941 : 16931 soit 41 %. Naturellement, ce pourcentage de lecture varie suivant les différents chapitres, allant d'un maximum de 69 % pour le chapitre 39 pour tomber à un minimum de 28 % pour le chapitre 16, mais la grande majorité des chapitres oscille assez régulièrement entre 30 et 50 % de lecture globale.

## 2. L'ANALYSE FACTORIELLE DES CORRESPONDANCES DU TABLEAU DE 89 X 66

Les dimensions mêmes d'un tableau comprenant 5874 cases, comme celui que nous présentons, sont un obstacle à l'assimilation rapide de l'information qu'il contient. Le but de l'analyse factorielle est de représenter, dans la mesure du possible, ces vastes ensembles de données dans un espace de très faibles dimensions et ce, en perdant le moins d'information possible.

Nous allons tenter d'illustrer par un exemple, que nous choisissons le plus simple possible, la démarche de la méthode que nous avons utilisée. Le mathématicien averti trouvera l'exposé complet de la méthode dans *l'Analyse des Données*, rédigé à son intention par J.-P. Benzécri, un exposé plus complet, particulièrement destiné aux recherches lexicologiques pour les littéraires et les historiens, a été donné par A. Prost dans son récent ouvrage (15) vers lequel pourra se retourner le chercheur que notre exposé simplifié n'aura pas entièrement satisfait.

Supposons pour notre exemple que nous ne nous intéressions qu'à l'emploi des deux mots  $\gamma\eta\lambda\lambda$  - %NY - et  $\Upsilon\eta\lambda\lambda$  - %R3 - dans les 66 chapitres du livre d'Isaïe. Nous devons compléter encore notre hypothèse en supposant que tous les chapitres sont de "*poids égal*", c'est-à-dire qu'ils ont sensiblement le même nombre de mots, pour que notre exemple soit plus simple à suivre (16). La représentation de la distribution de ces deux mots peut alors être portée pour chaque chapitre, comme dans les figures ci-après, sur un système de deux axes perpendiculaires avec une abscisse égale au nombre de fois où le mot %R3 apparaît dans ce même chapitre. Si les points se répartissaient uniformément dans tout le quadrant, comme dans la figure 1, la connaissance du nombre des occurrences d'un des deux termes n'apporterait que peu d'information sur le nombre des occurrences de l'autre terme. Si, au contraire, comme sur la figure 2, les points se répartissaient le long d'une droite, nous dirions que les variables sont liées linéairement entre elles et que la connaissance du nombre d'occurrences du mot %NY nous renseigne sur le nombre d'occurrences du mot %R3.



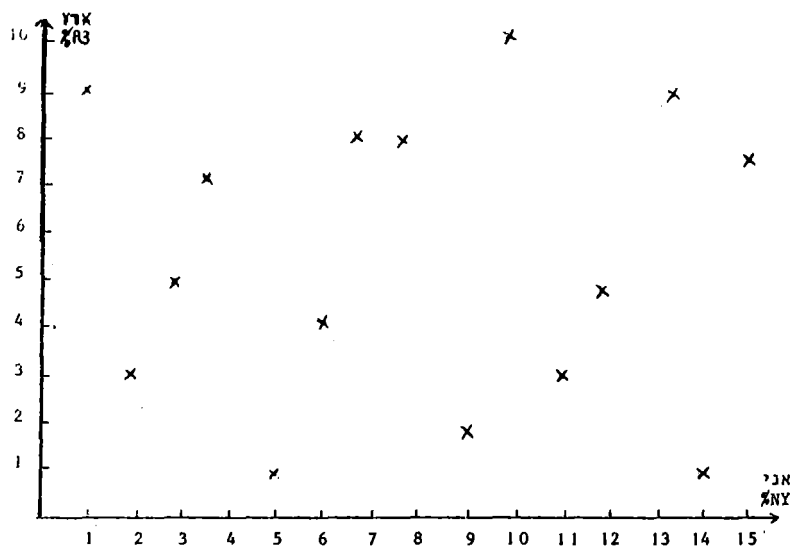


Fig. 1

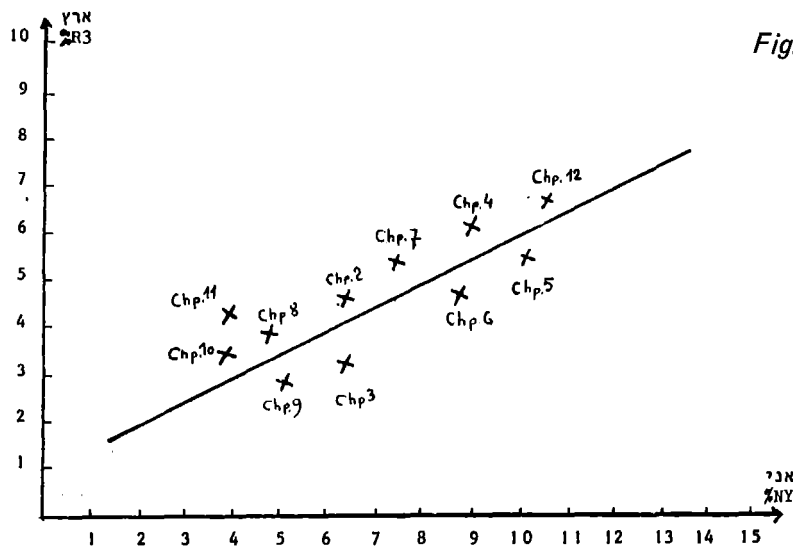
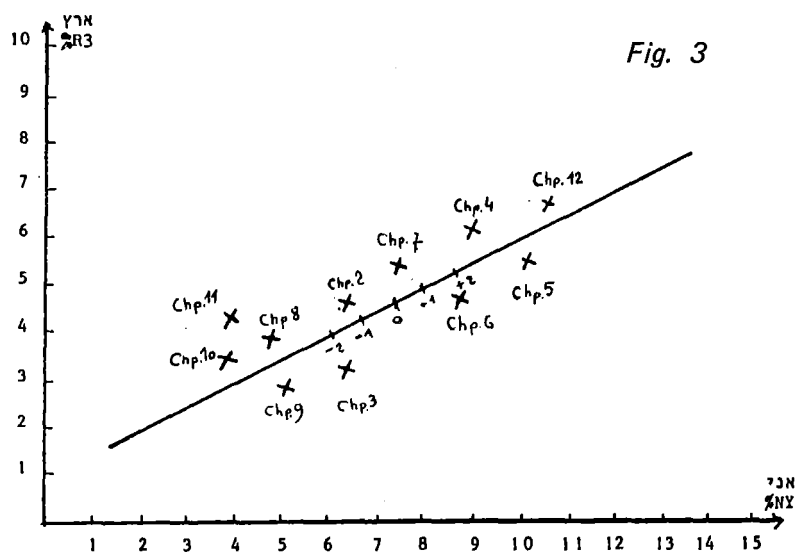


Fig. 2

Si nous traçons une droite qui passe *le plus possible par le milieu du nuage*, nous pourrions abaisser une perpendiculaire de chacun des points sur la droite comme dans la figure 3. La longueur du segment qui va du chapitre  $X_i$  à la droite est la distance du chapitre à cette droite  $\Delta X_i$  que nous venons de tracer. Il existe une droite qui minimise la quantité :

$$\sum_i (\Delta X_i)^2$$

représentant la somme des carrés des distances des différents chapitres à la droite; cette droite se nomme la droite des moindres carrés et c'est elle que nous avons tracée dans la figure 3.



Nous pouvons maintenant construire un nouveau système de coordonnées en prenant pour abscisse la droite  $aa'$ , pour ordonnée une droite qui lui est per-

pendiculaire et qui passe par le *centre du nuage*. Nous nommerons ces nouvelles directions respectivement le *facteur 1* et le *facteur 2*, nous pourrons alors, comme sur la figure 4, marquer la disposition de nos chapitres dans ce nouveau système de coordonnées.

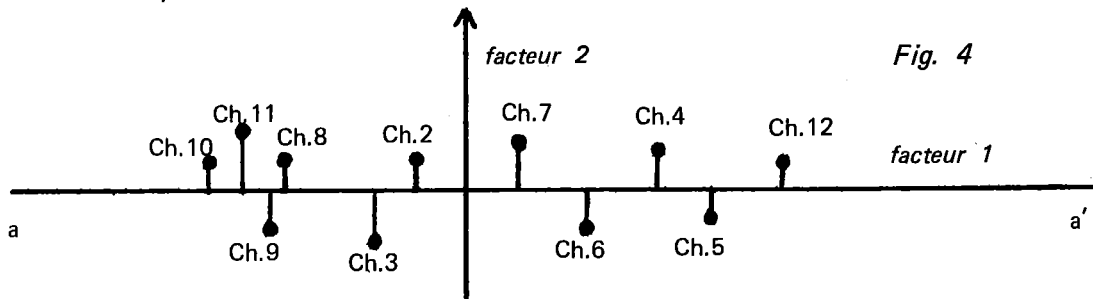


Fig. 4

On doit noter, qu'à ce stade, nous avons conservé toutes les informations sur la disposition relative des chapitres les uns par rapport aux autres.

Nous allons maintenant tenter de ne conserver qu'une seule des deux coordonnées et, ce faisant, tenter de projeter les points formant le nuage sur cette droite unique de manière à déformer le moins possible les distances entre les chapitres. Considérant la figure 4, on voit que la droite qui forme le *facteur 1* se prête relativement bien à cette opération. Le résultat obtenu est donné par la figure 5.

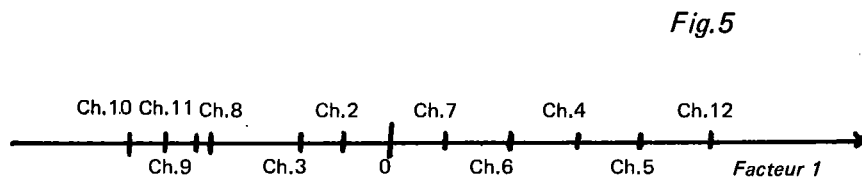


Fig.5

Plus les points sur le graphique sont proches de la droite représentant le *facteur 1*, moins importante est la perte d'information sur la position des chapitres les uns par rapport aux autres.

De fait, nous avons opéré ce qu'on nomme en statistique une régression (17). Les diverses méthodes d'analyse factorielle sont des généralisations de ce type de démarche aux espaces de plus grandes dimensions. Dans le cas qui nous intéresse, nous tenterons de résumer les positions relatives des 66 chapitres du livre d'Isaïe les uns par rapport aux autres, non plus d'après le nombre d'occurrences des deux mots -  $\Upsilon_{\text{R3}}$  - et -  $\Upsilon_{\text{NY}}$  -, comme dans notre exemple, mais en fonction des occurrences des 89 mots les plus fréquents dans chacun des 66 chapitres. Ces décomptes, sommes et distributions par chapitres, ont été obtenus automatiquement par l'ordinateur. Pour caractériser les positions relatives des 66 chapitres du livre, nous devons extraire les 2 ou 3 premiers facteurs qui permettent de visualiser de manière approchée les distances entre les chapitres.

Il va de soi que la généralisation de ces méthodes d'analyse aux espaces de plus grandes dimensions nécessite un outil mathématique beaucoup plus évolué que celui dont nous nous sommes servis dans notre exemple. L'emploi de cet outil requiert surtout un moyen de calcul électronique assez puissant pour pouvoir effectuer en quelques secondes des calculs extrêmement volumineux.

Parmi les différentes méthodes d'analyse, l'*analyse factorielle des correspondances* se distingue par les propriétés suivantes :

a. Elle permet d'établir une relation de proximité entre les *lignes-mots* et les *chapitres-colonnes* du tableau ainsi créé.

b. La distance employée pour mesurer le degré de disparité entre les colonnes, et respectivement les lignes, du tableau *tend à neutraliser* les disparités dues aux différences de tailles des textes ou au nombre d'occurrences des mots. Plus précisément, on dira qu'elle satisfait au principe d'équivalence distributionnelle. Les résultats de l'analyse ne changeraient pas si deux colonnes ayant *même profil*, c'est-à-dire si deux chapitres contenant les mêmes mots dans les mêmes proportions, étaient cumulés en une seule et même colonne. De la même façon, si deux lignes représentant deux mots répartis proportionnellement sur l'ensemble des chapitres étaient cumulées en une seule et même ligne, les résultats de l'analyse dans cette dernière hypothèse n'en seraient pas plus affectés.

En fonction de cette distance, deux chapitres seront d'autant plus proches qu'ils utilisent les mêmes mots dans les mêmes proportions. De même, deux mots seront d'autant plus proches qu'ils seront abondants ou absents dans les mêmes chapitres.

Les programmes d'*analyse factorielle des correspondances* fournissent toute une série de documents, édités automatiquement par l'imprimante de l'ordinateur, qui permettent de juger de la qualité de la représentation simplifiée que l'on obtient en projetant l'ensemble des représentations des données obtenues par l'analyse sur les premiers axes factoriels (18).

LES VALEURS PROPRES VAL(1)=0.9999577

NUM	ITER	VAL PROPRE	POURCENT	CUMUL	HISTOGRAMME DES VALEURS PROPRES DE LA MATRICE 89 x 66
2	0	0.14578593	9.872	9.872	*****
3	1	0.1130427	7.055	17.527	*****
4	0	0.06710625	4.544	22.071	*****
5	2	0.06599170	4.469	26.539	*****
6	1	0.05692331	3.855	30.394	*****
7	1	0.05375019	3.640	34.034	*****
8	2	0.05315527	3.599	37.633	*****
9	1	0.04971958	3.367	41.000	*****
10	2	0.04811621	3.258	44.258	*****
11	1	0.04343903	2.942	47.200	*****
12	1	0.04059688	2.749	49.949	*****
13	2	0.03971664	2.689	52.638	*****
14	1	0.03769159	2.553	55.191	*****
15	2	0.03684548	2.455	57.646	*****
16	2	0.03511297	2.378	60.064	*****
17	1	0.03236885	2.192	62.255	*****
18	1	0.03129366	2.119	64.375	*****
19	2	0.03059745	2.072	66.446	*****
20	1	0.02831811	1.918	68.364	*****
21	1	0.02667882	1.807	70.171	*****
22	1	0.02631408	1.782	71.952	*****
23	2	0.02517078	1.704	73.657	*****
24	2	0.02458383	1.665	75.322	*****
25	1	0.0232184	1.512	76.833	*****
26	2	0.02142255	1.451	78.284	*****
27	1	0.02017459	1.366	79.650	*****
28	1	0.01937541	1.312	80.962	*****
29	1	0.01849308	1.252	82.214	*****
30	2	0.01765647	1.176	83.410	*****
31	2	0.01730071	1.172	84.581	*****
32	1	0.01571578	1.064	85.644	*****
33	1	0.01517132	1.027	86.673	*****
34	2	0.01435054	0.972	87.645	*****
35	1	0.01313949	0.890	88.534	*****
36	2	0.01292032	0.875	89.409	*****
37	1	0.01155962	0.810	90.219	*****
38	2	0.01176336	0.797	91.016	*****
39	2	0.01129697	0.765	91.781	*****
40	3	0.01108157	0.750	92.531	*****
41	1	0.00965707	0.654	93.185	*****
42	2	0.00928867	0.629	93.814	*****
43	2	0.00858423	0.581	94.395	*****
44	1	0.00839763	0.569	94.964	*****
45	2	0.00753807	0.538	95.501	*****
46	2	0.00738590	0.500	96.001	*****
47	2	0.00685054	0.464	96.465	*****
48	2	0.00652750	0.442	96.907	*****
49	2	0.00563148	0.381	97.285	*****
50	2	0.00508431	0.344	97.633	*****
51	3	0.00453135	0.307	97.940	*****
52	2	0.00427850	0.290	98.230	*****
53	2	0.00376113	0.255	98.484	*****
54	3	0.00229408	0.223	98.707	*****
55	1	0.00277682	0.188	98.895	*****
56	5	0.00275109	0.186	99.082	*****
57	3	0.00247350	0.167	99.249	*****
58	2	0.00225233	0.153	99.402	*****
59	2	0.00150018	0.129	99.530	*****
60	3	0.00146548	0.099	99.629	*****
61	1	0.00141261	0.096	99.725	*****
62	2	0.00115838	0.081	99.806	*****
63	2	0.00100820	0.068	99.875	*****
64	2	0.00086406	0.055	99.929	*****
65	2	0.00065201	0.044	99.973	*****
66	2	0.00040887	0.028	100.001	*****

LA VARIATION DU TEMPS= 325

Fig. 6

### 3. L'HISTOGRAMME DES VALEURS PROPRES DE LA MATRICE

Si nous nous trouvions dans la situation d'hétérogénéité totale des lignes et des colonnes d'un tableau, le nuage des 66 chapitres et le nuage des 89 mots de notre tableau seraient tous deux contenus dans un espace à 65 dimensions. Dans le cas de l'hétérogénéité totale entre tous les éléments de l'ensemble, ces nuages de points prendraient, dans cet espace, une forme sphérique, c'est-à-dire que chacun des 65 axes porterait à peu près  $1/65$  ou 1,5 % de l'inertie du nuage. On appellera aléa la partie qui reviendrait à chacun des axes si le nuage était sphérique et qui, dans le cas de notre application, serait de 1,5 %.

L'histogramme des valeurs propres permet d'apprécier dans quelle mesure les données observées s'écartent de la situation d'hétérogénéité totale. Cette suite de valeur ne fournit pas d'indications impératives sur le nombre de facteurs à interpréter. La pratique montre cependant que lorsqu'un intervalle relatif important sépare deux facteurs consécutifs, on peut s'attendre à ce que le facteur ayant la valeur propre la plus élevée soit porteur d'une information qui pourra être interprétée par le spécialiste (fig. 6).

L'histogramme des valeurs propres de notre analyse fait apparaître deux facteurs particulièrement importants :

- le premier axe porte à lui tout seul 9,9 % d'inertie soit 6,5 fois l'aléa,
- le deuxième axe porte encore 7,7 % d'inertie soit 5,1 fois l'aléa,

alors que les pourcentages d'inertie portés par les axes suivants décroissent régulièrement de 4,5 % à zéro.

#### 4. ANALYSE DES DEUX PREMIERS AXES FACTORIELS

Le programme d'analyse factorielle des correspondances fournit, sous forme de tableaux (figures 7 et 8), pour chaque élément entrant dans l'analyse et de manière standard pour les cinq premiers facteurs, trois valeurs que nous allons tenter d'expliquer :

- $\#F$  est la valeur du facteur pour l'élément, *chapitre* ou *mot* dans le cadre de cette application (19).
- *COR* permet d'apprécier en quoi l'élément est bien représenté par l'axe.
- *CTR* est la part qui revient sur 1000 à cet élément dans la formation de l'axe, c'est la contribution relative de l'élément à l'axe.

Une lecture rapide des valeurs du premier facteur pour l'ensemble des chapitres permet de se rendre compte que la quasi-totalité des 39 premiers chapitres du livre a un premier facteur négatif (20). De même, on constate que les 27 derniers chapitres se trouvent presque tous du côté positif de l'axe (21). Peut-on alors mettre en évidence les chapitres et les mots qui contribuent le plus à créer cet axe ? Pour répondre à cette première question on extraira du tableau les éléments dont la contribution *CTR* est la plus forte.

Dans l'espace des *chapitres*, nous nous fixerons par exemple une limite inférieure de  $2 \times \frac{1000}{66}$  qui équivaut à deux fois la contribution moyenne, soit 30.

Dans l'espace des *mots*, nous nous fixerons une limite semblable : deux fois la contribution moyenne =  $2 \times \frac{1000}{89}$  soit 22.



J1	Q1T	POID	INR	1#F	COR	CTR	2#F	COR	CTR	3#F	COR	CTR	4#F	COR	CTR	5#F	COR	CTR			
11	11	226	18	101	52	3	01	-188	45	61	230	67	151	-49	3	11	292	108	281		
21	21	342	16	201	-448	112	221	-198	22	61	-261	38	171	-85	4	21	-546	166	851		
31	31	263	12	191	-23	0	01	-456	89	221	-244	26	111	92	4	21	-580	144	711		
41	41	283	5	101	-453	68	71	-663	146	201	-275	25	61	-272	25	61	239	19	51		
51	51	262	19	131	-19	0	01	-319	99	171	207	41	121	-328	105	311	131	17	61		
61	61	89	14	151	-83	5	11	265	45	91	237	36	121	30	1	01	51	2	11		
71	71	368	20	151	-543	277	421	180	30	61	-97	9	31	-215	44	141	-94	8	31		
81	81	377	19	101	-348	168	161	301	125	161	-8	0	01	245	83	181	28	1	01		
91	91	168	17	131	-128	14	21	-232	46	81	214	39	111	75	5	11	274	64	221		
101	101	291	22	141	-228	54	81	-438	200	371	-123	18	51	-64	4	11	129	17	61		
111	111	385	11	201	-345	43	91	-483	85	221	-555	112	501	-103	4	21	624	141	741		
121	121	117	4	91	-34	0	01	-324	34	41	-219	16	31	-451	67	141	-1	0	01		
131	131	93	13	141	-285	51	71	-173	19	41	-151	14	51	-5	0	01	-121	9	31		
141	141	399	23	131	-409	201	271	75	7	11	-188	42	121	328	128	371	-133	21	71		
151	151	196	6	131	-140	6	11	-415	54	91	376	44	121	212	14	41	-500	78	251		
161	161	121	8	81	-231	36	31	-221	34	31	231	36	61	88	5	11	120	10	21		
171	171	458	9	171	-419	63	111	-466	78	171	-255	23	91	-964	332	1261	78	2	11		
181	181	258	7	131	-413	66	91	-134	7	11	-538	113	321	391	59	171	180	13	41		
191	191	550	16	181	-565	200	371	-198	23	61	-305	54	221	-660	254	1051	182	19	91		
201	201	175	4	101	-601	98	101	451	55	81	145	6	11	-199	11	31	-140	5	11		
211	211	127	9	101	-20	0	01	390	91	121	-172	18	41	168	17	41	41	1	01		
221	221	444	17	151	-435	153	231	-123	12	21	-201	33	111	-446	161	521	-325	85	321		
231	231	357	12	121	-295	57	71	-266	46	71	-405	106	291	86	5	11	470	143	461		
241	241	396	13	171	-489	122	211	-337	58	131	-426	92	341	423	91	341	256	33	151		
251	251	209	8	131	-367	61	81	-449	91	151	-174	14	41	135	8	21	-278	35	111		
261	261	198	12	181	102	5	11	-543	129	311	-91	4	11	356	55	231	103	5	21		
271	271	307	10	141	-173	14	21	-414	81	151	-458	99	301	-487	112	341	42	1	01		
281	281	160	19	141	-164	25	31	-148	20	41	221	45	141	20	0	01	276	70	251		
291	291	246	20	111	-179	40	41	-44	3	01	329	135	331	-204	52	131	112	16	41		
301	301	241	28	131	-167	40	51	-205	59	101	318	141	411	-23	1	01	14	0	01		
311	311	116	10	91	-312	68	61	-103	8	11	-54	2	01	-232	38	81	-17	0	01		
321	321	256	11	161	61	2	01	-506	117	251	623	176	621	0	0	01	-55	1	11		
331	331	184	12	131	246	38	51	-240	36	61	310	60	171	199	25	71	200	25	81		
341	341	170	11	161	-16	0	01	-473	106	211	-128	8	31	344	56	191	-31	0	01		
351	351	92	6	151	319	28	41	-244	16	31	-19	0	01	80	2	11	-411	46	181		
361	361	755	28	251	-515	195	501	830	506	1681	90	6	31	232	40	221	104	8	51		
371	371	693	39	211	-335	141	301	657	540	1501	78	8	41	46	3	11	26	1	01		
381	381	414	17	111	-326	118	131	349	134	191	249	68	161	-208	48	111	-203	46	131		
391	391	526	13	171	-625	195	341	674	226	511	369	68	261	-205	21	81	-177	16	71		
401	401	198	21	131	384	160	211	-32	1	01	28	1	01	178	34	101	46	2	11		
411	411	642	21	251	862	427	1081	355	72	241	-300	52	291	-343	68	381	200	23	151		
421	421	325	19	161	576	265	431	14	0	01	-112	10	41	135	15	51	208	35	141		
431	431	541	21	201	705	353	721	274	53	141	-416	124	551	-40	1	11	-119	10	51		
441	441	221	22	161	413	153	251	334	100	211	-16	0	01	-245	54	201	-122	14	61		
451	451	419	25	201	606	310	631	203	35	91	-233	46	201	-151	20	91	96	8	41		
461	461	418	10	231	902	244	561	190	11	31	-216	14	71	-338	34	181	-619	115	681		
471	471	75	13	161	307	52	91	-30	1	01	144	12	41	-108	7	21	76	3	11		
481	481	460	20	231	780	368	851	261	41	121	-94	5	31	-272	45	231	-47	1	11		
491	491	257	23	121	78	8	11	326	141	211	-189	48	121	211	59	151	29	1	01		
501	501	254	10	131	451	106	141	165	14	21	-57	2	11	113	7	21	489	125	441		
511	511	280	19	151	51	2	01	57	8	21	-232	46	151	333	54	321	392	130	511		
521	521	94	14	131	81	5	11	-85	6	11	111	9	31	140	15	41	-275	55	201		
531	531	106	9	181	294	31	61	-184	12	31	297	32	121	229	19	71	183	12	61		
541	541	44	13	121	135	13	21	25	0	01	99	7	21	152	17	51	96	7	21		
551	551	143	14	151	161	16	21	-12	0	01	156	15	51	311	61	201	-285	51	201		
561	561	150	12	101	-2	0	01	117	12	11	92	7	21	-130	15	31	-369	116	291		
571	571	151	17	191	409	101	191	-135	11	31	85	4	21	229	32	131	-72	3	21		
581	581	247	12	181	156	11	21	-283	37	91	533	132	531	-240	27	111	-293	40	191		
591	591	494	16	211	450	102	221	-349	62	171	739	275	1301	-108	6	31	310	49	271		
601	601	293	18	181	-47	2	01	-134	12	31	-315	68	271	385	102	401	-358	109	501		
611	611	418	8	201	444	55	111	-661	121	321	-89	2	11	686	130	591	-629	110	571		
621	621	160	13	111	55	3	01	-169	23	31	9	0	01	410	131	331	63	3	11		
631	631	47	12	101	147	19	21	-166	24	31	-70	4	11	22	0	01	-14	0	01		
641	641	226	9	161	250	23	41	-539	109	221	413	64	221	-30	0	01	-282	30	121		
651	651	33	24	171	44	2	01	53	3	11	166	26	101	10	0	01	-47	2	11		
661	661	113	26	141	-3	0	01	273	93	171	126	20	61	-15	0	01	-16	0	01		
				6541.0	1000	1000				1000				1000				1000			

Fig. 7 LES 66 CHAPITRES



#### 4.1. AXE N° 1

4.1.1. Du côté des valeurs positives du premier axe (22), les *chapitres* qui répondent à cette condition sont classés par contributions décroissantes CTR :

<i>CHAP</i>	<i>I#F</i>	<i>COR</i>	<i>CTR</i>
<i>J1</i>			
41	862	427	108
48	780	368	85
43	705	353	73
45	606	310	63
46	902	244	56
42	576	245	43
			<hr/>
			428

Comme on peut le constater par CTR, ces six chapitres font à eux seuls 42,8 % de la contribution totale. Par ailleurs, on vérifiera par COR qu'ils sont très bien représentés sur cet axe.

4.1.2. Du côté des valeurs positives du premier axe, les mots qui contribuent le plus à la formation de l'axe sont :

<i>HEBREU</i>	<i>MOTS</i>	<i>I #F</i>	<i>COR</i>	<i>CTR</i>
	<i>II</i>			
ןןן	% NY	1160	497	108
ןן	% P	1597	371	76
ןןן	% YN	736	303	49
ןןן	NGD	1140	325	41
ןןן	YDJ	580	206	25
ןן	L %	239	197	25

4.1.3. Du côté des valeurs négatives du premier axe, les *chapters* les plus contributifs sont :

<i>CHAP</i>	<i>I #F</i>	<i>COR</i>	<i>CTR</i>
<i>J1</i>			
36	- 515	195	50
7	- 543	277	42
39	- 625	185	34
19	- 585	200	37
37	- 335	141	30
38	- 326	118	13

On voit que les contributions de ces *chapters* sont beaucoup moins fortes que celles des *chapters* qui contribuent à former le côté positif de l'axe. Le *chapter* 38 qui est cependant assez proche des *chapters* 36, 37 et 39 ne peut pas être pris en compte parce que sa contribution est nettement inférieure au seuil de 30 que nous avons fixé.

4.1.4. Du côté des valeurs négatives du premier axe, les *mots* les plus contributifs sont :

<i>HEBREU</i>	<i>MOTS</i>	<i>1 #F</i>	<i>COR</i>	<i>CTR</i>
	<i>II</i>			
ל	% L <sub>1</sub>	- 489	236	43
מלך	MLK	- 721	187	41
על	JL	- 344	243	40
ההוא	HHV%	- 893	185	36
ת	% T	- 329	211	30
יו	YVM	- 513	194	32

#### 4.2. AXE N° 2

4.2.1. Du côté positif trois *chapitres* contribuent très fortement à la formation du second axe; comme ces chapitres étaient déjà assez bien représentés sur le premier axe, on a noté sous l'appellation COR1 + COR2 leur représentation sur le plan des axes 1 et 2.

<i>CHAP</i>	<i>2 #F</i>	<i>COR</i>	<i>CTR</i>	<i>COR1 + COR2</i>
<i>J1</i>				
36	830	506	168	701
37	657	540	150	681
39	674	226	51	421
38	349	134	19	275

- Le chapitre 38 qui est assez proche, sur le plan des axes 1 et 2, des chapitres 36, 37 et 39 a cependant une contribution nettement plus faible : 19 qui est en dessous du seuil de 30 que nous avons fixé;
- on voit que les chapitres 36, 37 et 39 sont très bien représentés sur le plan des axes 1 et 2;
- on voit aussi, qu'à eux seuls les trois chapitres 36, 37 et 39 contribuent pour 36,9 % à la formation du second axe.

4.2.2. Les *mots* qui contribuent le plus à la formation du second axe du côté positif sont :

HEBREU	MOTS	2 *F	COR	CTR	COR1 + COR2
	<i>II</i>				
מֶלֶךְ	MLK	1056	400	114	587
אֶמֶר	%MR	583	595	104	604
אֶל	%L <sub>1</sub>	627	386	92	622
אֶת	%T	349	237	44	448
בֶּה	KH	695	359	32	359

Ces cinq vocables apportent une contribution au second axe de 38,6 %; on constate dans la dernière colonne qu'ils sont tous très bien représentés par leur projection dans le plan des axes 1 et 2.

4.2.3. Pour les *chapitres*, du côté négatif du second axe, les contributions les plus fortes sont beaucoup plus faibles que sur le premier axe :

<i>CHAP</i>	<i>2 #F</i>	<i>COR</i>	<i>CTR</i>
<i>J1</i>			
61	- 661	121	32
26	- 543	129	31
10	- 438	200	37

4.2.4. Du côté négatif du second axe, les *mots* les plus contributifs sont :

<i>HEBREU</i>	<i>MOTS</i>	<i>2 #F</i>	<i>COR</i>	<i>CTR</i>	<i>COR1 + COR2</i>
	<i>II</i>				
הָיָה	HYH	- 409	236	45	397
הָהוּא	HHV%	- 755	132	33	337
עַם	JM	- 401	175	27	208
עוֹלָם	JVLM	- 674	123	27	193
רוּחַ	RVX	- 602	135	24	140

Les deux premiers mots, que l'on peut déjà remarquer par leur contribution au côté négatif du premier axe, sont très bien représentés dans le plan des axes 1 et 2.

AXE HORIZONTALE 21--AXE VERTICALE 11--TITRE: 66 CHAPITRES DANS ISATE  
 LARGUEUR\* 1.81116 HAUTEUR\* 2.49058 -NOMBRE DE POINTS\* 99

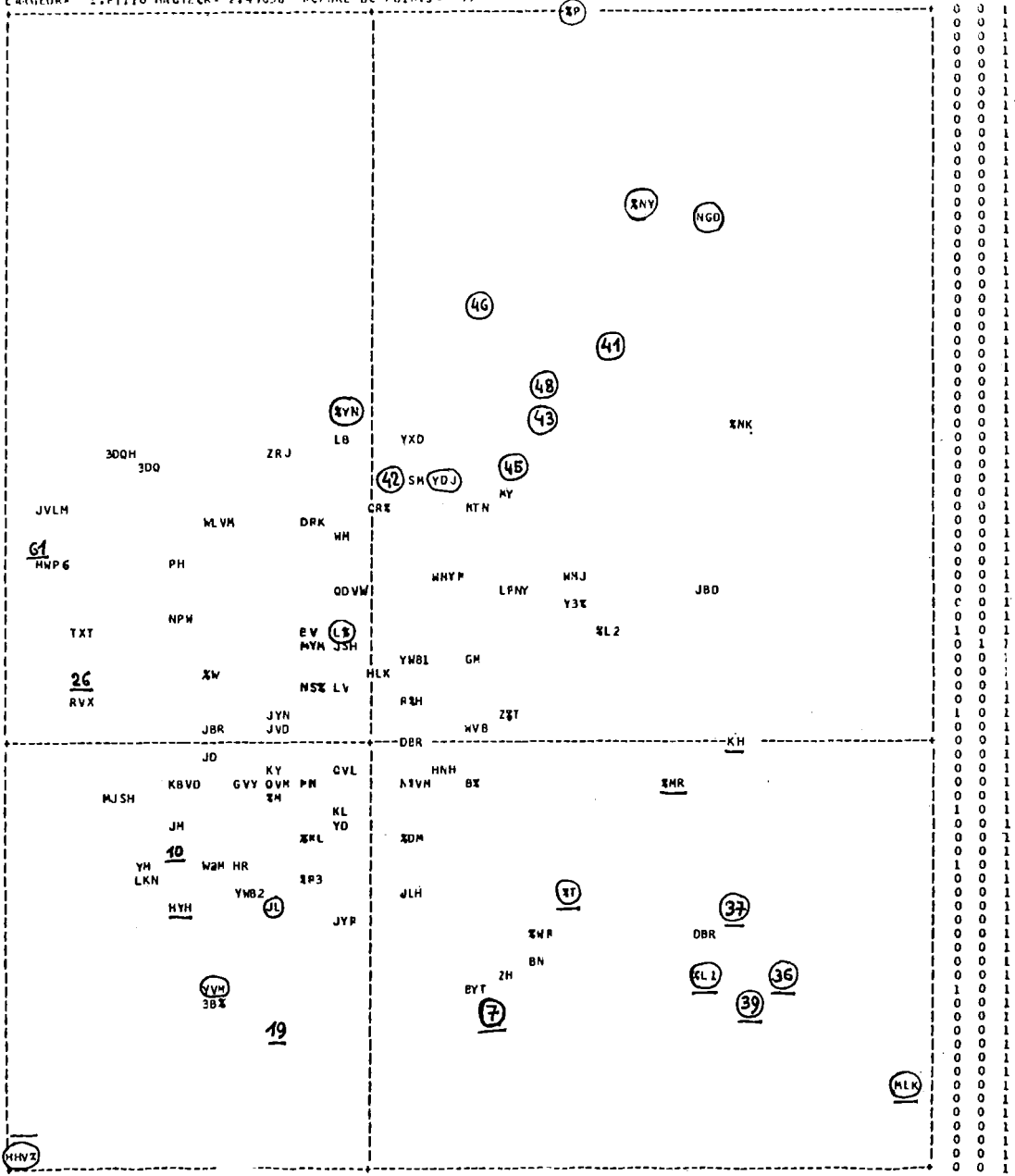


Fig. 9 LE PLAN DES AXES FACTORIELS



## 5. INTERPRETATION DES RESULTATS

Sur la figure 9 qui représente la projection des 89 *mots* dans le plan des axes factoriels 1 et 2, nous n'avons porté, pour ne pas alourdir le graphique, que les chapitres qui contribuent le plus à la formation des axes. Nous avons entouré d'un cercle les éléments qui contribuent le plus à l'axe n° 1 et souligné d'un tiret ceux qui contribuent le plus à l'axe n° 2. Le lecteur peut constater de lui-même que, dans certains cas, les mêmes éléments peuvent contribuer à la constitution de l'un et de l'autre axe. Il en va de même pour la figure 10 qui représente la projection des *chapitres*(23) sur le plan des axes 1 et 2.

Sur le graphique, certains mots occupent des positions extrêmes. On peut voir dans la position occupée par ces mots dans la représentation graphique le signe d'une répartition très irrégulière, voire quasi exclusive dans les chapitres les plus proches, représenté suivant la même projection.

C'est ce que nous allons vérifier en consultant le tableau des données de 89 x 66 :

- *MLK*, מלך (en bas à droite sur le graphique de la projection figure 9) sur 80 occurrences du mot dans tout le livre, 37 soit près de la moitié de l'effectif total du livre se trouvent dans les chapitres 36 à 39.
- *HHV%*, הנהיג (en bas à gauche sur le graphique de la projection, figure 9) sur 45 occurrences au total, 1 seule occurrence apparaît après le chapitre 32, dans le chapitre 52.
- *%P*, פא (en haut de la figure 9) sur 30 occurrences, 25 figurent entre les chapitres 40 et 48, les cinq autres se trouvent dans les chapitres 33 (1 fois), 35 (1 fois) et 26 (3 fois).

On doit voir dans les mots à répartition très irrégulière, comme ceux que nous venons d'examiner, plutôt une illustration de la typologie réalisée sur l'ensemble des mots que les raisons profondes de cette typologie. Pour s'en convaincre, on pourra vérifier qu'ils ne forment pas les axes à eux tout seuls, en consultant les tableaux de contributions relatives aux axes 1 et 2. Il est évident qu'au fur et à mesure où l'on examine les mots qui se trouvent proches du *centre du nuage*, on retrouvera ceux qui constituent le vocabulaire commun à l'ensemble des chapitres et qui est à peu près également réparti.

On retournera au texte pour voir si le groupe de mots qui caractérise les chapitres 41, 42, 43, etc., et qui sont : פֶּלֶם - %P -, יִבְנֵי - %NY -, דַּגָּד - NGD -, יִדְּי - YDJ -, יִבְנֵי - %YN -, peut donner lieu à une interprétation plus poussée. De même pour le groupe לֶלֶם - %L1 -, תֵּת - %T -, מִרְמֵר - %MR - qui caractérise les chapitres 36 à 39.

Quelles conclusions peut-on tirer quant aux chapitres ?

Considérons les chapitres dans le plan des axes 1 et 2 (figure 10).

Nous pouvons dire globalement que cette représentation fait apparaître, à quelques exceptions près, une proximité assez nette des chapitres consécutifs. Ceci est particulièrement vrai pour les groupes qui contribuent le plus à la formation des axes. Sur l'extrémité positive du premier axe, on peut noter la présence du groupe de chapitres 40 à 50, à l'exception du chapitre 49 (24).

On notera que l'on retrouve aussi sur la partie négative du premier axe, et positive du second axe, le groupe 36, 37, 38, 39.



Poursuivant l'analyse, on remarquera avec une certaine assurance, qu'au-delà de ces groupes de faible amplitude, il existe une division assez nette qui oppose tous les chapitres postérieurs au chapitre 39 à tous les chapitres antérieurs au chapitre 36. Cette division apparaît encore beaucoup plus marquée sur le premier axe lorsque dans un second temps on réitère l'analyse en annulant les chapitres 36 à 39 dans l'ensemble des chapitres (25). Comme on peut s'en rendre compte sur la figure 11, il est alors possible de séparer par un trait vertical les groupes cités plus haut en marquant deux ensembles assez bien différenciés. On a marqué par un cercle sur la figure 11 les exceptions à cette règle (26). Une telle répartition ne peut en aucun cas être le fruit du hasard. En effet, il faut rappeler ici qu'à *aucun moment* n'a été introduite dans la méthode de calcul la *notion de continuité des chapitres*. La méthode mise en oeuvre met donc en évidence les similitudes de la distribution des 89 vocables les plus fréquents, analysés ici, dans les chapitres consécutifs qui sont mis en évidence. Ce résultat nous conduit à rechercher des unités homogènes du point de vue du stock distributionnel du vocabulaire employé, formant des ensembles dont la taille dépassera le chapitre accepté au départ comme partition première du texte.

Nous serons, de plus, amenés à poser une question d'importance : certaines fluctuations de la place des chapitres sur les axes factoriels ne doivent-elles pas être attribuées à la trop petite taille du chapitre qui est notre unité de décompte ? En effet, il pourrait être possible que le comptage, dans ces unités, des mots les moins fréquents relève plus de l'observation de phénomènes de *type aléatoire* que de la mesure des fréquences relatives d'utilisation qui nous préoccupe. Prenons par exemple deux mots différents, l'un XYZ ayant 230 occurrences dans le livre. Ce mot devrait apparaître en moyenne 4 fois par chapitre si les chapitres étaient de tailles égales. On peut donc penser que si ce mot apparaît 20 fois dans chacun des  $N$  chapitres d'un groupe A et 5 fois dans chacun des  $N$  chapi-

AXE HORIZONTAL ( 1 )--AXE VERTICAL ( 2 )--TITRE: 62 CHAPITRES DANS ISAIE [ CHP. 36-39 EN SUPPLÉMENT ]

LARGEUR= 1.53797 HAUTEUR= 1.44384 -NOMBRE DE POINTS= 66

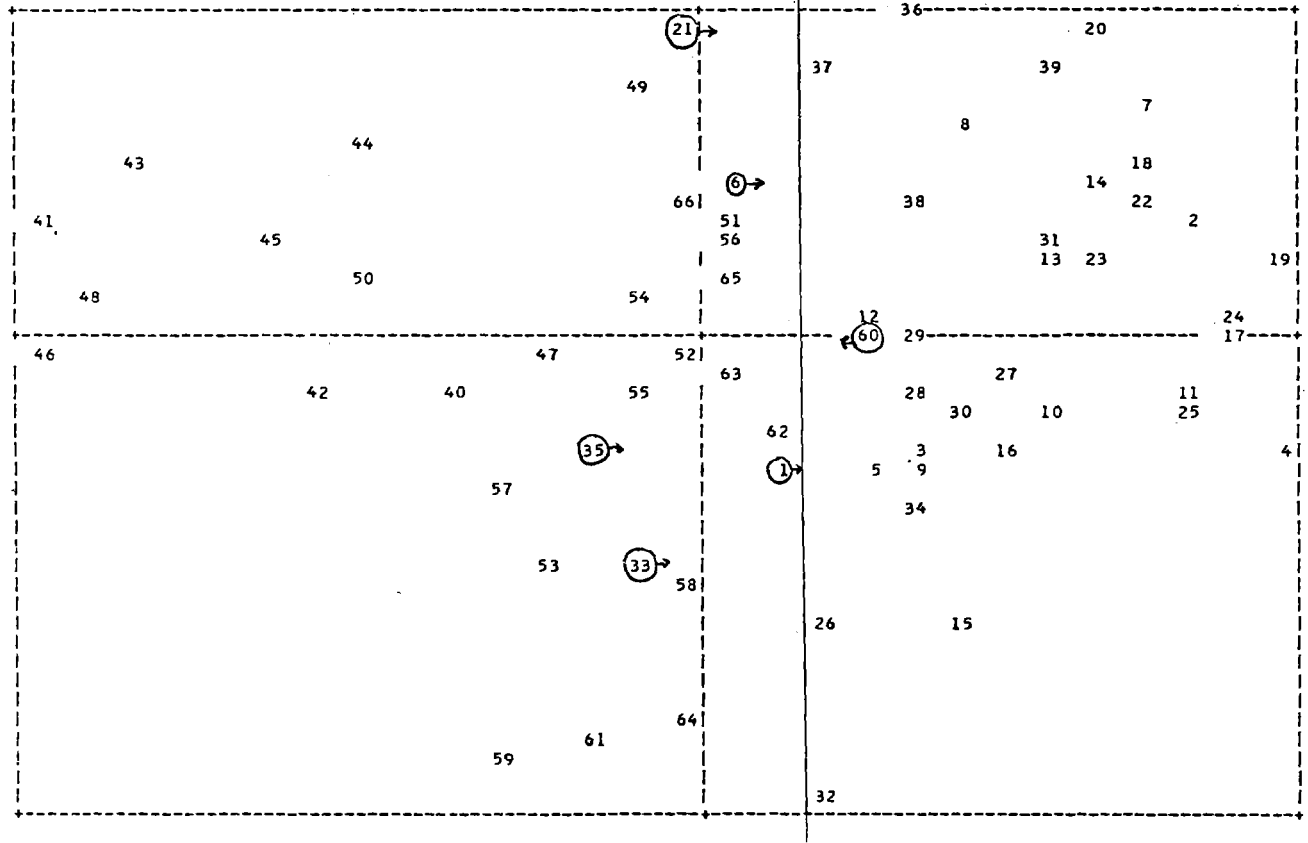


Fig. 11

LES CHAPITRES D'ISAIE SANS LES CHAPITRES 36-39

DANS LE PLAN DES AXES FACTORIELS

Extrait de la Revue (R.E.L.O.)  
XII, 1 à 4, 1976. C.I.P.L. - Université de Liège - Tous droits réservés.

tres d'un groupe *B*, nous avons affaire à des fréquences d'utilisation différentes provenant de deux "*situations*" différentes. Par contre, le mot *UVW* qui n'apparaît que 30 fois dans le livre devrait avoir une moyenne d'une "1/2 occurrence" par chapitre. Si ces occurrences étaient groupées dans quelques chapitres, elles seraient peut-être caractéristiques d'une situation thématique donnée, par contre, si elles se répartissaient à raison de 1 ou 2 par chapitre et étaient totalement absentes dans les autres, on ne pourrait accorder à cette répartition une importance décisive. C'est pourquoi, dans un second temps, ont été constituées des unités plus grandes que le chapitre qui ont été soumises au même type d'analyse.

## V

### *ANALYSE FACTORIELLE DES CORRESPONDANCES APPLIQUEES A DES FRAGMENTS DE N CHAPITRES CONSECUTIFS*

Dans la première partie de cette analyse, nous avons eu comme préoccupation première de ne pas avoir recours, dans nos essais d'analyse typologique au découpage que la critique biblique pouvait nous suggérer. C'est pour être sûr de ne pas céder à cette tentation, même à notre insu, que nous nous sommes orientés vers des découpages très arbitraires du livre en différentes parties de longueurs sensiblement égales. Le premier a été effectué en réalisant 13 fragments de 5 chapitres consécutifs. Nous avons effectué parallèlement un second découpage en 22 fragments de 3 chapitres consécutifs. Cette méthode de travail pourrait être critiquée par les philologues des écoles classiques car elle conduit à englober, dans une même unité, deux chapitres successifs dont les stocks lexicaux peuvent être très différents; elle ne représente qu'une étape dans la série des approxima-

tions qui conduisent au découpage le plus affiné possible auquel nous espérons arriver en constituant des ensembles très homogènes.

A l'issue de cette étape qui doit fournir une division typologique plus nettement délimitée des différentes parties qui composent l'ouvrage, l'unité des fragments constitués de façon aussi arbitraire sera remise en cause par la destruction des moins homogènes. L'analyse consistera alors à tester autour des coupures ainsi effectuées la proximité ou la non proximité des chapitres consécutifs.

#### *1. ANALYSE FACTORIELLE DU TABLEAU DE 13 FRAGMENTS DE 5 CHAPITRES*

##### *1.0. TABLEAU DES 13 FRAGMENTS DE 5 CHAPITRES (27)*

1	chp	1 - 5	8	chp	36 - 40
2	chp	6 - 10	9	chp	41 - 45
3	chp	11 - 15	10	chp	46 - 50
4	chp	16 - 20	11	chp	51 - 55
5	chp	21 - 25	12	chp	56 - 60
6	chp	26 - 30	13	chp	61 - 66
7	chp	31 - 35			

Dans ce premier découpage, le fragment le plus grand en nombre de mots comptés est de 2,8 fois plus important que le plus petit.

*1.1.* L'histogramme des valeurs propres de la matrice de 13 x 89 (figure 12) fait apparaître trois axes qui se détachent nettement avec 29,3, 17,6 et enfin 10 % d'inertie. On trouve ensuite deux pourcentages d'inertie à 6,9 puis 5,5; 5,3 %, etc.

LES VALEURS PROPRES					VAL(1)=C.59799321			
INDICE	ITER	VAL PROPRE	POURCENT	CUMUL	*)	HISTOGRAMME DES VALEURS PROPRES DE LA MATRICE 89 x 43		
1	2	0	C.17023065	25.324	29.324	*)	*****	*****
1	2	1	D.C7242657	17.664	45.938	*)	*****	*****
1	4	1	C.C3464338	10.878	57.836	*)	*****	*****
1	5	3	C.02645917	6.541	64.827	*)	*****	*****
1	6	1	C.02326546	6.874	71.721	*)	*****	*****
1	7	5	D.02275006	5.551	77.271	*)	*****	*****
1	8	2	D.C2157336	5.262	82.533	*)	*****	*****
1	9	2	C.C1726362	4.703	87.236	*)	*****	*****
1	10	2	C.C1753535	4.277	91.513	*)	*****	*****
1	11	2	C.C1417534	3.457	94.970	*)	*****	*****
1	12	2	C.C1071877	2.614	97.584	*)	*****	*****
1	13	1	C.C0950415	2.416	100.000	*)	*****	*****
LA VARIATION DU TEMPS=					7			

Fig. 12

1.2 En consultant sur le tableau (figure 13) les facteurs et les contributions aux deux premiers axes des mots soumis à l'analyse, on se rend tout de suite compte que le plan des deux premiers axes issu de cette analyse est le même que le plan des axes 1 et 2 que nous avons rencontré dans l'analyse précédente. Dans la mesure où nous avons fait des groupes de chapitres consécutifs et parce que des chapitres consécutifs sont *dans l'ensemble* semblables à l'intérieur de fragments opérés sur un ensemble du point de vue distributionnel, il n'y a pas lieu de s'étonner outre mesure de la similitude des résultats obtenus dans les deux analyses. On doit considérer ceci comme un raisonnement "parallèle" par rapport au principe d'équivalence distributionnelle énoncé plus haut.

1.3. Nous allons, dans cette analyse, nous consacrer plus particulièrement à l'étude des fragments sur les deux premiers axes; nous verrons ensuite qu'une interprétation peut être donnée au troisième axe, qui ne nous était pas apparue lors de la première analyse. Le petit nombre des fragments nous permet de parcourir beaucoup plus rapidement le tableau des contributions (fig. 14).





	J	Q	P	I	D	I	N	R	1#F	COR	CTR	2#F	COR	CTR	3#F	COR	CTR	4#F	COR	CTR	5#F	COR	CTR
1	1	690	72	59	210	132	27	-240	174	58	90	24	13	105	33	28	321	327	280				
2	2	358	95	50	279	363	62	84	33	9	8	0	0	-13	1	1	16	1	1				
3	3	464	95	52	328	290	53	-69	13	4	57	0	0	180	90	67	-139	54	41				
4	4	587	45	77	453	346	51	18	0	0	398	225	160	61	5	6	88	11	12				
5	5	722	61	69	375	304	71	-33	2	1	354	270	170	173	64	64	-194	62	82				
6	6	810	90	66	210	146	33	-179	107	40	13	1	0	-409	556	532	-3	0	0				
7	7	464	51	62	62	14	3	-300	181	63	-142	41	23	-257	133	118	-217	55	85				
8	8	574	121	138	117	29	14	642	882	688	-170	62	75	-14	0	1	15	1	1				
9	9	942	111	157	-654	831	445	48	4	4	248	106	153	24	1	2	-3	0	0				
10	10	719	79	79	-518	655	177	56	8	3	116	32	24	-94	22	25	-30	2	3				
11	11	300	71	56	-118	44	8	-77	19	6	-271	228	118	51	8	6	-20	1	1				
12	12	718	77	64	-131	51	11	-220	142	52	-260	198	117	95	26	24	321	301	282				
13	13	746	68	71	-95	24	6	-273	176	70	-301	214	139	229	124	126	-296	206	213				
I		6764.0	1000			1000			1000		1000			1000			1000			1000			

Fig. 14

1.3.1. On retrouve sur le premier axe l'opposition entre les premiers et les derniers fragments. Pour les 5 derniers fragments le facteur est très négatif alors qu'il est positif pour les premiers. Comme dans la première analyse, les contributions les plus fortes sont dues aux chapitres 41-50, représentés ici par les fragments 9 et 10, qui portent à eux seuls 60 % de l'inertie.

1.3.2. Le second axe est constitué par l'opposition, à tous les autres fragments, du fragment 8 constitué par les chapitres 36, 37, 38, 39 et 40. Comme précédemment, ces fragments sont très bien expliqués par le plan des axes 1 et 2, si l'on consulte la figure 15.

1.3.3. Le troisième axe oppose les fragments 9 (chapitres 41-45), 4 (chapitres 16-20) et 5 (chapitres 21-25) (28) du côté positif, aux fragments 11 (51-55), 12 (56-60) et 13 (61-66) du côté négatif. Nous allons chercher sur le troisième axe quels sont les mots qui ont la plus forte contribution à cet axe.

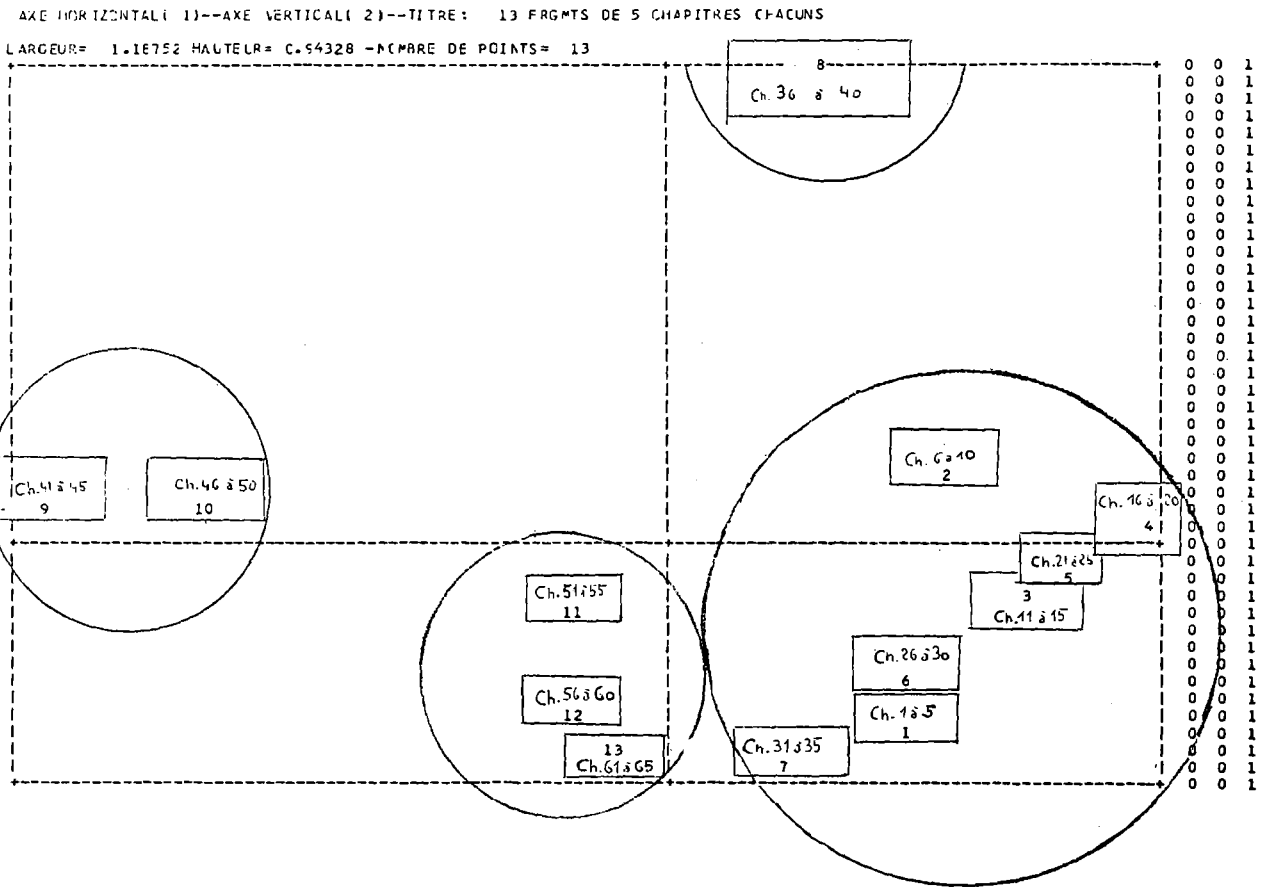


Fig. 15

LES 13 FRAGMENTS SUR LE PLAN DES AXES FACTORIELS

1.3.3.1. Du côté positif, ce sont :

<i>HEBREU</i>	<i>MOTS</i>	<i>3    F</i>	<i>COR</i>	<i>CTR</i>
	<i>II</i>			
הַהוֹא	HHV%	751	344	84
צָבָא	3B%	564	382	72
נִגְדָא	NGD	384	732	63
קָשָׁב	YWB2	556	84	35
אֲנִי	%NY	456	106	53

1.3.3.2. Du côté négatif :

עוֹלָם	JVLM	- 697	353	74
דָּרָךְ	DRK	- 500	492	39

Les deux derniers mots sont assurément caractéristiques du groupe des cinq derniers chapitres, mais ce groupe, on le voit, se caractérise aussi par l'absence quasi simultanée de הַהוֹא - HHV% - et de אֲנִי - %NY -, dont l'opposition forme le premier axe. On notera également que le fragment 7 (chapitres 31-35) se détache sur cet axe du groupe des chapitres 1-30.

1.4. L'étude à laquelle nous venons de nous livrer, concernant les résultats de l'analyse factorielle du tableau des 89 mots les plus fréquents regroupés en 13 fragments de 5 chapitres, confirme donc les grandes oppositions constatées par l'analyse factorielle du tableau des 89 x 66 dans le plan du premier et du second axe.

Un troisième axe, qui porte un pourcentage d'inertie important, oppose les fragments 4 et 5 (chapitres 16-25) aux fragments 11, 12, 13 (chapitres 51-65). Alors que le fragment 8 (chapitres 36-40) reste isolé, certains fragments se rassemblent en petits groupes compacts, tels les fragments 9 et 10 d'une part, 11, 12 et 13 d'autre part. On peut donc penser que pour ces derniers fragments, ce groupement de cinq chapitres en un fragment ne constitue pas l'unité homogène maximale et que l'on peut envisager la fusion en un seul fragment homogène du groupe 9-10 d'une part, et 11, 12 et 13 d'autre part, aux extrémités près.

## 2. ANALYSE FACTORIELLE DU TABLEAU DE 89 MOTS x 22 FRAGMENTS DE TROIS CHAPITRES

### 1.0. TABLEAU DES 22 FRAGMENTS DE 3 CHAPITRES

1	chp	1 - 3	8	chp	22 - 24	15	chp	43 - 45
2	chp	4 - 6	9	chp	25 - 27	16	chp	46 - 48
3	chp	7 - 9	10	chp	28 - 30	17	chp	49 - 51
4	chp	10 - 12	11	chp	31 - 33	18	chp	52 - 54
5	chp	13 - 15	12	chp	34 - 36	19	chp	55 - 57
6	chp	16 - 18	13	chp	37 - 39	20	chp	58 - 60
7	chp	19 - 21	14	chp	40 - 42	21	chp	61 - 63
						22	chp	64 - 66

1.1. L'analyse de ce tableau donne exactement les mêmes résultats. Pour s'en convaincre, on se reportera à la figure et au tableau des facteurs et des contributions que nous donnons ci-dessous.

	1#F	2#F	3#F	4#F	5#F
11	225	47	37	171	63
21	355	32	35	150	42
31	463	57	32	345	36
41	532	37	45	302	133
51	351	42	38	314	136
61	378	24	35	373	177
71	452	22	39	390	271
81	652	42	47	440	291
91	452	30	49	276	44
101	559	66	44	216	118
111	274	32	33	51	4
121	652	45	47	235	33
131	832	69	77	233	122
141	627	61	71	-624	569
151	754	62	73	-630	623
161	794	44	57	-733	546
171	549	52	43	-134	71
181	337	37	34	-124	29
191	410	43	37	-206	93
201	523	46	45	-114	23
211	534	33	35	-142	32
221	280	58	35	-74	15
	10001	10001	10001	10001	10001

Fig. 16

Rappelons seulement, pour mémoire, que les types de chapitres, que nous avons mis en évidence dans les analyses précédentes, portent, dans cette analyse, des identificateurs différents de ceux de l'analyse du tableau de 13 x 89 :

- Fragment 12 = chapitres 34, 35, 36
- Fragment 13 = chapitres 37, 38, 39
- Fragment 14 = chapitres 40, 41, 42
- Fragment 15 = chapitres 43, 44, 45
- Fragment 16 = chapitres 46, 47, 48
- Fragments 18-22 = chapitres 52-66.

AXE HORIZONTAL ( 2 ) -- AXE VERTICAL ( 1 ) -- TITRE: 22 FRGMS DE 3 CHAPITRES CHACUNS

LARGEUR= 1.14777 HAUTEUR= 1.14394 -NOMBRE DE POINTS= 22

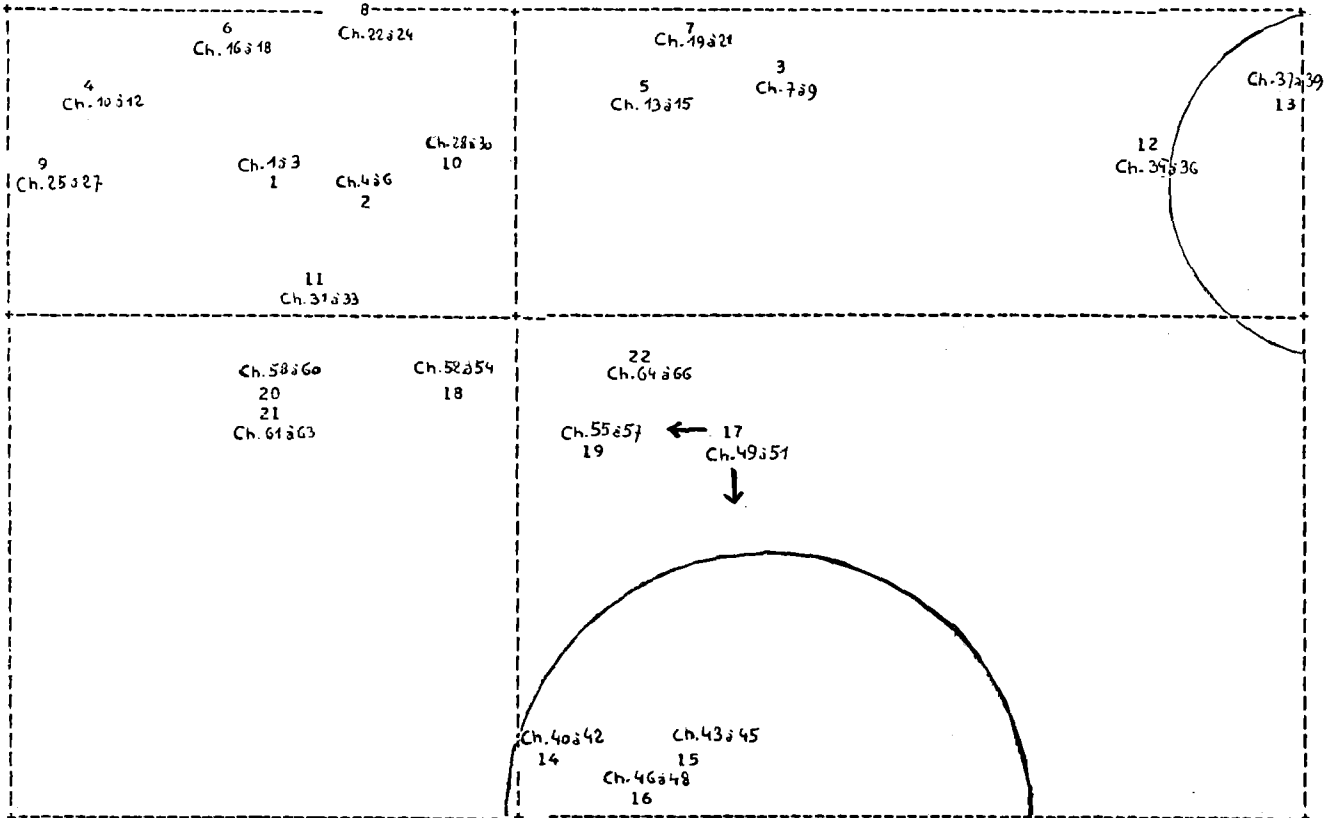


Fig. 17. - Les 22 fragments sur le plan des axes factoriels.

### 3. PREMIERES CONCLUSIONS SUR LES RESULTATS DE L'ANALYSE FACTORIELLE

3.1. Nous venons d'exposer les premiers résultats fournis par l'*Analyse Factorielle* à partir du tableau de données de 89 x 66 constitué par nous pour être soumis à l'analyse (29); la recherche des premières conclusions s'impose maintenant. Ce tableau a été analysé par lignes et par colonnes, ce qui représentait ce faisant un fractionnement en 66 fragments différents, semblables à la division capitulaire du livre. Les résultats de cette première analyse ont été confrontés aux différentes analyses du même tableau par lignes et par la somme représentant l'addition du groupement d'un certain nombre de colonnes consécutives : d'abord 5 colonnes représentant 5 chapitres divisant ainsi l'ouvrage en 13 fragments, puis 3 colonnes représentant 3 chapitres divisant l'ouvrage en 22 fragments. Les résultats obtenus ont été corroborés, par la suite, par les mêmes types d'analyses sur d'autres fragmentations des mêmes données. Nous avons dû constater qu'à travers tous ces résultats s'imposait une quadripartition du *Livre d'Isaïe* résultant de l'analyse de la répartition des 89 mots les plus fréquents. L'existence de quatre grands ensembles nous était imposée par l'hétérogénéité réciproque de ces ensembles, sensible pour chaque analyse à laquelle était soumis le tableau. Les quatre grands ensembles variaient peu quant au nombre de fragments qui les constituaient, mais ils variaient cependant sur leurs franges selon le découpage en fragments plus ou moins importants des données, soumis à l'analyse.

3.2. La première conclusion à tirer sera que pour user de l'*Analyse Factorielle* les fragments soumis à l'analyse ne peuvent être trop petits. Des fragments composés par l'addition des données dans 5 chapitres consécutifs formant un ensemble qui divise le livre en 13 fragments contenant théoriquement  $6941 : 13 = 590$  mots chacun (30), ont fait apparaître les oppositions les plus marquées entre des ensembles composés de chapitres consécutifs. L'*Analyse Factorielle* de



la distribution des occurrences des mots les plus fréquents a mis en évidence un certain nombre de regroupements entre des fragments, arbitrairement définis, en fonction de la plus ou moins grande similitude dans les distributions des mots, compte tenu du poids propre de chacun des fragments.

3.3. Cette méthode d'analyse appliquée à des données représentant la distribution des occurrences des 89 mots les plus fréquents dans les 66 chapitres du livre a été la plus objective possible, puisqu'elle a été conduite sans hypothèse explicite, donc sans modèle de référence ou sans découpage induit selon une école de critique textuelle quelconque. L'analyse factorielle a imposé à la lecture des tableaux calculés et édités par le programme (axes factoriels, représentation des éléments par les axes et contribution des éléments aux axes factoriels) la prise en compte assez large de quatre grands ensembles. Le graphique de la projection des éléments sur une surface, édité par le programme à l'intention des utilisateurs, n'est en fait qu'un artefact qui tente de représenter dans un espace à deux dimensions les projections des éléments contenus dans l'un ou l'autre des volumes multidimensionnels calculés par le programme (31). La mise à plat d'un tel volume ne peut être opérée sans quelques distorsions imposées aux positions des éléments projetés par rapport à leur position relative dans le volume calculé par le programme, et la lecture du seul graphique est insuffisante pour tirer un enseignement des premières constatations. Le graphique des projections dans le plan doit être considéré comme devant guider l'utilisateur dans sa lecture des tableaux calculés par le programme, en l'incitant à vérifier par les valeurs qui y figurent les franges imprécises des ensembles qui semblent se dessiner à la projection. En conclusion, le graphique ne peut jamais être considéré comme se suffisant à lui-même; s'il était utilisé seul, il conduirait l'utilisateur à affirmer des contre-vérités grossières.

3.4. De cette première démarche, il résulte que s'est imposée à nous par le calcul, de façon tout à fait objective, une certaine quadripartition dont nous allons donner ci-après la distribution. Cette quadripartition marque au premier chef la consécutive de certains groupes de chapitres et, par là, une certaine homogénéité intérieure dans les grands ensembles observés. Plus que l'homogénéité interne, ces distributions matérialisent la forte hétérogénéité, dans la distribution des mots les plus fréquents du livre, entre les quatre groupes ainsi mis en évidence par l'analyse.

3.4.1. La quadripartition la plus largement ressentie est :

- I Chp. 1 à 35
- II Chp. 36 à 39 (32)
- III Chp. 40 à 50
- IV Chp. 51 à 66

3.4.2. A l'intérieur de ces IV grands groupes on trouve des oppositions plus faibles qui semblent opposer des ensembles consécutifs de plus petites dimensions :

- 3.4.2.1. I<sub>1</sub> Chp. 1 à 30  
I<sub>2</sub> Chp. 31 à 35  
II<sub>1</sub> Chp. 36 à 39  
II<sub>2</sub> Chp. 38 dont le sensible éloignement semble marquer une certaine distance par rapport à II<sub>1</sub>.

3.4.2.2. Cet éloignement du groupe des chapitres historiques dont on sait qu'ils sont issus d'une même source que leurs parallèles du *Livre des Rois* est peut-être dû à la longueur de la lettre d'Isaïe à Ezéchias, dont le poids

l'emporte par rapport au reste du narratif des Chroniques royales du chapitre 38, à la différence de la première lettre d'Isaïe à Ezéchias dont le poids est moins fort par rapport au narratif des Chroniques royales du chapitre 37.

- 3.4.2.3. III Chp. 40 à 50 marquant toujours une grande cohérence  
sauf pour le Chp. 49 (Cf. Annexe 2)
- IV<sub>1</sub> Chp. 51 à 55
- IV<sub>2</sub> Chp. 56 à 60
- IV<sub>3</sub> Chp. 60 à 66

3.5. Nous ne voulons à ce stade de la recherche aller au-delà de ces constatations ni sur le plan statistique ni sur le plan de l'analyse critique textuelle. Nous ne pourrions encore tirer aucune conclusion quant à ce que peut exprimer cette partition sur le plan de la langue, de la thématique ou des auteurs supposés.

## VI

### *LA CLASSIFICATION ASCENDANTE HIERARCHIQUE SUR LES FRAGMENTS DE CHAPITRES CONSECUTIFS*

Le reproche le plus couramment fait à l'*Analyse Factorielle des Correspondances* résulte de la difficulté que rencontre le chercheur, une fois les grands ensembles typologiques définis, de cerner le contour exact de chacun de ces ensembles,

sans céder à la tentation toujours présente de regrouper ceux qu'il considère comme cohérents en intuitant le caractère commun des éléments en fonction de ses préoccupations. Le chercheur qui prescrit des ensembles en traçant des cercles autour des groupes de points présentant un même caractère sur un plan factoriel, a le sentiment d'accomplir sur ces projections une opération tout intuitive, dont le caractère subjectif, du fait que cette opération n'est pas du tout formalisée, est évident. L'intuition des regroupements qu'il propose et qui ne sont pas imposés par les résultats de l'analyse lui apparaît, de ce fait, comme un choix tout à fait conjectural qui gauchit la démarche mathématique dont il ne s'était pas départi jusque-là.

C'est à ce genre de préoccupations que s'efforcent de répondre les méthodes de *Classification Ascendante Hiérarchique* que nous allons exposer. Partant du même tableau de données que l'on a utilisé tout au long de cet article, ces méthodes d'analyse permettent de constituer des hiérarchies de classes aux contours très strictement définis (33). Comme cela est souvent le cas, en recourant à de nouvelles méthodes d'analyse le chercheur est confronté à d'autres inconvénients que ceux qu'il a pu déplorer par ailleurs. Aucune méthode ne permet, en vérité, d'atteindre un degré de perfection idéal dans l'analyse. La *Classification Automatique*, pour sa part, présente deux inconvénients majeurs. Le premier réside dans ce que les résultats obtenus sont beaucoup plus sensibles aux différences dans les découpages initiaux, car ceux-ci se répercutent en s'amplifiant à tous les niveaux de la classification hiérarchique, alors que les résultats obtenus par l'*Analyse Factorielle*, toutes conditions égales conservées, présentent plus de stabilité. Le second inconvénient réside dans la perte de la disposition relative des classes les unes par rapport aux autres, qui sont plus aisément perceptibles en *Analyse Factorielle*. Nous avons, néanmoins, utilisé cette seconde méthode pour corroborer les premiers résultats auxquels nous étions arrivés et dont nous venons de rendre compte.

1. LE TABLEAU DES DISTANCES

Le programme de Classification Automatique (34) commence par calculer toutes les distances qui existent entre les différents fragments utilisés précédemment, pris deux à deux. Ces distances sont rangées dans une matrice de 13 x 13.

Cette matrice est symétrique car

$\delta (i, i') = \delta (i', i)$ ; de plus, la diagonale principale est nulle car

$\delta (i, i') = 0$ .

TABLEAU DES DISTANCES  
\*\*\*\*\*

	1	2	3	4	5	6	7	8	9	10	11	12	13
1													
2	0.68												
3	0.77	0.71											
4	0.92	0.88	0.88										
5	0.83	0.69	0.72	0.86									
6	0.75	0.66	0.80	0.93	0.84								
7	0.91	0.88	0.89	1.12	1.00	0.82							
8	1.03	0.75	0.94	1.08	1.01	0.97	1.11						
9	1.08	1.07	1.17	1.33	1.19	1.10	1.13	1.11					
10	1.00	0.96	1.04	1.26	1.09	0.94	1.03	0.99	0.58				
11	0.87	0.78	0.86	1.13	1.02	0.83	0.91	0.94	0.93	0.86			
12	0.77	0.84	0.91	1.14	1.03	0.84	0.89	1.01	0.95	0.84	0.77		
13	0.91	0.87	0.91	1.19	1.01	0.89	0.91	1.08	1.01	0.93	0.80	0.80	

Fig. 18

En fonction du *Tableau des Distances* édité par le programme, on observera que les distances les plus petites sont, dans l'ordre :

plus petites distances	entre les fragments
0.58	9 et 10
0.66	6 et 2
0.68	2 et 1
0.69	5 et 2
0.71	2 et 3
0.72	3 et 5
0.75	1 et 6
0.75	2 et 8
0.77	1 et 3
0.77	1 et 12
0.77	12 et 11
0.78	2 et 11

On peut voir là toute la complexité du problème posé par la Classification Automatique : certains fragments sont assez "centraux", comme le fragment n° 2 et, de ce fait, ils sont assez proches de plusieurs fragments à la fois. Ainsi, par exemple, le fragment 2 est assez proche du fragment 8 (chapitres 36-40) dont la spécificité a été mise en évidence à plusieurs reprises au cours de cet article (35), mais il est beaucoup plus proche des fragments 6, 1, 5 et de quelques autres fragments.

Les distances les plus importantes sont :

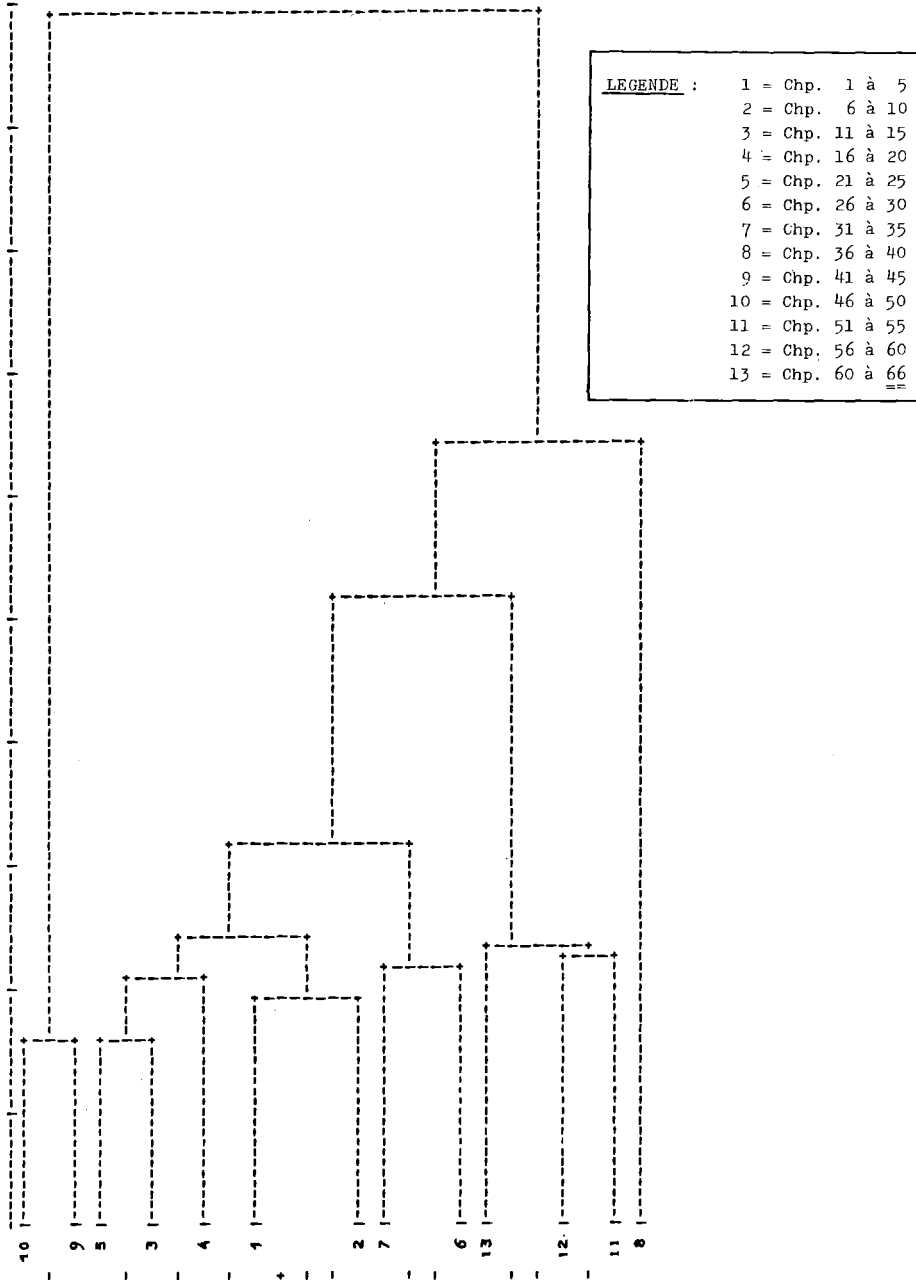
plus importantes distances	entre les fragments
1.33	9 et 4
1.26	10 et 4
1.19	13 et 4
1.19	9 et 5
1.17	9 et 3
1.14	12 et 4
1.13	11 et 4
1.13	9 et 7
1.12	7 et 4
1.11	9 et 8
1.11	8 et 7
1.10	9 et 6

Dans ces distances, les fragments 9 et 4 interviennent 6 fois chacun; la distance entre 9 et 4 est d'ailleurs la plus importante. Tout cela confirme bien les résultats de l'analyse factorielle du tableau 13 x 89.

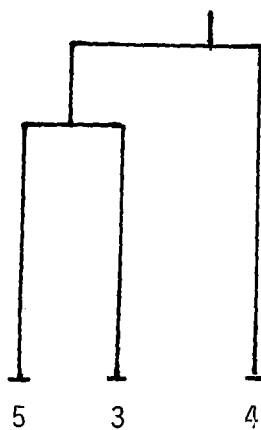
## 2. LE TRACE DE L'ARBRE

Considérons maintenant le tracé édité par le programme de Classification Arborescente. Pour comprendre cette représentation, il faut se souvenir en permanence que les seules proximités à interpréter sur cette arborescence sont celles qui font l'objet de la formation d'une classe. Par exemple, sur le graphique (figures 19 et 20) le fait que 5 et 3 soient réunis très bas dans l'arbre témoigne d'une proximité certaine entre eux. On ne peut cependant déduire, en aucune façon, de cette représentation que le fragment 4 est plus proche du fragment 3 que du fragment 5.

CLASSIFICATION HIERARCHIQUE ASCENDANTE  
DES 13 FRAGMENTS D'ISAIE







*Fig. 19*

On voit donc que le même arbre peut être représenté de plusieurs façons différentes, les classes et la hauteur des liaisons restant les mêmes. Il y a dans cette représentation une analogie certaine avec la mise à plat d'un mobile spatial.

### 3. ANALYSE DES STRATES DE LA HIERARCHIE

3.1. Si nous examinons les diverses strates de la hiérarchie sur le tableau édité par le programme, on pourra faire quelques remarques importantes sur l'algorithme d'agrégation hiérarchique (36).

CRITERE D'AGREGATION: MAXIMISATION DU MOMENT CENTRE D'ORDRE 2 D'UNE PARTITION																	
J	I(J)	A(J)	B(J)	C(J)	D(J)	E(J)	F(J)	G(J)	H(J)	I(J)	J(J)	K(J)	L(J)	M(J)	N(J)		
14	0.015	10	9	2	10	9											
15	0.015	5	3	2	5	3											
16	0.019	1	2	2	1	2											
17	0.021	5	3	3	5	3	4										
18	0.022	7	6	2	7	6											
19	0.022	12	11	2	12	11											
20	0.023	13	12	3	13	12	11										
21	0.024	17	16	5	5	3	4	1	2								
22	0.032	21	19	7	5	3	4	1	2	7	6						
23	0.052	22	20	10	5	3	4	1	2	7	6	13	12	11			
24	0.065	24	8	11	5	3	4	1	2	7	6	13	12	11	8		
25	0.101	14	24	13	10	9	5	3	4	1	2	7	6	13	12	11	8

Fig. 21

3.1.1. La classe 14 est formée par la réunion des fragments 9 et 10. Il faut noter tout de suite à ce sujet que cette classe ne sera réunie aux autres qu'à la dernière étape de l'algorithme de classification.

3.1.2. La classe 15 agrège les fragments 5 et 3.

3.1.3. La classe 17 agrège aux fragments 5 et 3 le fragment intercalé 4, créant ainsi la classe de fragments consécutifs 3, 4, 5; il faut noter que c'est le seul exemple d'agrégation de fragments de 5 chapitres non consécutifs.

3.1.4. La classe 16 réunit les fragments 1 et 2.

3.1.5. La classe 18 agrège les fragments 7 et 6.

3.1.6. Les classes 19 et 20 agrègent les fragments 11, 12 et 13.

3.1.7. La classe 21 constitue la classe de fragments consécutifs de 1 à 6.

3.1.8. La classe 22 agrège les fragments 7 et 6 à la précédente. Cette classe contient alors tous les chapitres de 1 à 35.

3.1.9. La classe 23 réunit la classe des chapitres consécutifs 51-65 avec cette dernière classe.

3.2. A partir de cette agrégation, on agrège des morceaux non consécutifs :

3.2.1. La classe 24 est formée par le rattachement du fragment 8 à la précédente.

3.2.2. La classe 25 représente la totalité des agrégations successives.

#### 4. ANALYSE DES TABLEAUX DE CONTRIBUTIONS

Les tableaux de contributions obtenus par le programme nous renseignent sur la contribution de chaque mot à la formation de chacune des classes, soit par une abondance marquée, soit par une défection marquée dans la classe considérée.

4.1. *La classe 14* - chapitres 41-50 - formée par les fragments 9 et 10 - est due tout particulièrement à l'abondance de %NY, %P, NGD, %YN, ce que nous savions déjà.

4.2. *La classe 22* - chapitres 1-35 - formée par les fragments 1 à 7 - est due à l'abondance des vocables HYH, JL, HHV%, 3B% et JM et à la sous-représentation de %NY et JBD.

4.3. *La classe 20* - chapitres 50 à 65 - formée par les fragments 11, 12 et 13 - est due principalement à l'abondance de JVLM, 3DQH, DRK, ZRJ, WM, TXT; par contre, 3B% y est sous-représenté.

#### 5. CONCLUSIONS DE L'ANALYSE DU TABLEAU DE 89 X 13 PAR LA METHODE DE CLASSIFICATION AUTOMATIQUE

5.1. Ces résultats confirment et complètent ceux que nous avons obtenus par l'Analyse Factorielle. On remarquera tout particulièrement l'homogénéité du couple 9-10 (chapitres 40-45 et 46-50) qui se différencie de tous les autres fragments. La spécificité du fragment 8 (chapitres 36-40) est confirmée. Les fragments 11, 12 et 13 (chapitres 50-65) semblent former un groupe assez distinct du groupe de chapitres 1-35.

5.2. La contre-épreuve qui consiste à réitérer l'analyse en donnant aux chapitres 36 à 39 une masse évanescence, c'est-à-dire en supprimant leur influence sur les résultats de l'analyse, a démontré que les trois grands ensembles de la quadripartition maintenant réduite se maintenaient avec la même stabilité (figure 22) (37). La contre-épreuve confirme comme précédemment les résultats obtenus. Une telle répartition ne peut en aucun cas être le fruit du hasard; il sera nécessaire d'en rechercher les causes, compte tenu de la nature des données qui ont été soumises à l'analyse.

5.3. On verra cependant d'autres données, telles que les données morphologiques dites *stylistiques*, proposées par Asa Kasher dans son article consacré à la

CLASSIFICATION HIERARCHIQUE ASCENDANTE  
 DES 12 FRAGMENTS D'ISAIE  
 (sans les Chp. 36-39)

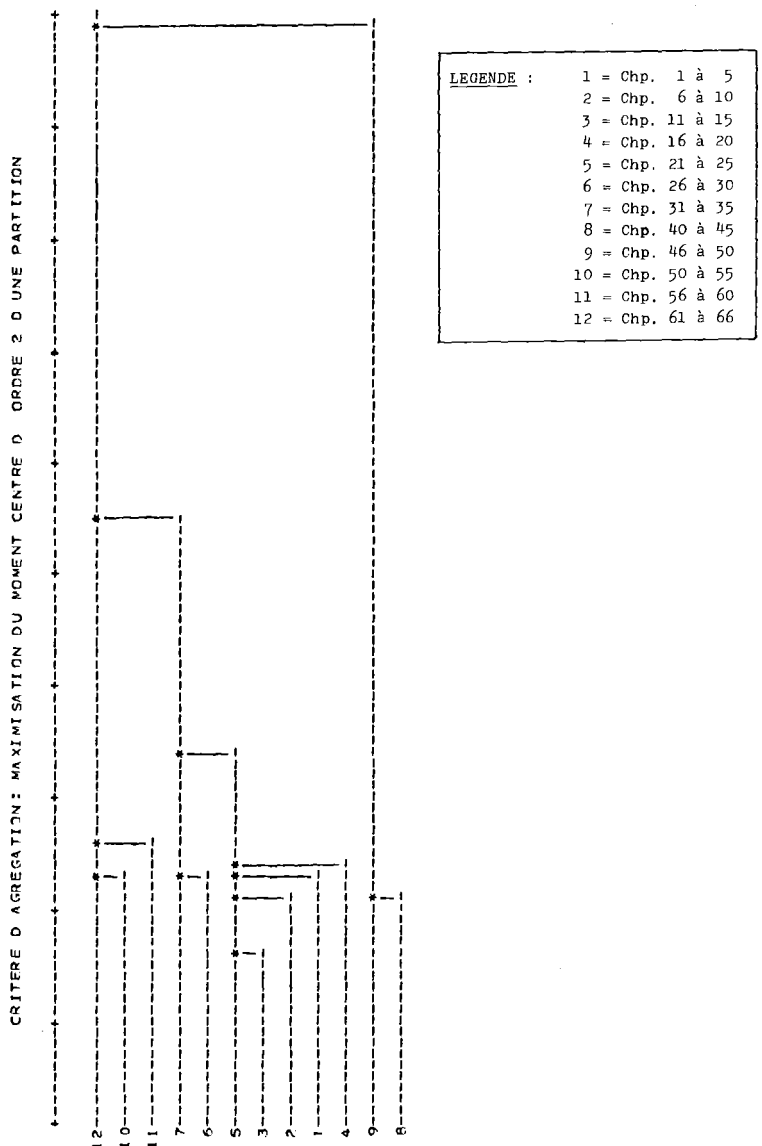


Fig. 22

recherche sur l'hétérogénéité d'auteurs dans le livre d'Isaïe (38), ont été soumises par nous à l'Analyse Factorielle. Les résultats obtenus à partir du vocabulaire des mots les plus fréquents (39). Dès l'instant où, partant de données d'autre nature tirées du même ouvrage et traitées par les mêmes méthodes, on peut obtenir des résultats différents quant à la distribution des fragments, la question qui doit être posée sera de rechercher dans les données traitées les raisons des résultats différents obtenus par les mêmes méthodes.

## VII

### *EN RESUME ...*

Le lecteur voudra bien reconnaître que les auteurs, au cours de cette étude, se sont toujours refusés à trancher en matière d'homogénéité ou d'hétérogénéité d'auteur. Ils se sont fait un devoir de rendre compte le plus objectivement possible de leurs recherches sur l'application des résultats qu'ils ont pu obtenir durant la phase exploratoire des méthodes d'*Analyse Factorielle* et de *Classification Ascendante Hiérarchique* en matière d'analyse textuelle. Usant pour la première fois des méthodes d'analyse multidimensionnelle pour analyser un texte à l'autorité d'auteur si controversée, nous avons voulu savoir, avant toute autre chose, quelle était la limite de ces méthodes et quels étaient les résultats que l'on pouvait en escompter. Nous avons cherché, en premier lieu, à mettre en évidence des fragments homogènes du point de vue du stock du vocabulaire employé, quelles que puissent être les raisons auxquelles peut être rattachée l'homogénéité des ensembles relevés. Nous avons pu constater que, plus que l'homogénéité interne des grands ensembles délimités par les méthodes d'analyse que nous avons employées, celles-ci faisaient ressortir la forte hétérogénéité

réci-proque des ensembles du point de vue de la distribution des mots les plus fréquents. Nous avons essayé pour cela d'éviter deux écueils que les essais tentés par les deux méthodes nous avaient laissé pressentir :

*1) Soumettre à l'analyse des unités trop petites de décompte de mots, qui pouvaient ne présenter que des variations de caractères purement aléatoires, nous entraînant hors des domaines d'application de la méthode.*

*2) Soumettre à l'analyse des unités trop importantes de décompte de mots, qui pouvaient nous amener à considérer comme un même objet ce qui n'était qu'une moyenne de deux objets très différents.*

Nous pensons que la classification typologique obtenue par l'*Analyse Factorielle* du tableau 89 x 66, affinée par la *Classification Automatique*, autant que les résultats de l'*Analyse Factorielle* sur 13 fragments de 5 chapitres chacun, confirmés en *Analyse Factorielle* sur 22 fragments de 3 chapitres chacun - bien que la classification automatique sur ces mêmes 22 fragments ait donné des résultats légèrement différents - reflète sans doute des différences réelles dans la distribution du stock du vocabulaire le plus fréquent, qui s'imposent à l'analyse philologique et dont le chercheur pourra tirer un utile parti.

Un point très important sur le plan méthodologique reste à résoudre et réside dans la possibilité d'arriver à affiner les découpages obtenus automatiquement, pour que les coupures ne soient plus effectuées de manière aussi arbitraire. Ces découpages doivent être nécessairement indépendants de tout découpage proposé par les méthodes de critique textuelle, auxquelles notre groupe de recherche souhaite confronter les résultats obtenus par des procédés automatisés nouveaux, dégagés de toute tentative impliquant un choix a priori. Les recherches qui sont

actuellement poursuivies tendront à développer des méthodes qui, partant des classifications obtenues sur la base de découpages très arbitraires, permettront d'obtenir des découpages plus nuancés en des fragments homogènes d'une plus grande précision. C'est en partie à l'exposé de ces méthodes que sera consacrée la suite de cette étude.

Par ailleurs, les parties distinctes et opposées mises en évidence par ces méthodes dans un livre donné ne pourront conduire à les interpréter comme le signe d'une hétérogénéité d'auteur que dans la mesure où l'emploi des mêmes méthodes sur des ouvrages d'auteurs dont l'unicité n'est pas mise en doute donnera un degré de cohérence interne nettement supérieur à celui obtenu dans l'ouvrage à l'autorité d'auteur controversée. Si les résultats obtenus par ces méthodes sur l'oeuvre d'un auteur dont l'unicité n'est pas contestée devaient être comparables, deux conclusions pourraient être possibles. La première serait que les oppositions des parties dans l'ouvrage à l'unicité d'auteur contestée ne seraient pas le fait d'une pluralité d'auteur, la seconde étant que ces méthodes souligneraient d'autres propriétés que celles permettant de déterminer l'autorité des parties, dont les implications resteraient à déterminer et qui pourraient porter sur des faits de langue, de thématique ou de genre littéraire.

*à suivre...*



## ANNEXE I

### CODE DE ROMANISATION DE L'HEBREU UTILISE AU CRAL. A NANCY

<i>HEBREU</i>	<i>NOM</i>	<i>PRONONCIATION</i>	<i>CODE</i>
ב	bèt	b	B
ג	gimel	g dur	G
ד	dalèt	d	D
ה	hé	h aspiré allemand	H
ו	waw	w anglais	V
ז	zaiñ	z français	Z
י	yod	y de yacht	Y
כ	kaph	k	K
ל	lamèd	l	L
מ	mèm	m	M
נ	nun	n	N
ס	samèkh	s dur	C
פ	pé	p	P
ק	qoph	q emphatique	Q
ר	rèsh	r	R
ש	sin	s dur	S
ת	taw	t	T

Deux autres consonnes ont été codées à cause de leur prononciation gutturale, dont l'une rappelle le χ (khi) grec :

ח	hèt	h aspiré emphatique	X
---	-----	---------------------	---

et l'autre la *jota* espagnole :

י	'ayin	aspiration sonore	J
		● emphatique très gutturale	

Deux consonnes ont été codées en tenant compte de ce que leur forme manuscrite hébraïque ressemblait à deux chiffres de la série des unités et qu'elles pouvaient ainsi être facilement mémorisées :

ט	ṭèt	ṭ emphatique	6
צ	ṣadé	s emphatique	3

Enfin l'*alèph* hébreu à cause de la ressemblance avec le code :

א	'alèph	gutturale faible	%
---	--------	------------------	---

et le *shin* hébreu à cause de la ressemblance avec le code :

שׁ	shin	se prononçant ch	W
----	------	------------------	---

## ANNEXE II

### ANALYSE DES CAUSES DE L'ECART DU CHAPITRE 49 DANS LE GROUPE DES CHAPITRES 40 A 50(40)

Considérons la quantité  $\frac{P(I, J)}{SOMC(J)} \times 1000$

où  $P(I, J)$  représente le nombre d'occurrences du mot  $I$  dans le chapitre  $J$  et  $SOMC(J)$  représente le nombre de mots comptés dans le chapitre  $J$ . Cette quantité nous donne le profil de la colonne  $J$ , multiplié par 1000 pour faciliter les calculs. Il faut cependant noter que cette quantité n'exprime pas la fréquence relative du mot, mais qu'elle est toutefois représentative par rapport à l'ensemble des mots du chapitre.

On trouvera dans le tableau donné ci-dessous les valeurs de cette quantité à propos des chapitres 40 à 49 pour les mots dont la contribution au premier axe factoriel (1 # F) est la plus forte, axe sur lequel s'écarte particulièrement le chapitre 49.

	(-)HYH	YDJ	NGD	(-)JL	L%	%YN	%P	%NY	YVM
CH40	7.	28.	7.	21.	90.	42.	21.	0.	0.
CH41	14.	27.	34.	7.	34.	61.	41.	54.	7.
CH42	8.	23.	15.	38.	136.	15.	8.	23.	0.
CH43	7.	14.	14.	0.	95.	20.	14.	54.	7.
CH44	7.	20.	20.	47.	73.	20.	27.	20.	0.
CH45	6.	29.	17.	23.	81.	52.	6.	64.	6.
CH46	14.	0.	14.	29.	100.	14.	71.	100.	0.
CH47	32.	43.	0.	97.	129.	43.	0.	22.	11.
CH48	21.	35.	35.	7.	106.	7.	21.	57.	7.
CH49	25.	13.	0.	25.	32.	0.	0.	38.	6.

Dans le tableau les quantités marquant une fréquence très différente dans l'utilisation d'un mot dans un chapitre donné par rapport à son utilisation dans les autres chapitres ont été marquées par des caractères en italique. On constate que le chapitre 49 est souvent cité :

NGD = 0

L % = 32 C'est le chapitre qui présente la plus grande proportion

% YN = 0 d'occurrences nulles.

% P = 0

Vient ensuite le chapitre 47 avec NGD = 0; % P = 0; JL = 97.

Sur le second axe factoriel, on notera un suremploi du verbe % MR pour le chapitre 49 en commun avec le chapitre 44.

*TABLEAU DE PARTITION D'ISAIE D'ASA KASHER*

A	1 - 2,5	O	41
B	2,6 - 4	P	42
C	5	Q	43
D	6 - 9,6	R	44
E	9,6 - 12	S	45
F	13 - 14	T	46 - 48
G	15 - 18	U	49 - 51,11
H	19 - 21	V	51,12 - 52,12
I	22 - 23	W	52,13 - 55
J	24 - 27	X	56 - 57
K	28 - 30	Y	58 - 60
L	31 - 33	Z	61 - 63,6
M	34 - 37,35	A <sub>1</sub>	63,7 - 64
N	40	B <sub>1</sub>	65
		C <sub>1</sub>	66

### ANNEXE III

#### DONNEES MORPHOSTYLISTIQUES D'ASA KASHER

#### ANALYSEES PAR L'ANALYSE FACTORIELLE DES CORRESPONDANCES

En complément de notre étude et comme première étape critique de la méthodologie nouvelle mise en oeuvre par nous, il nous a semblé intéressant de soumettre au même type de méthode, qui était en cours de test, les données réunies par Asa Kasher dans son article sur l'"Hétérogénéité d'auteurs dans le livre d'Isaïe" (41). La division des colonnes (chapitres) était relativement différente de celle que nous avons observée dans notre étude; l'auteur proposait 29 fragments dont il garantissait l'unité logique, division que nous avons conservée dans l'intérêt de la recherche pour le découpage nouveau qu'elle apportait aux données à traiter. L'auteur fournissait dans le même temps, parmi d'autres, les occurrences des sept voies (42) fondamentales du verbe hébreu, auxquelles il ajoutait une voie "fourre-tout" pour les formes verbales difficiles à classer. Notre analyse a porté dans un premier temps sur cette partie des données à l'exclusion de toute autre.

A partir de ces données nous avons constitué un tableau de contingence de 29 x 8 et, faisant nôtre l'hypothèse de travail de l'auteur en ce que "*le style c'est l'homme*", nous avons soumis le tableau au même programme d'*Analyse Factorielle des Correspondances* que nous avons utilisé pour nos propres expériences.

## 1. L'HISTOGRAMME DES VALEURS PROPRES

Cet histogramme fait apparaître deux facteurs particulièrement importants :

- 1.1. Le premier axe porte à lui tout seul 31,8 % de l'information.
- 1.2. Le second axe porte encore 26,9 % de l'information.

LES VALEURS PROPRES VAL(1)=0,99999726

NUM ITER	VAL PROPRE	POURCENTI	CUMUL	HISTOGRAMME DES VALEURS PROPRES DE LA MATRICE
2	0,00187515	31,817	31,817	
3	0,00189511	26,930	58,748	
4	0,00349112	12,174	70,922	
5	0,00654710	11,476	82,398	
6	0,00718351	9,196	91,595	
7	0,0053354	5,892	97,487	
8	0,00195934	2,513	100,000	

LA VARIATION DU TEMPS= 3

Fig. 1

## 2. ANALYSE DES AXES FACTORIELS

Dans l'espace des *fragments* nous nous sommes fixé comme limite inférieure pour la contribution le produit  $\frac{1000}{29} = 35$  qui équivaut à la contribution moyenne des 29 fragments.

Dans l'espace des *voies du verbe hébreu* nous nous sommes fixé comme limite inférieure une limite semblable soit  $\frac{1000}{8} = 125$  qui équivaut aussi à la contribution moyenne des 8 voies du verbe.

2.1. *Axe n° 1*

2.1.1. Du côté des valeurs positives du 1er axe dans l'espace des *fragments* (Cf. figure 2), répondent à cette condition les fragments :

<i>FRAGMENTS</i>	<i>I # F</i>	<i>COR</i>	<i>CTR</i>
O	282	592	86
P	392	512	142
R	249	521	78
T	135	411	36
X	213	689	55

qui correspondent à nos chapitres 41; 42; 44; 46-48, d'une part, et 56-57, d'autre part.

2.1.2. Du côté des valeurs positives du 1er axe dans l'espace des *voies du verbe* (Cf. figure 3), la seule voie qui répond aux conditions posées est le *Hiph'il*.

<i>VOIES</i>	<i>I # F</i>	<i>COR</i>	<i>CTR</i>
HIPH	151	373	131

La forme *Hiph'il* semble marquer particulièrement le style de ces chapitres.

2.1.3. Du côté des valeurs négatives du 1er axe dans l'espace des *fragments* (Cf. figure 2), répondent aux conditions les fragments :



I I1	Q	L	T	P	O	I	D	I	N	R	1#F	COR	CTR	2#F	COR	CTR	3#F	COR	CTR	4#F	COR	CTR	5#F	COR	CTR
11A	888	28	14	-118	374	161	-125	423	221	-20	11	11	-52	75	91	14	5	11							
21B	885	25	26	14	2	01	129	208	201	27	9	21	209	546	1231	98	120	341							
31C	810	23	14	30	19	11	58	70	41	-58	71	81	158	518	641	80	132	201							
41D	884	46	101	90	463	151	20	23	11	54	168	141	57	186	171	-27	44	51							
51E	534	48	251	66	109	81	-56	83	81	52	67	141	-53	72	151	-89	203	541							
61F	962	38	551	-199	359	621	-166	251	511	0	0	01	-107	105	501	166	247	1471							
71G	971	31	571	-296	614	1101	-199	280	591	97	66	311	31	7	31	25	4	31							
81H	988	37	291	-82	112	101	62	62	71	13	3	11	211	724	1861	73	87	281							
91I	920	33	361	-175	369	421	-43	23	31	-110	147	431	32	12	41	-175	369	1441							
101J	981	56	781	-288	660	1631	-15	2	11	-94	83	531	-148	204	1391	-58	32	271							
111K	945	71	271	61	126	111	-92	291	291	56	104	231	-1	0	01	112	424	1241							
121L	997	41	241	-147	483	361	-70	113	101	2	0	01	119	313	641	-62	88	221							
131M	943	36	141	-12	5	01	160	849	441	27	25	31	42	59	71	12	5	11							
141N	946	26	521	-134	116	191	326	673	1301	-151	146	631	-25	4	21	34	7	41							
151O	978	27	461	282	592	861	-153	177	301	-46	16	61	96	69	281	-128	124	621							
161P	961	23	881	392	512	1421	-318	340	1111	-94	30	221	-153	79	611	11	0	01							
171Q	850	22	361	167	222	251	-125	126	171	-231	426	1251	-30	8	21	93	68	271							
181R	962	31	481	249	521	781	-109	101	181	194	317	1241	-6	0	01	-52	23	121							
191S	915	22	481	80	38	61	229	305	541	222	285	1121	-167	163	681	146	124	641							
201T	924	49	281	135	411	361	62	86	91	-134	409	951	-1	0	01	-27	18	51							
211U	589	48	31	-21	94	11	5	6	01	13	34	11	-46	424	121	13	31	11							
221V	891	21	421	-120	95	131	102	68	111	239	368	1281	-48	16	61	-230	344	1581							
231W	947	43	281	-142	406	351	-153	467	481	2	0	01	61	73	181	9	1	01							
241X	983	30	251	213	689	551	-49	38	41	-115	206	431	49	37	81	-28	13	31							
251Y	977	56	831	85	62	181	309	829	2571	-9	1	11	-95	79	581	-26	6	61							
261Z	363	25	251	41	21	21	98	124	111	68	59	121	-81	86	181	75	73	191							
271A1	729	19	101	100	256	81	91	211	81	-75	148	121	-35	34	31	56	80	81							
281B1	840	23	151	-33	22	11	170	552	321	-14	4	11	116	257	351	-16	5	11							
291C1	830	21	141	69	91	41	-63	79	41	168	548	641	-2	0	01	-75	112	171							
I	I	3617	0	10001			10001			10001			10001			10001			10001						10001

Fig. 2

I J1	Q	L	T	P	O	I	D	I	N	R	1#F	COR	CTR	2#F	COR	CTR	3#F	COR	CTR	4#F	COR	CTR	5#F	COR	CTR
11KAL	973	647	411	22	103	131	-16	60	91	27	148	491	50	517	1831	-26	145	641							
21NIPH	931	73	1181	-238	454	1691	103	85	371	10	1	11	79	50	521	207	341	4381							
31PIEL	991	91	1971	-41	10	61	372	825	6031	-114	79	1281	-107	69	1181	-36	8	181							
41RUAL	874	13	2131	-968	750	5021	-377	114	911	-43	2	31	-188	28	531	-316	80	1861							
51HIPH	990	141	1521	151	273	1311	-170	349	1871	-132	211	2641	-104	132	1751	46	25	421							
61HOPH	710	8	831	-391	190	501	-312	121	371	177	39	271	-262	86	621	470	274	2471							
71ITP	948	13	1191	407	227	851	187	48	211	586	473	4611	-380	200	2061	13	0	01							
81AUT	536	13	781	-290	182	441	-85	16	51	224	107	681	-322	224	1511	-57	7	61							
I	I	3617	0	10001			10001			10001			10001			10001			10001						10001

Fig. 3

<i>FRAGMENTS</i>	<i>1 # F</i>	<i>COR</i>	<i>CTR</i>
F	- 199	359	62
G	- 296	614	110
I	- 175	369	42
J	- 268	660	163
L	- 147	483	36

qui correspondent à nos chapitres 13-14; 15-18; 22-23; 24-27; 31-33.

2.1.4. Du côté des valeurs négatives du 1er axe dans l'espace des *voies du verbe* (Cf. figure 3), celles qui répondent à l'attente sont :

<i>VOIES</i>	<i>1 # F</i>	<i>COR</i>	<i>CTR</i>
NIPH	- 238	454	169
PUAL	- 968	750	502

#### 2.1.5. *Conclusions de l'analyse du premier axe*

On voit sur cet axe une forte opposition entre *grosso modo* les chapitres 13 à 33 et les chapitres 41-48 et 56-57 d'une part et entre la voie factitive du verbe et les deux voies passives, simple et intensive.

#### 2.2. *Axe n° 2*

2.2.1. Du côté des valeurs positives du 2° axe dans l'espace des *fragments*

(Cf. figure 2), répondent aux conditions les fragments :

<i>FRAGMENTS</i>	<i>2 # F</i>	<i>COR</i>	<i>CTR</i>
M	160	849	44
N	326	673	130
S	229	305	54
Y	309	829	257

qui correspondent à nos chapitres 34-37, 5; 40; 45; 58-60.

2.2.2. Du côté des valeurs positives du second axe dans l'espace des *voies du verbe* (Cf. figure 3), la voie suivante répond aux conditions posées :

<i>VOIE</i>	<i>2 # F</i>	<i>COR</i>	<i>CTR</i>
PIEL	372	825	603

2.2.3. Du côté des valeurs négatives du second axe dans l'espace des *fragments* (Cf. figure 2), on trouve :

<i>FRAGMENTS</i>	<i>2 # F</i>	<i>COR</i>	<i>CTR</i>
F	- 166	251	51
G	- 199	280	59
P	- 318	340	111
W	- 153	467	48

qui correspondent à nos chapitres 13-14; 15-18; 42; 52, 13-55.

2.2.4. Du côté des valeurs négatives du second axe dans l'espace des *voies du verbe* (Cf. figure 3), on trouve :

<i>VOIES</i>	<i>2 # F</i>	<i>COR</i>	<i>CTR</i>
HIPH	- 170	349	197

#### 2.2.5. *Conclusions de l'analyse du second axe*

On note sur cet axe une opposition assez mal définie entre, d'une part, des chapitres considérés comme appartenant à la première partie du livre d'Isaïe, 34-37, 5 d'un côté et 13-14; 15-18 de l'autre, et d'autre part entre les chapitres 40; 45; 58-60 d'un côté et 42; 52, 13-55 d'autre part.

Par ailleurs, on trouve du côté positif la voie active intensive du verbe en opposition avec la voie active factitive du côté négatif.



AXE HORIZONTAL( 1)—AXE VERTICAL( 2)—TITRE: ~~RENE~~ DONNEES ASA KASHER —ANNEXE 2 TABLE 1— ~~RENE~~

LARGEUR= 1.37574 HAUTEUR= 0.75079 —NOMBRE DE POINTS= 8

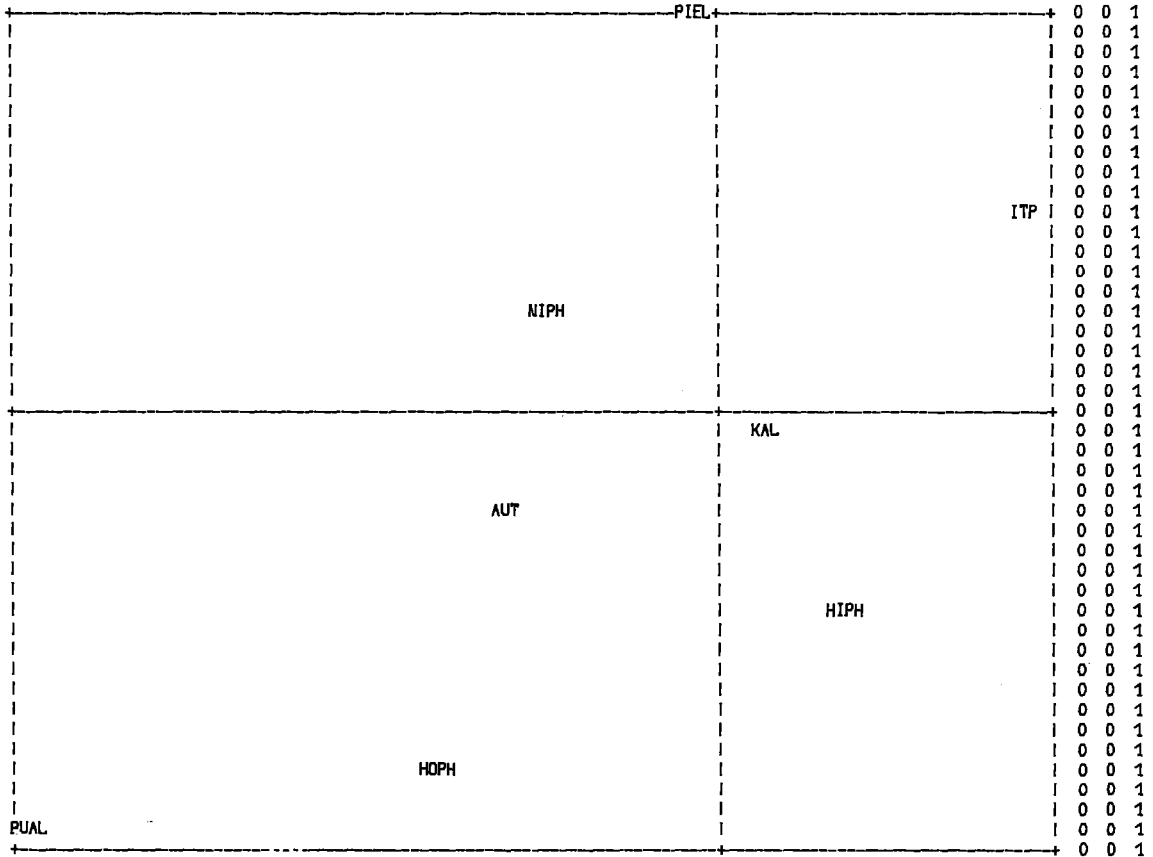


Fig. 5

### 3. PREMIERES CONCLUSIONS SUR LES DONNEES D'ASA KASHER

Les résultats de l'analyse du tableau 29 x 8 proposé par Asa Kasher, dont les données sont constituées par des éléments dits "*stylistiques*", ne semblent pas déterminants pour créer des ensembles intrinsèquement homogènes, dont l'hétérogénéité relative permettrait d'opposer de façon nettement différenciée de grands ensembles constitués. Il semble plutôt que cette analyse met en évidence dans le livre d'Isaïe l'opposition stylo-thématique entre les formations actives *Kal*, *Pi'el*, *Hiph'il* et *Itpa'el* et les parties rédigées à partir des formes passives *Nif'al*, *Hoph'al* et *Pu'al* qui entraînent avec elles les rares autres formes. Il ne semble pas pouvoir être tiré de ces constatations des instruments permettant de faire apparaître la disparité d'auteurs dans le livre, mais bien plutôt des faits de syntaxe ou de style liés à des contextes thématiques.

G.E. WEIL

A. SALEM

M. SERFATY

## NOTES

- (1) CH. V. LANGLOIS et CH. SEIGNOBOS, *Introduction aux études historiques*, Paris, 1898, p. 107.
- (2) E. RENAN, *L'avenir de la science*, Paris, 1898, p. XIV.
- (3) Cf. par exemple CH. MULLER, *Initiation à la statistique linguistique*, Paris, 1968.
- (4) P. GUIRAUD, *Problèmes et méthodes de la statistique linguistique*, Paris, p. 5.
- (5) G. E. WEIL, "La Massorah", *REJ*, CXXXI, 1972, pp. 44ss.
- (6) L'histoire des sociétés ne viendra jamais démentir une semblable démarche. Il est connu, et on voit encore des groupes humains, des peuples, des nations, des militants ou des adeptes se référer, sans aucun souci critique, à l'oeuvre scripturaire théologique, philosophique, politique ou scientifique d'un maître et de ses disciples proches.
- (7) *Talmud de Babylone* : Traité de *Baba' Batra'* 14b. Il est tout à fait raisonnable de penser que l'on ne peut pas plus assigner de *terminus ad quem* à ces sociétés savantes, si même leur dénomination tend à définir leur *terminus a quo*, que l'on ne peut assigner de *terminus ad quem* à la Compagnie ou à l'Académie fondée par Richelieu ou à la Petite Académie surajoutée par Colbert, dont les dénominations ont varié avec le temps et les situations historiques.



- (8) De fait, la division qui s'impose à première lecture est une tripartition de l'ouvrage : chp. 1 à 35 les premières prophéties; chp. 36 à 39 les parties historiques formant appendice dont la recension originelle semble être apportée par les chp. 2 Rois 18, 13 à 20, 19 et par le Psaume 116 qui, dans le Psautier, restitue le Cantique d'Ezéchias d'Is. 38, 9-20; enfin, la seconde partie formée par les chp. 40 à 66.
- (9) E. J. YOUNG, *Studies in Isaiah*, Londres, 1955; O. EISSFELDT, *The Old Testament, An Introduction*, translated by P. R. Ackroyd, Oxford, 1965, pp. 303-346.
- (10) LARRY L. ADAMS, *A statistical analysis of the book of Isaiah in relation to the Isaiah problem*. Thèse, Brigham Young University, Provo, USA; Y. T. RADDAY, *The Unity of Isaiah in the light of statistical linguistics*, collection *Massorah*, Hildesheim, 1973; A. KASHER, "The Book of Isaiah : Characterization of authors by morphological Data processing", *Revue*, n° 3, Liège, 1972.
- (11) La dépendance au sens statistique du terme découle nécessairement de la dépendance affirmée des éléments analysés dans le cadre d'un corpus textuel réalisé en situation d'énonciation et non des éléments d'un corpus créé *ad rem*.
- (12) Certains statisticiens, cependant, ont souhaité appliquer les méthodes statistiques aux petits nombres, cf. L. von Bortkewicz, *Das Gesetz der kleinen Zahlen*, 1898; G. U. YULE, *The statistical study of Literary vocabulary*, Cambridge, 1944.

- (13) J.-P. BENZECRI, professeur à l'Université de Paris VI, *L'Analyse des Données*, vol. I : *La Taxinomie*; vol. II : *L'Analyse des Correspondances*, Paris, 1973.
- (14) Cf. p. 14.
- (15) A. PROST, *Vocabulaire des proclamations électorales de 1881*, Paris, 1974, et BENZECRI, *L'Analyse des données. o.c.*
- (16) Le code utilisé pour la saisie des données en hébreu est reproduit en Annexe I.
- (17) Opération ayant pour objet de rapporter de façon approximative la grandeur d'un phénomène inconnu en la reliant à la grandeur certaine d'un phénomène connu. Dans le cas de notre exemple, nous sommes passés de façon approximative de deux grandeurs connues, nombre d'occurrences du mot %NY et %R3 respectivement dans chacun des chapitres, à un facteur qui est une combinaison linéaire des deux.
- (18) Le programme d'analyse factorielle des correspondances que nous avons utilisé dans le cadre de cet article est dû à N. TABET, l'un des collaborateurs de J.-P. BENZECRI.
- (19) Les chiffres 1, 2, 3, etc. qui figurent dans les tableaux avant le symbole # F indiquent le numéro de l'axe pour lequel sont données les trois valeurs.

- (20) Les chapitres : 1, (+ 52); 26, (+ 102); 32, (+ 61); 33, (+ 246); 35, (+ 319) ont, seuls, un premier facteur positif.
- (21) Seuls ont un premier facteur négatif les chapitres : 56, (-2); 60, (-47); 66, (-3).
- (22) Chaque fois où il est fait mention des valeurs positives ou négatives d'un axe, il faut entendre les valeurs positives ou négatives du facteur correspondant à cet axe.
- (23) Pour la clarté de l'exposé, on notera seulement que la figure de la projection des chapitres s'est inversée après avoir tourné de 90° à droite, l'axe 1 étant devenu dans cette figure l'axe horizontal. Cette inversion n'a aucune importance en elle-même, elle est due au programme qui a cherché à donner toutes les informations dans une projection occupant, en surface, la place la plus adéquate possible dans les dimensions les plus réduites possibles.
- (24) L'explication de l'écart remarqué du chapitre 49 par rapport au groupe des chapitres 40 à 50 nécessitant un développement par lui-même, celle-ci a été donnée en *Annexe II*.
- (25) Plus exactement on supprime leur influence sur les résultats de l'analyse en leur donnant une masse évanescence. Cf. 5.2, p. 54.
- (26) Cf. notes 20 & 21.

- (27) La division capitulaire d'Isaïe en 66 chapitres nous a contraint à constituer le 13<sup>o</sup> fragment avec les 6 chapitres restants.
- (28) Les trois fragments cités ne constituent pas un groupe homogène puisque le fragment 9 s'oppose très fortement aux fragments 4 et 5 sur le premier axe.
- (29) Cf. pp. 14-15.
- (30) De fait, en réalité on se trouve placé dans le cas de 13 fragments entre un maximum représenté par le fragment 8 (chp. 36 à 40) contenant 817 mots et un minimum représenté par le fragment 4 (chp. 16 à 20) contenant 305 mots dont la moyenne observée est de 560 mots, peu éloignée de la moyenne théorique.
- (31) Cf. pp. 19-20.
- (32) Le chapitre 40 ayant presque toujours été entraîné par les découpages avec le groupe des chapitres 36 à 39.
- (33) BENZECRI, *o.c.*, T. I DMCL. p. 119-152; CAH. p. 153-206; etc.
- (34) Le programme que nous avons utilisé pour cette analyse est dû à MICHEL JAMBU, collaborateur de J.-P. BENZECRI, et a été implanté dans la Bibliothèque des Programmes du Centre Inter-Régional de Calcul Electronique du CNRS à Orsay.

- (35) Il faut noter ici qu'en fonction des différents découpages qui ont été arbitrairement opérés, le chapitre 40 s'est toujours trouvé lié au groupe des chapitres 36 à 39 bien qu'à la simple lecture il s'en sépare fondamentalement. Ayant admis *ne varietur* un découpage arbitraire, nous n'avons jamais séparé volontairement ce chapitre des autres chapitres composant les fragments.
- (36) Le critère d'agrégation retenu dans le programme qui a fourni les résultats exposés est le minimum de variance entre deux fragments. Ce critère est utilisé pour combiner les fragments deux à deux jusqu'à épuisement et achèvement de la classification. Ainsi le programme recalcule à chaque pas la nouvelle variance pour la classe ainsi créée et tous les fragments subsistants, et agrège les deux éléments présentant à nouveau la variance interne la plus faible.
- (37) Cf. pp. 34-35
- (38) A. KASHER, "The book of Isaiah : Characterization of authors by morphological data processing." in *Revue*, n° 3, Liège, 1972.
- (39) Cf. Annexe III.
- (40) Cf. p. 34.
- (41) *Revue*, n° 3, Liège, 1972.
- (42) Nous rendons par ce terme, le mot hébreu "BINYAN".