



CLASSIQUES
GARNIER

TOLLENAERE (F. de), MULLER (Ch.), DUGAST (D.), CANDEL (D.), « Chroniques et comptes rendus », *Cahiers de lexicologie*, n° 48, 1986 – 1, p. 117-146

DOI : [10.15122/isbn.978-2-8124-4301-5.p.0119](https://doi.org/10.15122/isbn.978-2-8124-4301-5.p.0119)

La diffusion ou la divulgation de ce document et de son contenu via Internet ou tout autre moyen de communication ne sont pas autorisées hormis dans un cadre privé.

© 2012. Classiques Garnier, Paris.
Reproduction et traduction, même partielles, interdites.
Tous droits réservés pour tous les pays.

CHRONIQUES ET COMPTES RENDUS

CIVICIMA

COMITÉ INTERNATIONAL DU VOCABULAIRE DES INSTITUTIONS ET DE LA COMMUNICATION INTELLECTUELLES AU MOYEN AGE

Le Comité International du Vocabulaire des Institutions et de la Communication Intellectuelles au Moyen Age (CIVICIMA) a été créé lors d'un 'workshop' sur le thème de la «Terminologie de la vie intellectuelle au moyen âge», qui s'est tenu aux Pays-Bas, les 20 et 21 septembre 1985.

Le CIVICIMA ne vise qu'une partie du vocabulaire de la vie intellectuelle, à savoir les points de convergence des diverses disciplines, non pas leur contenu, mais le système du travail intellectuel.

Les membres du CIVICIMA présents aux Pays-Bas se sont mis d'accord sur le programme de travail que nous présentons ici, tout en précisant que le schéma ne se veut ni définitif ni exhaustif : il s'agit d'un programme ouvert mentionnant simplement les premiers terrains de recherche que nous nous proposons d'aborder.

L'objectif principal est la publication d'études interdisciplinaires sur les différents éléments du programme, qui devront bien entendu être différenciés géographiquement et chronologiquement. Ces études pourront être publiées ou non, selon les vues des auteurs, et sous la forme qui leur paraîtra la plus indiquée : livres, articles, microfiches, etc.

De plus, nous avons adopté la suggestion de constituer un fichier centralisé des termes ayant trait aux terrains déterminés, à établir selon des normes uniformes et soumis à la limite chronologique de 1520. Le traitement des fiches informatisées se fera à l'Université de Birmingham, le fichier traditionnel aura peut-être sa place au bureau du Comité Du Cange à Paris.

Le CIVICIMA étant une initiative néerlandaise, la présidence en a été confiée au professeur L.M. de RIJK, de l'Université de Leyde, et le secrétariat à Olga WEIJERS, du Lexique du latin médiéval des Pays-Bas. Les différents pays de l'Europe occidentale, ainsi que les Etats-Unis et le Canada, sont représentés par des membres. Pour coordonner et pour stimuler les recherches envisagées, le Comité a l'intention de se réunir périodiquement.

Nous demandons à tous les chercheurs de nous mettre au courant d'éventuelles recherches analogues en cours ou prévues.

Secrétariat du CIVICIMA : Olga Weijers, Bibliothèque Royale, Prins Willem Alexanderhof 5, 2595 BE Den Haag, Pays-Bas.

Programme de travail

1. Vocabulaire des écoles (institutions, enseignants, élèves, méthodes).

- IX - XIe siècles. L'école de type carolingien.
- XIIe siècle.
- XIII - XIVe siècles (écoles urbaines et studia des mendiants).

2. Vocabulaire des universités (institutions, personnes, méthodes d'enseignement).

- XIIIe siècle. L'espace européen paraît conserver une certaine unité qui permet une étude d'ensemble.
- XIV - XVe siècles, où les diversités régionales ou nationales devront être prises en compte (France, Angleterre, Italie, Empire, Péninsule ibérique).
Ad 1 et 2 : Dans le domaine de l'enseignement, il faudra traiter des termes concernant les méthodes d'enseignement, en laissant de côté le vocabulaire technique propre aux diverses disciplines.
 Sur le terrain institutionnel, il faudra incorporer les aspects socio-culturels, comme par exemple le vocabulaire relatif à la vie des étudiants (notamment dans les collèges).

3. Vocabulaire du livre et de l'écriture.

- le livre (fabrication, composition, édition, circulation).
- autres formes de documents écrits (chartes, tablettes, fiches).
- écriture (matériaux, styles d'écriture, copistes).
- le texte et sa tradition (copie, faute, correction).
- les bibliothèques et les archives (cadre matériel, classement, catalogues).

4. Vocabulaire des méthodes, instruments et produits du travail intellectuel.

- alphabétisation, annotation, traduction, information, communication.
- concordances, tables, glossaires, encyclopédies.
- traités, commentaires, sommes, compendia, etc.

5. Les appellations des disciplines et de leurs étudiants.

- disciplines et sciences.
- personnes s'y consacrant (artista, decretista, physicus, etc.).

Réalisations.

ad 2 : une étude d'Olga Weijers sur la *Terminologie des universités au XIIIe siècle* paraîtra en 1986 dans la collection du *Lessico Intellettuale Europeo*.

Fr. B R O M E

Histoire du mot

Selon O. BLOCH et Walther von WARTBURG, *Dictionnaire étymologique de la langue française*, 6e éd. (1975), le mot *brome*, datant de 1826, aurait été emprunté du grec *brômos* 'puant' par le chimiste BALARD, qui découvrit ce corps.

Il est en effet exact que la découverte de cet élément chimique fut faite par A.-J. BALARD dans le laboratoire de M. ANGLADA. Seulement ce n'est point le mérite de BALARD d'avoir proposé le nom *brome* pour la nouvelle substance.

Quand on se donne la peine de lire le mémoire de BALARD, pharmacien et préparateur de chimie à la Faculté des Sciences à Montpellier, «Sur une Substance particulière contenue dans l'eau de la mer», notamment le paragraphe II, où il traite «De la Dénomination de la Substance rouge, retirée de l'eau mère des salins, après l'action du chlore», on y lit : «M. Anglada me conseilla d'appeler cette substance *Brôme*, en déduisant cette dénomination du grec *βρωμος* (*foetor*). Ce nom se prête à merveille à la formation des dénominations composées que nécessitent ses combinaisons, et je l'adopte pour la facilité du langage» (*Annales de Chimie et de Physique* par MM. GAY-LUSSAC et ARAGO, Tome 32, p. 341, Paris, 1826).

Mais dans le même tome 32, à la page 382 et suivantes, on trouve le «Rapport sur le Mémoire de M. Balard relatif à une nouvelle Substance (Extrait du procès-verbal du lundi 14 août 1826)» par VAUQUELIN, THENARD et GAY-LUSSAC «chargés par l'Académie ... de lui faire connaître notre opinion sur un Mémoire de M. Balard ..., ayant pour objet la description des propriétés d'une nouvelle substance qu'il a trouvée dans les eaux de la mer». Le Rapport continue : «M. Balard a donné à la nouvelle substance le nom de *muride* ; mais plusieurs objections pouvant être faites contre cette dénomination, nous l'avons remplacée, avec le consentement de l'auteur, par celle de BROME, de *βρωμος*, mauvaise odeur».

Dans le tome 33 (Paris, 1826) des *Annales*, la correction de la commission de l'Académie, quant à la paternité du nom de *brome*, se trouve corroborée par une «Lettre de M. Anglada aux Rédacteurs des *Annales de Chimie et de Physique*», dont je cite le premier paragraphe :

«Messieurs,

En parcourant le Mémoire de M. Balard que vous avez inséré dans le dernier Cahier de vos *Annales*, je n'ai pas été peu surpris d'apprendre

que c'est moi qui avais donné le nom de *brôme* à la nouvelle substance à laquelle ce Mémoire est consacré. Cette assertion est évidemment une inexactitude dont une distraction de l'Editeur aura été cause, pendant le voyage que vous avez fait tous deux dans les Départemens. Il m'importe qu'elle soit relevée, et vous en fais expressément la demande. Cette rectification se fera d'abord dans les intérêts de la vérité. L'on sait, en effet, que la substance dont il s'agit était sortie de mon laboratoire, et avait été présentée à l'Académie avec le nom de *muride* que j'avais cru devoir proposer».

En 1825 Carl LÖWIG, un étudiant qui venait d'entrer au laboratoire de chimie à l'Université d'Heidelberg, suscita l'intérêt de son professeur Leopold GMELIN. LÖWIG apporta en effet de chez lui à Kreuznach un liquide rouge qu'il avait préparé par l'action du chlore sur l'eau d'une source saline. Le professeur GMELIN demanda à LÖWIG de consacrer une étude à cette trouvaille, mais entretemps le mémoire de BALARD parut dans les *Annales de Chimie et de Physique*.¹ Par la suite le mot français *brome* fit carrière en allemand sous la forme *Brom* (Carl LÖWIG, *Das Brom und seine chemischen Verhältnisse* (1829), en néerlandais comme *broom* (1857).

F. de TOLLENAERE
Pays-Bas

NOTE

1. Mary Elvira WEEKS and Prof. Harry M. LEICESTER, *Discovery of the Elements* 719 (1967).

Avec mes remerciements à Madame H. A. BOSMAN-JELGERSMA, professeur de l'histoire de la pharmacie à l'Université de Leyde, ainsi qu'à Monsieur J. W. van SPRONSEN, docteur en chimie à La Haye.

A PROPOS DE LA SPÉCIALISATION LEXICALE

Dans son article «Sur la spécialisation lexicale»,¹ M. D. DUGAST formule un certain nombre d'observations qui lui ont été inspirées par une «(re)lecture» de mon ouvrage sur le vocabulaire de CORNEILLE.²

Je m'accorde avec lui pour estimer qu'un index (ou une concordance) doit se prêter à d'autres exploitations que celles pour lesquelles il a été établi, et je suis le premier à souhaiter que le recensement des 532 800 mots du théâtre de CORNEILLE (32 pièces), qui occupe une centaine de pages de mon livre, puisse encore fournir à d'autres chercheurs des données utiles et des exemples de répartitions lexicales.³ Raison de plus pour y pourchasser les erreurs qui ont pu s'y glisser. M. D.D. en a relevé quelques-unes (pas toutes, je le crains), et cela me conduit à donner ici un bref erratum ; que les utilisateurs de cet index veuillent bien en tenir compte pour les pages 281, 293, 294, 306 et 318 :

avorter	: col. E :	lire 1 (et non 11)
décider	: col. D :	14 (et non 16)
	col. E :	2 (et non 12)
déloyal	: col. B :	6 (et non 1)
exécuter	: col. B :	24 (et non 27)

Ce sont là des erreurs d'impression, que l'on évite aujourd'hui en reproduisant directement les chiffres sortis de l'ordinateur. Reste le cas des vocables *interpréter* et *interprète* ; on notera que ces deux entrées ne sont pas à leur place alphabétique ; or les colonnes F (fréquence) et R (répartition), issues d'un autre listage, ont un ordre différent ; les données sont donc simplement interverties ; on lira :

	F	R	
interdire	10	8	
interdit adj.	8	8	
interprète	18	11	
interpréter	1	1	<i>Hor. 223</i>

Je confirme également que p. 135, dernière ligne, il faut bien remplacer 241 par 238 (note 4 de M. D.D.).

Dans sa note 16, M. D.D. conteste ma décision de compter dans le vocabulaire caractéristique d'une des 32 pièces les vocables de fréquence 1 qui y figurent (hapax). D'un point de vue strictement probabiliste, cela peut se défendre ;

mais je considère que c'est une donnée utile, qui devient valable si l'utilisateur la regroupe avec d'autres. En voici des exemples, que l'on multiplierait sans peine.

Dans les hapax de *Théodore*, je note le mot (rare) de *macule* ; réuni à *impollu*, à *impudicité*, à *prostituer*, qui figurent dans la même liste, il illustre un des thèmes de la pièce (qui se manifeste également, par antonymie, avec *vierge*, *virginité* et *sainteté*, et que l'on retrouve dans les vocables «statistique-ment caractéristiques») ; mais il dénote aussi, avec *impollu*, *chatouillement*, etc., d'un goût nouveau de l'auteur pour l'emploi d'un vocabulaire peu courant.

Le *Menteur* et sa *Suite*, on le sait, marquent une étape intéressante dans l'oeuvre théâtrale de CORNEILLE. On est frappé par le nombre exceptionnel des hapax dans chacune de ces deux comédies : 98 et 92, nombres qui ne sont dépassés que par *Médée*. Sont-ils tous utiles ? J'en extrais l'adverbe *secrètement* (*Menteur*, v. 744) ; du point de vue thématique, il peut rejoindre *secret*, n. masc., dont la fréquence est significativement élevée ; mais c'est sans grand intérêt. Mais la même liste nous offre *impudemment*, *naturellement*, et, dans la *Suite*, *agréablement*, *naïvement*, *passablement*, *ponctuellement*, *volontairement* ; un coup d'oeil sur le vocabulaire caractéristique des autres pièces, et surtout des tragédies qui précèdent ou suivent, suffit pour suggérer qu'il y a là une des marques du style des comédies, moins contraint, plus libre, plus proche de la prose. Une piste s'ouvre ; suivons-la. On constate sans peine que sur les 34 adverbes en *-ment* qui figurent dans une seule des pièces de notre corpus, les comédies (24 % de l'ensemble) en ont 15 ; les tragédies qui en ont le plus (*Nicomède*, 5 ; *Agésilas*, 3) sont justement celles dont le lexique se rapproche le plus, par divers autres caractères, de celui des comédies. Si on considère tous les adverbes de ce type, quelle que soit leur fréquence, l'index en fournit 115 ; les occurrences sont au nombre de 454 dans les tragédies, 349 dans les comédies (je néglige ici les «pièces diverses», alors qu'une répartition régulière donnerait 562 et 241). L'écart est donc très sensible, et permet de dire en toute vraisemblance que le style tragique évite ces mots. On peut étendre la recherche aux noms en *-ment* (de 3 syllabes au moins) : il s'en trouve 83, avec 555 occurrences dans le genre tragique, 348 dans le comique (théor. : 632 et 271). L'écart, un peu moins fort, reste très significatif, et permet de conclure que, comme les adverbes, ces mots sont moins admis dans l'une des séries, évidemment à cause de leur lourdeur.⁴ On voit que l'examen des hapax peut se révéler suggestif, et qu'ils peuvent contribuer efficacement à caractériser le style des parties d'un corpus.

Je dois relever enfin la note 22 et le texte de la p. 122 auquel elle renvoie. M. D.D. estime que, en matière de richesse lexicale, et jusqu'à son invention de l'échelle Uber, «on a piétiné quelque temps» ; par exemple «en forçant les textes sur un lit de Procuste (on coupait, on étirait...) pour leur donner la taille qui convenait aux comparaisons». Et, dans la note : «C'est l'approche de Muller. Elle relève d'une erreur de méthode fondamentale : sauf à admettre une homogénéité de bout en bout, amputation et étirement se font sur des critères parfaitement différents : la première ne s'exerce que sur une partie de texte indépendante, celle que l'on supprime ; le second à partir d'une synthèse du texte : étant donné ce qu'il a été jusque là, voici ce qu'il pourrait être sur une longueur plus grande».

Je ne sais si l'erreur de mon contradicteur est «fondamentale», mais elle est évidente.

« Mettons les choses au point, car il s'agit là d'un problème de méthode qui intéresse de nombreux chercheurs.

Etant donné un texte de N mots, dont on connaît l'échelle des fréquences (V_i , nombre de vocables de fréquence i ; V , nombre total des vocables du texte), il s'agit de construire un modèle pour une étendue N' différente de N ; cela pour comparer sa richesse lexicale et sa distribution avec celles d'un autre texte d'étendue N' . Pour reprendre la terminologie de M. D.D., on va soit « étirer » le texte, si $N' > N$, soit l'« amputer », si $N' < N$.

J'ai proposé, en 1964, un procédé qui, destiné surtout à comparer entre elles les parties d'un texte par référence à son ensemble, ne s'applique que pour $N' < N$. Depuis, il a été pratiqué par de nombreux auteurs, tant français qu'étrangers, et parfois avec des variantes intéressantes. Et il a été soumis à des vérifications expérimentales et à des simulations (tirages aléatoires sur ordinateur), qui ont fourni des confirmations d'une précision que je n'aurais pas espérée.⁵

Mais je n'ai jamais pratiqué l'« étirement », et j'ai douté longtemps qu'il fût possible d'en construire un modèle. Première erreur.

Ensuite il est faux de dire que, quand on applique la loi binomiale pour calculer ce que serait un texte raccourci, on « supprime » une certaine partie de ce texte, en l'« amputant » de sa fin, ou de son début, ou de tout autre fragment. Les estimations obtenues sont des « espérances mathématiques », $E(V')$ et $E(V_i')$, c'est-à-dire des moyennes de tous les tirages possibles de $(N - N')$ mots ; évidemment elles reposent sur une « hypothèse nulle » qui suppose l'homogénéité du texte de départ ; ce qui ne signifie en rien que l'opération n'est valable que si cette homogénéité est réelle ; bien au contraire, cette opération sert la plupart du temps à faire apparaître et à mesurer des écarts stylistiques entre les parties du texte ou du corpus étudié. C'est enfoncer des portes béantes que de déclarer que l'homogénéité lexicale n'est jamais réalisée dans un texte réel, et surtout dans une oeuvre littéraire. Chacun sait que c'est une référence idéale, mais nécessaire pour mesurer les variations lexicales au fil du texte.

M. DUGAST a parfaitement le droit de préférer sa propre méthode pour l'étude quantitative des faits de lexique. Et d'autres que lui travaillent dans la même voie, où il reste certainement beaucoup à faire.⁶ Mais cela ne l'autorise pas à donner des appréciations hâtives et une présentation inexacte d'autres approches que la sienne.

Ch. MULLER
Strasbourg

NOTES

1. *Cah. de Lexicol.*, XLVI (1985) 1, pp. 109-131.
2. *Le vocabulaire du théâtre de P. Corneille. Etude de statistique lexicale*, 1967. Réimp. en 1979 aux éd. Slatkine-Champion, Genève-Paris.
3. L'index (avec annexes) donne les fréquences et sous-fréquences des 5347 vocables (plus ceux des variantes) ; il localise (pièce, n° du vers) environ 2700 occurrences

(fréq. 1, 2 et 3), et situe (pièce) les vocables de fréquence 4 et les noms propres. Cet index a pris une valeur supplémentaire par la publication de la thèse de Ch. BERNET, *Le vocabulaire des tragédies de Racine* (*ibid.*, 1983), qui met en parallèle les données quantitatives de ces deux oeuvres.

4. Tragédies et comédies représentent 0,7 et 0,3 de l'ensemble qu'elles forment (pièces diverses exclues). Pour les adverbes, le X^2 dépasse 60 ; pour les noms, 30. Il semble donc que l'adverbe soit évité, dans le style le plus soigné, à la fois pour sa forme et pour son rôle syntaxique.
5. Voir N. MÈNARD, *Mesure de la richesse lexicale. Théorie et vérifications expérimentales* (*ibid.* 1983), et M. DUBROCARD, «Evaluation de l'étendue du lexique. Essai de simulation», in *TraLiLi XXI*, (1983), pp. 235-246 (en particulier, le tableau des pp. 244-245). On trouvera de nombreuses références dans une petite bibliographie de 1981 qui est reproduite p. 83 de mon recueil d'articles, *Langue française, linguistique quantitative, informatique* (*ibid.* 1985).
6. Voir B. DOLPHIN, *Vocabulaire et lexique* (*ibid.*, 1983). Parmi les recherches en cours, celles de Ph. THOIRON (Lyon), et de deux mathématiciens, D. NAJOCK (Berlin) et D. SERANT (Lyon).

Centre : Informatique et Bible, Abbaye bénédictine de Maredsous, Brepols publishers, 1981, 180 pages.

Ce ne serait pas jouer sur les mots de dire que le volume intitulé *Centre : Informatique et Bible* qui présente le travail effectué à l'abbaye bénédictine de Maredsous à Denée en Belgique, est, pour qui travaille en linguistique quantitative et computationnelle, une véritable « bible » : on y trouve probablement, ramassés en quelque 180 pages, le maximum de renseignements rassemblés en une seule publication. On avait la *Bibliographie critique de la statistique linguistique* de Pierre GUIRAUD (Utrecht, 1954), le *Style and Stylistics : an Analytical Bibliography* de Louis-Tonko MILIC (New York, 1967) ; depuis 1981, on a *Centre : Informatique et Bible*.

Soit donc d'abord sa *bibliographie* : elle occupe les pages 87 à 155, comporte 591 titres appartenant à 427 auteurs, et couvre les années 1960 à 1980. Elle n'est pas complète, bien entendu, et les auteurs s'en expliquent (p. 87) : il fallait choisir parmi des milliers de titres, et tout choix comporte sa part de sacrifices. On pourrait être étonné d'une prolifération des travaux quantitatifs et computationnels en linguistique et en littérature, quand on sait la résistance opposée à cette nouvelle approche de la chose littéraire, résistance qui tient plus du manque de familiarité et de connaissance des possibilités offertes que du rejet pur et simple : le lecteur trouvera amplement ici de quoi étendre de telles connaissances.

La bibliographie comporte 19 sections ; la présentation est bilingue, et quasiment chaque section présente deux parties, l'une intitulée « application à la Bible », l'autre, qui est d'un intérêt plus général, intitulée « théorie-theoretical background ».

Prenons comme exemple les sections intitulées « analyse stylistique » (n° 9), « analyse statistique » (n° 11) et « lexicographie et lemmatisation » (n° 15). Après une présentation bilingue de la section, on trouve, respectivement : un titre (n° 209) et 22 titres (n° 210 à 231) pour les deux parties de la section « analyse stylistique » ; 18 titres (n° 246 à 263) et 24 titres (n° 264 à 287) pour la section « analyse statistique » ; enfin, 8 (n° 407 à 414) et 17 titres (n° 415 à 431) pour la section n° 15. C'est dire que la partie « théorie » est amplement développée par rapport à la partie « application à la Bible » ; c'est dire aussi la richesse des titres pour chacune de ces sections. On constatera sans doute que les titres sont essentiellement anglo-saxons, ce qui ne tient pas seulement à l'intérêt que l'on porte à la Bible outre-Atlantique, puisque justement la partie la plus nombreuse de cette bibliographie n'a pas trait à la Bible comme on vient de le voir. Beaucoup de titres rédigés en français sont d'ailleurs belges. Les études quantitatives en France, bien que Pierre GUIRAUD les y ait introduites voici trente ans déjà, y

restent encore peu développées : nous n'avons pas de revues telles que *Computers and the Humanities*, *CHum*, maintes fois citée dans la présente bibliographie, ou *Quantitative Linguistics*, titre qui ferait encore peu recette chez nous, nous le regretterons.

Cette importante bibliographie (en ce qui nous concerne, la partie la plus intéressante de l'ouvrage), est suivie d'un *index des auteurs* (pp. 156 à 163) qui contient donc 427 noms ; d'un *glossaire* (pp. 165 à 172), et d'un dossier trilingue d'*informations pratiques* (pp. 173 à 178) sur les services disponibles au centre.

On peut s'arrêter un instant au *glossaire*. Il contient 57 entrées parmi lesquelles nous relevons : «dictionnaire, forme, forme canonique ou lemme, homographe, lemmatisation, lemme et lemme composé, lexique, mot-clef, mot-outil, mot significatif, translittération...». On reparlera de la lemmatisation plus loin ; on peut s'interroger doublement sur la portée de l'appellation «mot significatif» qui nous est présenté seulement comme le «véhicule d'un contenu sémantique», et quand on sait que, dans la *Concordance de la Bible de Jérusalem* (Cerf, Brepols, 1982) on nous présente les listes des mots significatifs relatifs aux deux Testaments à partir d'une sélection basée sur l'écart entre les pourcentages que ces mots représentent dans chacun des deux corpus. Une telle procédure ne peut mener qu'à des erreurs absolument rédhitoires : un pourcentage non contrôlé statistiquement n'a jamais eu aucune signification. C. MULLER (*Initiation à la Statistique Lexicale*, Paris, 1968) l'a dit dans une formule quasi lapidaire : «les pourcentages, dont sont prodigues de bons linguistes, amis des statistiques, mais rebelles à la statistique» (p. 74).

De même, sommes-nous arrêté par la définition du «mot-outil», «plutôt choisi en fonction de sa fréquence d'utilisation» (p. 170). Si on a cru pouvoir placer quelques espoirs de définition de ce côté-là il y a encore une vingtaine d'années, on s'est vite rendu compte de son irréalité, d'où a d'ailleurs découlé la certitude maintenant admise que la fréquence n'a pas d'existence en langue, mais seulement en texte.

Le *glossaire*, qui donne une idée des intérêts des ouvriers de Maredsous, et est intéressant en ce sens, ne nous paraît pas être la partie la plus utile de l'ouvrage.

Beaucoup plus intéressant, la description des *textes bibliques* qui font partie de ce qu'on appellera le «trésor» de Maredsous, avec la présentation des textes (essentiellement, les sources utilisées), les modes d'accès aux fichiers, et peut-être surtout leurs *volumes*. On possède ainsi 12 tels corpus, avec le nombre de mots (le volume) pour chaque livre les composant. En voici la liste :

39 textes en hébreu totalisant 416 872 mots ; 4 en araméen avec 7 165 mots ; 10 en sirac (le livre de Ben Sira) avec 15 626 mots ; 4 en arabe (75 120 mots) ; 4 et 4 en syriaque (57 340 et 33 248 mots) ; 86 en grec (761 697 mots) ; 79 en latin pour la Vulgate (666 246 mots) ; 78 en anglais pour la Revised Standard Version (891 578 mots) ; et 4 corpus français : la Bible de Jérusalem (851 641 mots), celle de Maredsous (852 940 mots), celle de Segond (775 370 mots) et la T.O.B. -traduction oecuménique de la Bible (877 935 mots), les 3 premiers avec 73 textes chacun, le dernier avec 76.

On regrettera sans doute que l'on ne nous donne pas en même temps le nombre des mots différents, mais il ne fait aucun doute qu'ils soient disponibles : la *Concordance* (voir plus haut), avec un nombre de mots légèrement différent (851 400) en donne 12 866 ; on connaît aussi les comptes de MORGENTHAUER publiés en 1958 pour le texte grec du Nouveau Testament (utilisés dans notre thèse, *La Description Quantitative des Textes ; Essai de Stylistique Expérimentale*, 1981, à l'Université de Montpellier) qui donne 5 436 vocables pour 137 328 mots (Maredsous donne 138 013 mots pour le Nouveau Testament grec).

Les simples comparaisons des longueurs, entre les versions, i.e. selon les langues et les traductions dans une même langue, peuvent être source de renseignements sur les langues, mais aussi sur les qualités, lexicologiques et stylistiques, des traductions, voire, indirectement, sur les «cultures» qui les ont produites. Mais le nombre des mots différents ou vocables (ou lemmes) utilisés, et la relation de ces nombres aux longueurs dans lesquelles ils sont utilisés, relation qu'on s'accorde à appeler «richesse lexicale» ; plus encore sans doute, et similairement, la saisie de l'accroissement du vocabulaire au long des textes à comparer, ne peuvent qu'enrichir de telles comparaisons.

Voici, pour les 5 Rouleaux ou Megillot, souvent étudiés (voir les n° 260 et 374 de la bibliographie, entre autres), les longueurs données dans le volume pour 8 versions, à quoi nous nous permettons d'ajouter notre propre décompte effectué sur le texte de la *Version Autorisée* du Roi Jacques, texte anglais de 1611, avec le nombre des mots différents.

Ruth	: 1816	2072	1819	2453	2526	2350	2448	2441	2581	(441)
Cant.	: 1641	2025	1786	2523	2579	2600	2731	2605	2671	(562)
Eccl.	: 4168	4546	3813	5277	5503	5552	5970	5752	5572	(774)
Lam.	: 1931	2391	2389	3250	3185	3262	3432	3271	3425	(682)
Esther	: 4578	5843	6176	5430	(8230)	(7917)	5590	5337	5701	(697)

Hébr. Grec Vulg. RSV Jér. Mard. Seg. TOB A.V. (Voc)

Longueurs comparées des Cinq Rouleaux.

(Les longueurs entre parenthèses concernant Esther comportent le supplément grec absent du texte hébreu et des bibles protestantes).

Ajoutons que les machines de Maredsous sont équipées pour traiter des textes en caractères spéciaux, dont des exemples, malheureusement seulement esthétiques pour nous, sont donnés pour le grec (p. 83), l'hébreu (p. 84), l'arabe (p. 85) et le syriaque (p. 86). Ces exemples sont donnés en appendice à un article (traduction anglaise d'un original flamand) de Marc VERVENNE de l'Université de Louvain intitulé : «Computers and literature ; an application in the field of biblical sciences», pages 52 à 71, qu'il nous faut maintenant résumer.

Cet article comporte 4 sections, et une introduction. Les 4 sections concernent la banque des données, son utilisation, le traitement des caractères spéciaux et le matériel dont dispose le centre. Nous nous contenterons de parler de la banque des données, après un aperçu sur l'introduction.

Celle-ci, citant un texte de W. FUCHS dans *Reden am Beispiel Dürers*, München, 1972, donne immédiatement le ton en parlant «d'une science littéraire exacte, qui consiste à traiter la littérature d'une manière systématique en utilisant des outils mathématiques de la même manière que le font les sciences exactes». On y insiste sur l'importance que prend l'ordinateur dans les études linguistiques et littéraires en général, pour la critique littéraire et l'approche scientifique des textes, tant en ce qui concerne la recherche en grammaire qu'en philologie par exemple. On y cite les travaux qui se font tant aux USA qu'en Allemagne, à Amsterdam ou à Haïfa, en ce qui concerne la Bible, et, en France, à Nancy (où on connaît, entre autres, de B. BARC, une thèse de 3ème cycle de 1972 sur le texte hébreu de *Jérémie*). On y précise que la linguistique computationnelle n'est pas limitée au stockage des informations, et d'autre part qu'elle permet l'analyse exhaustive des textes. Le Centre trouve son origine dans un premier projet d'index multilingual qui date de 1958, et une première publication, en 1974, d'un ouvrage de 1214 pages contenant 9000 entrées se rapportant à 150 000 références bibliques.

La banque des données regarde 8 corpus de 42 à 73 livres et de quelque 850 000 mots chacun, totalisant environ 6 800 000 mots. On insiste justement, dans sa description, pour dire que les premières permutations du texte en langage machine déterminent toutes les applications consécutives, et qu'elles méritent donc le plus de soin et de réflexions. Puisqu'aucun lecteur optique ne peut lire directement le texte, il est nécessaire de dactylographier le corpus entier.¹ Un système de références par numéros renvoie non seulement aux versets, mais aussi aux «unités de sens» ; les accents (comme en français), les signes diacritiques (en grec), les accents musicaux (en hébreu) sont rendus également par des numéros ou des signes spéciaux : c'est dire la richesse de l'information stockée dans la machine.

On est particulièrement intéressé par l'approche pragmatique de la lemmatisation. On insiste peut-être un peu trop sur le côté conventionnel et arbitraire de la définition du lemme : on aimerait que des critères grammaticaux et lexicographiques à vocation taxinomique président davantage à cette définition ; mais on sait les difficultés que rencontre la seule conception d'une norme en ces domaines.

L'idée a été d'économiser le plus possible le temps et le travail de programmation. A cet effet, on a constitué (en ce qui concerne le français, pour prendre le meilleur exemple), un dictionnaire des formes du français à partir des quelque 850 000 mots du texte de la Bible de Maredsous, d'où sont sorties 30 000 formes correspondant à 14 000 lemmes, dictionnaire qui permet de lemmatiser automatiquement de 60 à 90 % d'un nouveau texte ; le choix des cas ambigus et des formes absentes dans le corpus de base doit être traité manuellement, avec toutefois l'aide d'un traitement statistique des possibilités des cas d'ambiguïté qui simplifie les vérifications subséquentes. On ajoute que, les langues modernes ayant été fortement influencées par le vocabulaire biblique, ce dictionnaire de base doit pouvoir être utilisé pour n'importe quel texte littéraire français. Nous en prenons acte, et souhaitons vivement la mise à l'épreuve répétée du dictionnaire : par exemple, dans quelle mesure un index, non lemmatisé comme on sait, du *Trésor de la Langue Française* de Nancy, (on peut penser à un Victor HUGO, qui ferait suite au PROUST traité par E. BRUNET récemment

pour les Editions Slatkine à Genève) pourrait être opposé au dictionnaire des formes et des lemmes de Maredsous à la satisfaction de l'analyste ? On sait l'importance que le vocabulaire biblique prend chez HUGO, et l'expérience pourrait être riche.

Centre : Informatique et Bible n'est pas un volume important par le nombre de ses pages ; il l'est certainement par l'information qu'il contient. Ce qui est le propre d'un ouvrage de référence. «Allah n'a pas permis qu'un seul livre fût parfait» a-t-on dit. Aussi n'est-ce pas diminuer les grands mérites de celui-ci que de souhaiter, comme nous l'avons fait, un peu plus de rigueur dans certaines définitions, ou un complément d'information à propos des corpus (il y a la place pour une colonne concernant les lemmes, données plus utiles au lecteur que les pourcentages de longueur qu'il peut calculer lui-même au besoin). On louera par contre les grandes possibilités offertes par le Centre, y compris sous un aspect pragmatique, comme nous avons dit à propos de la lemmatisation, ce qui relève d'un réel professionnalisme à quoi les travaux et critiques universitaires français ne nous ont pas toujours habitués. Ce choix des auteurs est aussi une ouverture à la critique ; mais il s'agit alors d'une critique éminemment constructive qui ne peut que participer à l'enrichissement mutuel des chercheurs.

Intéressé ou non par les études bibliques, tout chercheur engagé dans le travail computationnel appliqué à la linguistique et à la littérature aura le plus grand intérêt à avoir ce petit volume sur sa table de travail.

D. DUGAST
Paris

NOTE

1. Rappelons que le texte date de 1981.

Concordance de la Bible de Jérusalem réalisée à partir de la Banque de données bibliques de l'Abbaye de Maredsous. Cerf, Brepols, 1982, 1230 p.

Avec la première concordance française¹ complète d'une traduction catholique de la Bible -il s'agit du texte de la *Bible de Jérusalem* (Paris, 1956) et non de la traduction oecuménique en cours depuis 1967- le Centre d'Informatique Biblique de Maredsous (5642, Denée, Belgique) nous offre un remarquable travail lexicographique, et aussi lexicométrique, l'angle sous lequel nous nous proposons de l'examiner ici.

On sait que le texte biblique est un des textes, et probablement le premier, à avoir été l'objet de comptages d'unités, non seulement les mots, mais même les lettres, et cela dès avant la traduction grecque de la Septante (Septuagint) au III^{ème} siècle av. J.C. Il s'agissait tout simplement du respect et de la conservation du texte sacré, ce qui peut nous rappeler la description phonétique du sanscrit dans la grammaire de PANINI, dans le souci de protéger la prononciation d'autres textes sacrés. Si PANINI est l'ancêtre des phonéticiens, ceux qui ont compté page après page le texte de la Bible sont sans doute aussi les ancêtres des lexicométriciens.

On peut d'ailleurs se demander comment le texte hébraïque antérieur aux premiers Ptolémées a bien pu être soumis à une traduction, au nom justement du caractère sacré du Texte, quand on sait qu'il a fallu attendre 1977 pour qu'un autre texte sacré, le Coran, réputé intraduisible, reçoive de la plus haute autorité morale de l'Islam son imprimatur pour la traduction en français que nous a donnée La Pléiade en 1967 sous la signature d'un «D. MASSON» dont l'initiale du prénom a longtemps caché celui d'une femme.

Pour en revenir aux comptages des mots de la Bible, on peut ajouter ceci, qu'au XVIII^{ème} siècle, deux gentlemen anglais, l'un en 1712, et l'autre à Amsterdam en 1778, ont compté, pendant trois années chacun, le nombre des livres, des chapitres, des versets, des mots et des lettres de la Version Autorisée du Roi Jacques (1611). On peut donner ici le nombre des mots : 592 439 pour l'Ancien Testament (A.T.), 181 253 pour le Nouveau (N.T.), soit un total de 773 692.

Disons peut-être au passage que le découpage en versets auquel nous sommes accoutumés est dû à la manie de la clarification des scolastiques du XVI^{ème} siècle héritée de la passion des subdivisions (trinitaires si possible) de Thomas d'AQUIN, et à qui on doit tous les découpages des anciens textes de HOMÈRE et PLATON à VITRUVÉ. On dit que c'est ÉRASME qui aurait fait le découpage de la Bible pendant ses voyages. Mais celui-ci a été exécuté par le Rabbi NATHAN, et, en Angleterre, a été adopté par Robert STEPHENS pour le

Testament grec en 1551, et pour la Vulgate en 1555. C'est un découpage tout récent, et parfois arbitraire.

Maredsous nous donne 851 400 mots pour l'ensemble de la Bible, soit 667 742 pour l'A.T. et 183 658 pour le N.T. La différence avec les chiffres précédents ne tient pas seulement à la langue : on sait que le canon de l'A.V. (Authorized Version) n'est pas celui de la Bible catholique, pour l'A.T. principalement, où il manque les belles histoires de Tobie, de Judith, de Suzanne et quelques autres textes.

Maredsous nous donne aussi le nombre des vocables. Il y en a 12 866. Ce qui, pour les lecteurs qui utiliseraient l'échelle Uber de mesure de la richesse lexicale (voir par ex. *Cahiers de Lexicologie*, XXXVI, (1980) 1, «Qu'est-ce qu'un vocabulaire théorique ?», p. 5) met le texte biblique à 19,315, comparable au N.T. grec de MORGENTHALER (137 328 mots pour 5 436 vocables, soit 18,821), aux oeuvres de CHAUCER (417 564 ; 8 072 ; 18,435) et de GIRAUDOUX (671 364 ; 15 771 ; 20,842), mais le fait plus riche que CORNEILLE (532 800 ; 5 347 ; 16,409) et plus sobre que SHAKESPEARE (884 647 ; 31 534 ; 24,423), même si on tient compte des diversités de la lemmatisation. On peut rappeler qu'avec 32 1/2 millions de mots, le XIX^{ème} siècle du *T.L.F.* est à 19,907 ; avec 37 1/2 millions, le XX^{ème} est à 20,501 ; et avec plus de 70 millions de mots, les deux siècles sont à 20,571.

La première qualité que nous trouvons au travail de Maredsous est d'avoir justement lemmatisé le texte. C'est important, les usagers des dictionnaires et des index du *T.L.F.*, qui doivent lemmatiser eux-mêmes, le savent bien. Chaque entrée du dictionnaire est précédée de sa fréquence dans l'A.T. et le N.T. Les mots-outils (135, qui totalisent 517 808 emplois), dont les nombres d'occurrences descendent de 71 538 pour «le, la, les» à 41 pour «duquel», ne sont pas accompagnés de leurs références ni de leurs contextes, mais le Centre d'Informatique (adresse plus haut) est à même de fournir rapidement à qui les désirerait les listes des références, avec ou sans contextes.

Bien entendu, dès qu'on parle «lemmatisation», on entre dans un monde d'où l'arbitraire ne peut pas être absent. On peut discuter à l'infini sur l'intérêt qu'il y a à distinguer ou à ne pas distinguer «le, la, les» pronoms et articles, «que» relatif et conjonction, etc. Pourquoi rassembler «le, la, les», mais non «elle, elles» et/ou «il, ils» ? Fallait-il dissocier «au» et «du», et distinguer «des = de + les» de «des» pluriel de «un» ? Il faut admettre, actuellement du moins, que ceci est finalement assez secondaire. Il était autrement important de donner des unités sémantiques cohérentes (toutes les formes d'un même verbe sous la même entrée), et de distinguer les homographes (e.g. «conseiller» verbe et substantif), ce qui représente un effort à mettre au crédit des ouvriers de la Concordance.

Les sémanticiens apprécieront aussi le souci d'utiliser une graphie différente pour signaler, sous une même entrée, des distinctions du genre : «seule voie d'accès vers la Judée» et «il en prit un accès de fureur» : la dernière utilisation est en italique.

Au total, on ne peut que louer un travail aussi remarquable pour ses qualités à la fois au niveau de la conception (ce n'est pas toujours là qu'on trouve les

qualités les plus fréquentes) et de la précision du travail. Quand on apprend que la présente Concordance est le premier ouvrage publié du projet d'un outil multilingue pour lequel plus de 8 textes différents de la Bible sont déjà enregistrés (en hébreu, grec, latin, français, anglais, arabe, syriaque), on ne peut qu'attendre avec impatience la réalisation du projet qui ouvrira la porte à de passionnants travaux de comparaisons des langues, des stylistiques et des cultures par exemple, par delà les distances dans le temps et dans la géographie.

Nous n'avons aucunement l'intention de déprécier tant soit peu la qualité d'ensemble de ce travail en signalant un objet de déception. Le volume se termine par 5 tables statistiques dans lesquelles on trouve les 75 premiers mots par ordre décroissant des fréquences (I) ; ceux qui appartiennent en propre à l'un des Testaments (II) ; les comparaisons écartant les mots grammaticaux (IV). 75 mots n'est pas énorme, certes, mais comment tout mettre ?

Par contre, on est arrêté par les tableaux III et V, qui donnent les mots présentant un *écart significatif* dans un Testament par rapport à l'autre. Voilà qui est tout de suite intéressant. Malheureusement, en fait d'écart significatif, on a utilisé la différence entre les pourcentages que font les occurrences du mot dans chacun des deux Testaments. C'est une déception.

Il faudrait faire tout un historique des *écarts significatifs* destinés à préciser un vocabulaire caractéristique ou distinctif, ou à tracer un profil lexical ; à dégager des mots-clés ou des mots-thèmes. On a utilisé les 50 fréquences les plus importantes d'un texte ; on a comparé les fréquences dans un texte à celles d'une liste-type. Le problème n'a pas de solution idéale. On convient actuellement de repousser toute liste-type externe au corpus, et de considérer le corpus comme un tout à l'intérieur duquel les occurrences des vocables lui appartenant seraient, a priori, réparties entre les parties le composant selon les lois du hasard. Si on met en évidence des entorses graves à ces lois-là, on accepte de parler de vocables caractéristiques ou significatifs.

Ce n'est pas ici le lieu de tenter de faire le point d'une question très débattue. De quel droit décide-t-on que les vocables se répartiraient au hasard ? Pourquoi le degré de caractérisation d'un vocable serait-il fonction uniquement du nombre de ses occurrences ? D'autres critères, à définir, comme le poids sémantique du vocable, qui, à fréquence basse, le destine en quelque sorte à entrer dans telle partie du corpus et non dans telle autre, ne sont-ils pas au moins aussi importants ? Mais comment, actuellement, mesurer cela ?

On ne peut donc, actuellement, s'en remettre qu'aux fréquences ; encore convient-il de contrôler la signification des nombres. On distingue à cet effet entre les hautes et les basses fréquences. On appelle «haute fréquence» toute fréquence d'un vocable assez forte pour que, dans une répartition régulière hypothétique, la plus petite partie du corpus considérée en reçoive au moins 5. La plus petite partie, ici le N.T., faisant 21,57 % de l'ensemble, seront considérées comme «hautes fréquences» les fréquences supérieures à $23 \text{ (} 5 \div 2157 = 23,18 \text{)}$.

On calcule alors pour ces fréquences des répartitions théoriques entre les parties concernées ; ici : $F_i(A.T.) = .7843 F$ et $F_i(N.T.) = .2157 F$.

Il s'agit ensuite de comparer ces fréquences théoriques aux fréquences réelles dans les parties. Soit, par exemple, le vocable «fils» dont on nous donne :

$$3825 + 262 = 4087$$

et pour quoi on calcule :

$$3205 + 882 = 4087$$

Il faut alors comparer la fréquence réelle dans l'A. T. (3825) à la fréquence calculée ou théorique (3205).

On utilise à cet effet soit l'*écart-réduit*, dont un calcul simplifié et suffisant donne :

$$(3825 - 3205) / \sqrt{3205} = 10,9516$$

chiffre dont on lit la signification dans une table, et qui nous dit ici que la probabilité d'atteindre un tel écart du seul fait du hasard est très faible ; on en conclut, un peu facilement certes, mais utilement, que cet écart suffit à faire de ce vocable un mot bien plus présent ou fréquent dans l'A.T. qu'on ne pouvait s'y attendre, donc, si l'on veut, à en faire un mot caractéristique de cette partie du corpus. Il faut toujours veiller à ne pas être trop affirmatif...

Un autre test, un peu plus précis et plus puissant, est celui du KHI^2 ; en réalité, pour un degré de liberté, ce qui est le cas ici, la valeur du KHI^2 est simplement le carré de la valeur de l'écart-réduit. On calcule :

$$\frac{(3825 - 3205)^2}{3205} + \frac{(667742 - 3825) - (667742 - 3205)}{667742 - 3205}$$

$$\text{soit : } \frac{(3825 - 3205)^2}{3205} + \frac{(3205 - 3825)^2}{667742 - 3205}$$

$$\text{ou : } 119,9376 + .578 = 120,516$$

(On note que 119,9376 est le carré de 10,9516)

chiffre dont on lit la signification dans une table dont voici un extrait :

KHI^2 :	2,70	3,84	6,64	7,88	10,83	12,12
Probab. :	.10	.05	.01	.005	.001	.0005

probabilité signifiant probabilité pour que cet écart ne soit pas dû au hasard. On voit que 120 recouvre une probabilité extrêmement faible.

Les résultats sont parfois différents. Soit le mot «jour».

On nous donne : $1861 + 426 = 2287$.

On calcule : $1794 + 493 = 2287$ (arrondi)

et : $KHI^2 : 2,43$. On ne peut donc pas affirmer, à moins d'accepter plus de 10 % de chance de se tromper, risque que personne n'acceptera de courir, que si il y a dans l'A.T. 67 occurrences de plus que ne le voudrait une répartition strictement équilibrée entre les deux parties, ce nombre soit dû à autre chose que le hasard des répartitions. On ne pourra donc pas dire que «jour» caractérise l'A.T. On remarque d'ailleurs que ce vocable est presque toujours et partout réparti aléatoirement dans les textes.

Or, la table V met ce vocable au 23ème rang des mots significatifs pour l'A.T. La différence d'appréciation est très grave.

Nous nous permettons de donner ci-après les listes du tableau V dans l'ordre des KHI², et en signalant les changements de rangs. On pourra remarquer, pour le N.T. par exemple, que les mots qui «gagnent» des rangs, sans que le nombre de rangs gagnés soit significatif en soi, nous semblent de fait caractériser fortement le texte :

'foi, vérité, disciple, croire, annoncer, glorifier, signe, enseigner, comprendre, guérir, ressusciter, témoignage, baptiser, démon, foule, charité, crucifier, ange...'

alors que ceux qui en «perdent» jouent ce rôle à un moindre degré assurément :

dire, seul, oeuvre, grâce, prier, chair, écrire...

mais aussi : 'demander, vouloir, aller, venir, entendre, entrer, appeler, pouvoir, devoir, ciel, homme...'

Il reste peut-être à parler des mots de basses fréquences. Le tableau II, qui donne les mots propres à un Testament, en est l'occasion. Tout en réservant la discussion évoquée plus haut de savoir si le contenu sémantique d'un vocable ne donne pas à ce dernier vocation à être utilisé dans telle partie du corpus et non dans telle autre, ce qui serait un important critère d'appréciation, on peut s'interroger pour savoir si, par exemple, les mots qu'on ne rencontre que 4 fois dans le N.T. sont bien caractéristiques de ce texte.

Il n'est plus question de s'en remettre, comme ci-dessus, à la loi normale et au KHI². On a longtemps laissé ces basses fréquences hors du champ des analyses, du moins au niveau des mots individuels. Sans entrer dans une querelle qui voudrait que seule la loi dite hypergéométrique soit applicable (ce qui est d'ailleurs soutenu par très peu de monde), on peut s'en remettre à la lecture des tables de la loi binomiale (voir le supplément de «Définition des notions de répartition...» dans les *Cahiers de Lexicologie*, 42, 1983-I, pp. 24-27).² On y lit par exemple que si la partie du texte fait 26,59 % du corpus, les vocables de fréquence 4 dont toutes les occurrences s'y trouvent ne le sont qu'avec une probabilité inférieure à 1 % de chance pour que ce soit le fait du seul hasard. On les admettra donc au rang des mots caractéristiques du N.T. C'est ce que fait le tableau II, mais sans nous dire pourquoi, et on craint que la raison pour laquelle on nous donne ces mots est qu'ils se rangent parmi les 75 premiers qui répondent à la seule définition d'être absents dans l'autre partie du corpus. Pour l'A.T., la même lecture n'est valable, nous montrent les tables, que pour les vocables de fréquence au moins égale à 22.

Ceci dit, donnons-nous, dans nos listes rectifiées, un rendu satisfaisant de la caractérisation des deux Testaments, une fois admises les réserves énoncées plus haut quant à la seule référence aux fréquences ? Non, bien entendu, car il manque certainement à ces listes des mots de fréquences inférieures qui devraient prendre la place de ceux dont la caractérisation est moindre. C'est un travail de détection à confier à l'ordinateur, seul capable d'appréhender l'ensemble du vocabulaire.

C'est une rectification qu'on serait très heureux de trouver dans une prochaine édition de la Concordance, et dans celles, attendues, qui seront consacrées

à d'autres langues. Mais, même sans une telle rectification, le Centre d'Informatique Biblique de Maredsous vient de nous offrir un ouvrage de valeur dont aucun lexicométricien ne saurait maintenant se passer.

D. DUGAST
Paris

NOTES

1. On connaît, en langue anglaise, la *Young's Analytical Concordance to the Bible (Authorized Version)*, et les Concordances de la *Revised Standard Version*, de la *New International Version* et de la *Good News Bible*. Adresses : The Bible Society, 146 Queen Victoria Street, London EC4V4BX, G.B. (pour la G.N.B.), Bible Division, Zondervan, 1415 Lake Drive SE, Grand Rapids, Michigan 49506, USA (pour la N.I.V.).
2. Une présentation de ces tables avec mode d'emploi paraîtra prochainement ici même.

Dans le tableau suivant, qui n'est que le reclassement des unités du tableau V, nous donnons successivement : le rang nouveau, le vocable, (le rang du tableau V), le signe + ou - disant si le rang a monté ou descendu, les valeurs du KHI², comparées pour les 2 Testaments à lire comme suit : s'il s'agit de l'A.T. (première partie), présence significative dans l'A.T. et absence significative dans le N.T. ; s'il s'agit du N.T., présence dans le N.T. et absence dans l'A.T. Nous ne donnons pas les fréquences réelles et théoriques par raison d'économie, mais le lecteur les trouvera ou les calculera facilement.

75 Mots caractéristiques de l'un ou l'autre Testament.

Ancien Testament					Nouveau Testament		
			A.T.	N.T.		N.T.	A.T.
1	roi	(2) +	124,1	451,0	disciple	(4) +	744,2
2	fils	(1) -	120,4	437,6	foi	(3) +	647,4
3	pays	(3) =	83,3	302,8	dire	(1) -	467,3
4	peuple	(4) =	52,6	191,1	croire	(7) +	435,1
5	prêtre	(8) +	23,5	85,4	foule	(48) +	329,6
6	filles	(9) +	22,8	82,9	monde	(6) =	328,3
7	épée	(11) +	21,9	79,6	ressuscit.	(26) +	290,4
8	ville	(5) -	20,7	75,2	vérité	(10) +	242,8
9	armée	(19) +	17,5	63,8	apôtre	(17) +	228,1
10	main	(6) -	17,2	62,4	baptiser	(49) +	191,1
11	holocauste	(28) +	17,1	62,3	synagogue	(52) +	191,1
12	oracle	(17) +	17,1	62,1	non	(9) -	184,4
13	dieu	(20) +	16,8	61,0	ange	(70) +	160,0
14	autel	(13) -	16,6	60,4	charité	(59) +	150,8
15	père	(7) -	16,6	60,3	corps	(21) +	148,0

75 Mots caractéristiques de l'un ou l'autre Testament (suite).

	Ancien Testament				Nouveau Testament			
			A.T.	N.T.			N.T.	A.T.
16	clan	(33) +	16,0	58,0	venir	(2) -	147,9	40,7
17	lévite	(34) +	15,1	55,0	recevoir	(17) =	140,6	38,7
18	or (subs.)	(12) -	15,0	54,7	crucifier	(71) +	139,7	38,4
19	an	(14) -	13,9	50,6	témoignage	(33) +	137,6	37,9
20	coudée	(36) +	13,9	50,5	démon	(60) +	132,8	36,5
21	argent	(18) -	13,6	49,5	prêtre-grd	(35) +	132,6	36,5
22	camp	(38) +	13,5	49,2	glorifier	(34) +	126,1	34,7
23	montagne	(15) -	13,4	48,6	accueillir	(53) +	117,6	32,3
24	feu	(59) +	13,2	47,9	autre	(5) -	109,9	30,2
25	sacrifice	(24) -	12,5	45,6	grâce	(19) -	109,1	30,0
26	bétail	(53) +	12,0	43,6	heure	(40) +	108,0	29,7
27	habiter	(30) +	11,6	42,3	vouloir	(14) -	105,6	29,0
28	famille	(43) +	11,5	41,6	falloir	(32) +	104,9	28,9
29	maison	(10) -	11,4	41,3	mort	(38) +	103,9	28,6
30	face	(25) -	11,3	41,1	entendre	(12) -	99,7	27,5
31	ennemi	(27) -	11,3	40,9	saluer	(67) +	96,5	26,5
32	prince	(39) +	11,2	40,7	contraire	(70) +	91,2	25,1
33	revenir	(22) -	10,9	39,6	enseigner	(48) +	89,6	24,6
34	coucher	(55) +	10,7	38,8	prier	(28) -	88,6	24,4
35	année	(31) -	10,6	38,5	écrire	(26) -	87,4	24,0
36	arche	(52) +	10,3	37,3	païen	(55) +	86,8	23,9
37	chef	(26) -	10,2	37,3	savoir	(11) -	82,5	22,7
38	bélier	(69) +	10,1	36,7	esprit	(25) -	82,4	22,7
39	nation	(16) -	9,9	36,1	saint(s)	(66) +	80,6	22,2
40	mois	(44) +	9,9	36,0	nouveau	(54) +	79,6	21,9
41	tribu	(41) =	9,9	35,9	bop	(20) -	77,1	21,2
42	officier	(73) +	9,7	35,4	siècle	(74) +	70,8	19,5
43	bête	(47) +	9,7	35,2	guérir	(61) +	70,6	19,4
44	entrée	(66) +	9,1	33,0	voir	(8) -	70,5	19,4
45	méchant	(72) +	8,7	31,6	aimer	(27) -	67,2	18,5
46	guerre	(60) +	8,3	30,3	advenir	(73) +	61,6	16,9
47	bois	(65) +	8,2	29,7	ciel	(18) -	58,5	16,1
48	habitant	(56) +	8,0	29,0	devoir	(23) -	57,8	15,9
49	huile	(70) +	7,6	27,8	chair	(41) -	53,7	14,8
50	bénir	(37) -	7,3	26,5	remettre	(56) +	51,0	14,0
51	briser	(74) +	7,3	26,4	juger	(51) =	50,8	13,8
52	eau	(32) -	7,1	25,8	frère	(15) -	49,7	13,7
53	côté	(54) +	7,0	25,3	demeurer	(42) -	47,7	13,1
54	serviteur	(29) -	6,6	24,1	gloire	(37) -	42,3	11,6
55	porte	(45) -	6,5	23,5	annoncer	(58) +	41,4	11,4
56	héritage	(67) +	6,5	23,5	signe	(68) +	40,4	11,1
57	alliance	(58) +	6,4	23,1	entrer	(36) -	40,1	11,0
58	colère	(48) -	6,3	22,8	connaître	(31) -	37,8	10,4
59	ordre	(57) -	6,2	22,6	jeter	(50) -	34,5	9,5

75 Mots caractéristiques de l'un ou l'autre Testament (suite).

Ancien Testament					Nouveau Testament		
			A.T.	N.T.		N.T.	A.T.
60	offrir	(51) -	6,1	22,1	comprendre	(75) +	33,7 9,3
61	ordonner	(64) +	5,7	20,9	grand	(22) -	29,5 8,1
62	tête	(40) -	5,5	20,2	péché	(43) -	26,4 7,3
63	désert	(75) +	4,4	15,9	rendre	(30) -	25,9 7,1
64	terre	(21) -	4,3	15,6	seul	(63) -	25,7 7,1
65	écouter	(49) -	4,2	15,2	oeuvre	(62) -	25,1 6,9
66	lieu	(61) -	4,0	14,3	demander	(57) -	23,6 6,5
67	rester	(62) -	3,9	14,2	appeler	(46) -	20,9 5,8
68	milieu	(63) -	3,9	14,2	personne	(72) +	20,3 5,6
69	monter	(50) -	3,8	14,0	maître	(65) -	20,1 5,5
70	mal	(71) +	3,6	13,3	vie	(44) -	18,8 5,2
71	parler	(35) -	3,1	11,4	parole	(29) -	16,9 4,6
72	coeur	(42) -	3,1	11,3	pouvoir	(47) -	11,8 3,2
73	jour	(23) -	2,5	9,2	tenir	(69) -	9,7 2,7
74	oeil	(68) -	2,3	8,3	homme	(45) -	4,9 1,3
75	prendre	(46) -	1,7	6,3	aller	(64) -	4,1 1,1

Remarques

Les valeurs des KHI² sont à lire selon les données de la table rapportée plus haut. Les signes + et - sont à apprécier en fonction du nombre de places gagnées ou perdues, sans que toutefois cela n'ait d'autre valeur que vaguement indicative. On n'oubliera pas que le traitement de tous les vocables introduirait dans les listes ci-dessus des vocables intermédiaires ; ces listes ne sauraient donc constituer un rapport définitif. Enfin, on n'oubliera pas qu'il s'agit ici d'indications, et qu'on peut avoir quelques raisons, sur des bases sémantiques par exemple, de modifier quelque peu le contenu des listes.

Cependant, on rejettera comme caractéristiques de l'A.T. les vocables «prendre, oeil, jour, coeur, parler, mal, monter», et on sera très prudent pour les vocables qui vont du rang 54 au rang 68. De même (mais on peut déjà s'interroger sur ce que peut signifier une caractérisation due à l'absence...), on rejettera les vocables «aller, homme, tenir, pouvoir», et on sera prudent pour «parole, vie, maître, personne, appeler, demander».

La caractérisation du N.T. semble plus affirmée ; c'est que le texte est plus court, et le hasard y a moins sa place quand les occurrences sont nombreuses. Les conclusions sont plus assurées dans un texte long. Cependant, on considérera avec quelque prudence le caractère distinctif de «aller, homme», et, du côté de l'absence, de «prendre».

UNE NOUVELLE REVUE DE LEXICOGRAPHIE :

Lexicographica, Niemeyer, Tübingen, 1985.

En avril 1985 est paru le n° 1 de la «Revue Internationale de Lexicographie», *Lexicographica*. Cette revue est publiée par Antonin KUČERA, Alain REY, Herbert Ernst WIEGAND et Ladislav ZGUSTA, en coopération avec la Dictionary Society of North America (DSNA) et la European Association for Lexicography (EURALEX). Les éditeurs appellent *Lexicographica* une publication soeur des *Cahiers de Lexicologie* d'une part, et de *Dictionaries : Journal of the Dictionary Society of North America* de l'autre.

Lexicographica est rédigée en allemand, en anglais ou en français ; elle comporte essentiellement deux parties : une section à thème et une section non thématique. Pour présenter *Lexicographica* 1, je décrirai en -I- la première de ces sections, qui forme le coeur de l'ouvrage. J'évoquerai ensuite, de -II- à -VI-, les autres sections.

-I- «THEMATIC PART : Monolingual and bilingual lexicography of German Language». Partie à thème : lexicographie monolingue et bilingue de la langue allemande, sous la direction de Herbert Ernst WIEGAND , pp. 1. 224.

Huit études composent cette partie. Une seule est en français (F.J. HAUSMANN), trois sont en allemand (P. KÜHN, W. RETTIG, W. SEIBICKE), quatre en anglais (H. BERGENHOLZ et J. MUGDAN, G. STEIN, J.C. TRAUPMAN, H.E. WIEGAND). Parmi leurs titres, on relève :

- (a) une étude sur les langues allemande, anglaise et française (HAUSMANN),
- (b) une contribution sur l'allemand et le français (RETTIG),
- (c) trois articles sur l'allemand et l'anglais (BERGENHOLZ et MUGDAN, STEIN, TRAUPMAN),
- (d) trois analyses sur l'allemand (KÜHN, SEIBICKE, WIEGAND).

C'est dans cet ordre que je procéderai pour ce compte rendu.

(a) *Trois paysages dictionnaires : la Grande-Bretagne, la France et l'Allemagne*, Franz Josef HAUSMANN, pp. 24-50.

Faisant suite à de nombreux travaux du même auteur, c'est ici l'ensemble des dictionnaires allemands et anglais qui est analysé, à l'attention du public français. L'étude de la typologie et de la qualité des dictionnaires prend son véritable intérêt une fois replacée dans la perspective dictionnaire. La dictionnaire - on doit ce néologisme à Bernard QUEMADA - c'est l'ensemble des activités menant à des dictionnaires, et touchant aux dictionnaires, la lexicographie étant plus particulièrement l'activité scientifique qui consiste à traiter des textes (en général pour la production de dictionnaires). F. J. HAUSMANN apprécie ce néologisme puisqu'il parle de *bataille*, de *culture*, de *domaine*, d'*échiquier*, de *miracle*, de *passion*, de *paysage*, de *phénomène* et de *regard dictionnaire* !

Les dictionnaires généraux monolingues des trois pays choisis sont caractérisés par la fulgurante montée du dictionnaire en Grande Bretagne de 1978 à 1983, l'importance tellement plus grande de la dictionnaire en France avec le *Grand Robert*, le *GLLF* et le *TLF*, et, en Allemagne, la carence du *DUDEN*. Dans l'ensemble des dictionnaires pédagogiques on relève «l'excellent *Harrap's 2000 Word English Dictionary*» (1981), et les *Dictionnaires du français langue étrangère (DFLE 1 & 2)* de chez Larousse. La richesse de l'illustration par l'exemple est fortement appréciée. C'est, là encore, sur le «paysage» allemand que l'auteur reste le plus réservé, mentionnant toutefois le dictionnaire de français langue étrangère de LÜBKE. L'étude systématique se poursuit par celle des dictionnaires de locutions, de constructions, de collocations (il n'en existe pas en français), des dictionnaires de synonymes-synonymie cumulative et distinctive-, des dictionnaires bilingues. Le statut du dictionnaire est ainsi souligné : «La France, pays du dictionnaire» ; «la Grande-Bretagne, futur pays du dictionnaire ?» (fait intéressant, le dictionnaire y joue le rôle d'«école parallèle»). Et l'Allemagne ? Le miracle dictionnaire ne s'y est pas produit ; la maison Langenscheidt reste cependant le géant mondial du bilingue. Et tous les espoirs restent permis avec le dictionnaire informatisé.

Le lecteur apprécie les planches et les tableaux, suivis de huit pages d'illustrations.

(b) *Die Zweisprachige Lexikographie Französisch-Deutsch, Deutsch-Französisch. Stand, Probleme, Aufgaben*. La lexicographie bilingue français-allemand, allemand-français. Etat actuel, problèmes, perspectives, Wolfgang RETTIG, pp. 83-124.

C'est une étude souvent très critique qui est faite, extrêmement détaillée et fouillée, illustrée de nombreux et riches tableaux. 40 dictionnaires (et éditions) ont été analysés : ce sont des dictionnaires généraux, à l'exclusion des dictionnaires de spécialité (sciences, techniques), des dictionnaires spécialisés (locutions, par exemple). W. RETTIG prend en compte les dictionnaires portant sur les seules deux langues annoncées. Il a voulu compléter le panorama donné par HAUSMANN en 1977. Faisant le point sur l'état de la lexicographie franco-allemande, l'auteur analyse les problèmes d'éditions successives. Des tableaux résument une étude comparative précise des caractéristiques matérielles des dictionnaires.

W. RETTIG s'est attaché à une étude typologique des relations d'équivalence entre la langue source et la langue cible, en distinguant les relations de convergence, de divergence, de congruence. Une précision d'ordre terminologique : on parlera de la *traduction d'un texte*, mais de l'*équivalent d'un lexème*, en accord, ici, avec HAUSMANN. Mais W. RETTIG reproche à ce dernier plusieurs points de la typologie des relations d'équivalence donnée en 1977 ; et il s'en prend en particulier à ce qu'il dit du traitement bilingue des termes scientifiques et techniques en montrant que les relations de congruence existent pour ces termes (*Dezimeter/décimètre*, *Gramm/gramme*, *Helium/hélium*), même si des variations existent (le fr. *hydrogène* a pour équivalent allemand *Hydrogen*, mais également *Wasserstoff*, plus courant).

Autre problème épineux : la structure syntaxique des lexies complexes. Une lexie simple en langue source doit-elle toujours correspondre à une lexie simple en langue cible ? Les cas des faux-amis est également pris en compte, et l'auteur analyse les commentaires explicites apparaissant dans les dictionnaires, souvent sous la forme de marques de niveau de langue et d'indicateurs de domaines.

D'une façon générale, W. RETTIG suggère d'élargir le traitement des équivalences par un recours systématique aux syntagmes illustratifs. Choix judicieux à notre avis, puisque c'est une incitation à mieux tenir compte des contextes élargis. Le problème de la nomenclature des dictionnaires est également analysé. L'auteur recommande de renoncer aux entrées constituées de vedettes complexes, rares ou vieilles. Il suggère une meilleure distinction entre les syntagmes figés et les syntagmes libres. Il constate que les dictionnaires aux nomenclatures étendues contiennent souvent peu de syntagmes, et inversement. Une distinction intéressante pour le classement des lexies construites : le classement « alphabétique par nids » (« nestalphabetisch »), comme dans le dictionnaire de WEIS-MATTUTAT 1978 et de MATTUTAT 1979 où les composés en *Bau-* de *Baubschnitt* à *Bauzeit* sont rangés dans le « nid » de *Bau*, avant la vedette *Bauch*. Dans le dictionnaire de WEIS-MATTUTAT 1979, c'est le classement « alphabétique par niche » (« nischenalphabetisch ») qui est adopté : la « niche » de *Bau* ne va que jusqu'à *Baubüro*, qui est suivi de la vedette *Bauch*. Dans les dictionnaires « glattalphabetisch », les lemmes apparaissent tous alphabétiquement et séparément. Les deux dernières catégories de dictionnaires suivent un ordre alphabétique strict (« striktalphabetisch »). Mais sous quelle vedette, à quelle lettre placer les lexies complexes ?

Dans un dernier paragraphe W. RETTIG rassemble une série d'autres paramètres, les indications métalinguistiques telles que le genre, la flexion, la prononciation, les marques d'écart, les illustrations. Et dans tous les cas il déplore les lacunes, les manques, les insuffisances, dont le catalogue serait à dresser.

W. RETTIG présente un tableau de ce qui forme le marché actuel (1984) des dictionnaires étudiés : un dictionnaire électronique, et 45 autres dictionnaires. Pour l'avenir, une meilleure information des futurs utilisateurs serait souhaitable, ainsi qu'une meilleure homogénéité des données métalinguistiques.

En conclusion on peut dire qu'on retrouve là nombre de caractéristiques des dictionnaires en général.

(c) Etudes portant sur les langues allemande et anglaise.

● *Linguistic Terms in German and English Dictionaries*. Nous adoptons la traduction «Termes de linguistique dans le dictionnaire : langues allemande et anglaise» (de préférence à celle donnée par les éditeurs : «Termes linguistiques en dictionnaires : domaine allemand et anglais»), Henning BERGENHOLTZ, Joachim MUGDAN, pp. 3-23.

L'auteur a réalisé un vaste tableau comparatif et critique du traitement de la terminologie linguistique dans les dictionnaires généraux, en offrant au lecteur 4 pages de tableaux tout à fait bienvenus dans une étude qui frôlait, sans cela, l'aridité ; ce genre de question tente tout linguiste : comment définit-on les termes de linguistique ? Les dictionnaires sont en général bien peu logiques. Ils utilisent des termes sans les définir, ou les emploient dans un sens différent ; les termes relevant du domaine de la linguistique dans les dictionnaires généraux ne sont pas toujours ceux qui constituent les nomenclatures des dictionnaires spécialisés. Les auteurs de dictionnaires ne s'expliquent guère, pas plus dans les préfaces, les listes d'abréviations ou de domaines, que dans leur métalangage. Et d'un dictionnaire à l'autre, les conceptions et les réalisations diffèrent encore. Notons que H. BERGENHOLTZ et J. MUGDAN défendent la nécessité de recourir à un *corpus de textes* représentatif (p. 13, note 5).

Pour une information claire et utile sur tous ces points, le meilleur moyen est de se reporter aux tableaux (pp. 5, 9-12, 18).

● *English-German/German-English Lexicography : Its early beginnings*. Les débuts de la lexicographie anglais-allemand, allemand-anglais, Gabriele STEIN, pp. 134-164.

C'est là une étude chronologique descriptive des dictionnaires. Les premiers dictionnaires bilingues ou multilingues allemand-anglais datent du XVI^e s. G. STEIN donne une liste de 138 dictionnaires datant de 1530 environ à 1716. Notons qu'il faut attendre 1687, avec Henry OFFELEN (professeur de droit et de sept langues) pour rencontrer le premier ouvrage anglais-allemand. G. STEIN souligne le rôle du public visé par les dictionnaires et met en lumière les principes lexicographiques qui sous-tendent les oeuvres étudiées. La connaissance des sources des dictionnaires permet souvent un meilleur éclairage de l'apport effectif de nouvelles productions.

L'auteur note qu'en comparaison avec le français, l'espagnol et l'italien, l'allemand arrive tard en lexicographie bilingue allemand-anglais (pour le français, on note les oeuvres de John PALSgrave, 1530, et Claudius HOLLYBAND, 1593).

Une étude historique fouillée sur les dictionnaires, avec une réflexion sur leurs origines, sur leurs principes conducteurs et sur leurs utilisateurs.

● *Modern Trends in German-English Lexicography*. Tendances actuelles de la lexicographie allemand-anglais, John C. TRAUPMAN, pp. 165-171.

L'auteur note que les techniques et méthodes lexicographiques sont en général mal diffusées parmi les auteurs de dictionnaires à venir. La plupart des dictionnaires généraux et techniques publiés depuis 1945 ne satisfont pas aux besoins. Cependant, depuis la seconde guerre mondiale, les dictionnaires tendent à accorder plus d'importance à la langue de la communication quotidienne et, surtout, à l'anglais américain. C'est au *COLLINS German-English/English-German Dictionary* (1980) que J. C. TRAUPMAN décerne la palme.

(d) Etudes portant sur l'allemand

● *Gegenwartsbezogene Synonymenwörterbücher des Deutschen :*

Konzept und Aufbau. Nous proposons la traduction «Dictionnaires de synonymes de l'allemand d'aujourd'hui : conception et structure» (de préférence à celle des éditeurs : « Le dictionnaire des synonymes en Allemagne à l'heure actuelle : conception et structure») Peter KÜHN, pp. 51-82.

P. KÜHN présente une étude distinctive des différentes sortes de synonymies dans l'allemand contemporain. Une première distinction à noter : celle qu'on peut faire entre les dictionnaires de synonymes et les dictionnaires notionnels (*Begriffswörterbücher*). On répartit les dictionnaires de synonymes en deux classes, selon qu'ils donnent ou non des informations sémantiques : ce sont respectivement les dictionnaires de synonymie distinctive (comparative), et cumulative. La première catégorie est peu représentée, la plupart des dictionnaires ici étudiés appartiennent à la synonymie cumulative, comme les deux grands «*Begriffswörterbücher*» de DORNSEIFF (1970) et de WEHRLE/EGGERS (1967). Certains dictionnaires se situent à mi-chemin entre les deux.

L'auteur étudie les précurseurs des dictionnaires actuels, avant d'envisager les principes de l'articulation de la macrostructure : les synonymes peuvent être classés par ordre alphabétique ou d'après un ordre notionnel systématique «ontologique». La «*Begriffssystematik*» de P. M. ROGETS (1852) comporte 1000 notions, ce qui rend complexes les critères de présentation et de traitement... L'auteur relève la dénomination de «*register-alphabetisiertes Synonymenwörterbuch*» (dictionnaire de synonymes alphabétisé par registres). L'ordre alphabétique reste cependant le plus fonctionnel, seul à permettre une consultation rapide du mot cherché (cf. le *DUDEN*). P. KÜHN analyse les critères de choix de la documentation lexicale. Le dictionnaire de DORNSEIFF (1970) *Der deutsche Wortschatz nach Sachgruppen* prend tous les registres en compte, tous les niveaux de langue et d'expression, et se fait fort de ne pas négliger les parlars oraux. Les grands dictionnaires alphabétiques comportent également un vocabulaire très étendu (langue courante, langues techniques, régionales). Mais, et cela les différencie des dictionnaires notionnels, on note l'ampleur des attestations de la langue parlée.

Passant à la microstructure, il convient de rappeler la difficulté qu'il y a à définir un synonyme. Existe-t-il de vrais synonymes ? Ne s'agit-il pas, plus précisément, de ressemblances, d'analogies ? Les dictionnaires actuels s'entendent sur une conception restreinte de la synonymie. Mais dans l'ensemble les auteurs s'expliquent peu. P. KÜHN étudie les caractéristiques et explications sémantiques, stylistiques, de niveaux de langue, fournies par les dictionnaires, et qui n'ont parfois qu'une valeur très relative.

Les dictionnaires de synonymes sont en général considérés comme des recueils de variantes stylistiques, ou comme moyens d'atteindre le mot juste : ce sont alors des objets normatifs, prescriptifs. Les dictionnaires de synonymes ne sont pourtant pas des guides normatifs, mais plutôt des manuels d'interprétation. L'article, qui comporte de nombreux extraits de dictionnaires ainsi que des tableaux, se termine par une importante bibliographie de dictionnaires et de littérature secondaire. C'est un très vaste ensemble qui a été analysé là, et qui aborde des problèmes rencontrés par maint auteur ou lecteur de dictionnaire.

● *Deutsche Schimpfwörterbücher*. Dictionnaires d'injures allemands, Wilfried SEIBICKE, pp. 125-133.

L'auteur, spécialiste bien connu de la langue allemande de la technique et de la science, décrit ici, successivement, 15 dictionnaires de termes d'injures (de 1797 à 1981). Ils sont souvent constitués par des linguistes amateurs, et n'apprennent rien, selon W. SEIBICKE, sur la lexicographie générale de l'allemand actuel. Certains ont tout juste un intérêt historique, ou un intérêt juridique pour qui est enclin aux jurons tout en craignant les poursuites judiciaires vers lesquelles cette fâcheuse habitude pourrait les mener... Quant aux psycholinguistes, ils pourraient eux aussi trouver là matière à étude. Mais la plupart de ces dictionnaires n'ont d'intérêt que par le charme de l'allemand régional. Avouons que cette lecture, au milieu d'un ensemble fort sérieux par ailleurs, ménage un bien agréable moment de détente !

● *German Dictionaries and Research on the Lexicography of German from 1945 to the Present. With a select Bibliography*. Dictionnaires et métalxicographie en Allemagne depuis 1945, Herbert Ernst WIEGAND, pp. 172-224.

Cet article, conçu pour un public non allemand, donne une vue d'ensemble. Il présente tous les dictionnaires allemands existants en 1945, et expose ensuite l'ensemble de la lexicographie jusqu'en 1984. Une place importante est accordée au dictionnaire de GRIMM et à sa révision, aux dictionnaires étymologiques, aux dictionnaires de sémantique historique, aux dictionnaires basés sur des études de corpus d'auteurs, à la lexicographie du vieux-, du moyen-, du nouveau-haut-allemand précoce, du langage socio-politique. C'est en particulier le *Wörterbuch der Deutschen Gegenwartssprache* qui est analysé, ainsi que le dictionnaire de WAHRIG, les 6 volumes du *DUDEN*, le *BROCKHAUS-WAHRIG* et le *DUDEN-UNIVERSALWÖRTERBUCH*. Nul doute que ce qui importe à l'heure actuelle, c'est la constitution d'un dictionnaire assisté par ordinateur.

D'importants tableaux récapitulatifs illustrent cet article, qui comporte par ailleurs de très nombreux renvois bibliographiques. Enfin, plus de 22 pages sont consacrées à une bibliographie sélective sur le sujet.

Toute étude portant sur l'histoire des dictionnaires ou, plus généralement, sur la dictionnaire allemande, devra s'inspirer désormais de cette mine de renseignements, mise à notre disposition d'une manière aussi claire et systématique.

-II- «NON-THEMATIC PART », pp. 225-242

Cette partie non thématique se compose de deux articles :

- *Enzyklopädische Lexika als Hilfsmittel zum kritischen Lesen*. Les dictionnaires encyclopédiques au service d'un lecteur critique, Walter LENSCHEN, pp. 225-238. L'auteur s'interroge sur l'apport des dictionnaires encyclopédiques, en partant d'un texte trouvé dans un livre d'école de 1941.

- *Samuel Johnson (1709-1784) : Bicentenaire de sa mort*, Franz Josef HAUSMANN, pp. 239-242.

-III- «DISCUSSION»

Notes sur la réception du «Dictionnaire françois» (1680) de Pierre Richelet, Laurent BRAY, pp. 243-251. L'auteur analyse les discours critiques sur l'oeuvre de RICHELET.

-IV- «REPORTS»

Il est fait état de deux colloques :

- *International Conference on Lexicography*. Conférence internationale sur la lexicographie, Exeter, 9-12.09.1983, Alan KIRKNESS, pp. 252-253.

- *Historisches Wörterbuch des deutschen Gefühlswortschatzes*. Dictionnaire historique du vocabulaire émotionnel allemand, Aachen, 23-25.11.1983, Sabine PLUM, Gabriele WAND, pp. 252-257.

-V- «REVIEWS»

- Kita TSCHENKELI, *Georgisch-deutsches Wörterbuch*, 1965, 1970, 1974, par Georg BOSSONG, pp. 257-262.

- Arley GRAY, *Longman dictionary of American English*, 1983, par Gerald M. DALGISH, pp. 262-267. C'est surtout le traitement grammatical et le vocabulaire qui sont étudiés. Le dictionnaire comporte 38 000 entrées, ce qui est peu par rapport aux 55 000 entrées du *Oxford Student's Dictionary of Current English*.

- Manfred HÖFLER, *Dictionnaire des anglicismes*, 1982, et Josette REY-DEBOVE, Gilberte GAGNON, *Dictionnaire des anglicismes*, 1980, par Franz Josef HAUSMANN, pp. 267-268. Ces deux dictionnaires au titre identique sont pourtant complémentaires, dans la mesure où J. REY-DEBOVE et G. GAGNON utilisent des fonds d'exemples déjà fournis par des travaux antérieurs, alors que M. HÖFLER a lui-même dépouillé nombre de nouvelles sources contemporaines ; «le rapport de forces entre les deux dictionnaires est inversé» : celui de 1980 veut être «un dictionnaire d'usage», un ouvrage «pratique et vulgarisateur», l'autre ouvrage se voulant «strictement historique et scientifique».

- *Studien zur neuhochdeutschen Lexikographie*. Etudes sur la lexicographie du nouveau-haut-allemand, sous la direction de H. E. WIEGAND, 1979-1983, par Reinhold WERNER, pp. 268-277. Ces études apparaissent alors même

que d'autres événements historiques interviennent, tels que la fondation de la *European Association for Lexicography*, la fondation des revues *Lexicographica* et *Lexicographica Series Maior*, ainsi que le projet d'une *Encyclopédie internationale de la lexicographie*. On retrouve dans les 4 volumes l'analyse critique des prestations lexicographiques des dictionnaires d'allemand (surtout dans les dictionnaires monolingues de langue générale). Cet ensemble constitue un ouvrage de référence pour l'analyse de ces dictionnaires. C'est là une somme lexicographique qui devrait développer la communication entre lexicographes et «méta-lexicographes» de différentes langues.

- Fritz NEUBAUER, *Die Struktur der Explikationen in deutschen einsprachigen Wörterbüchern*, 1980, par Ladislav ZGUSTA, pp. 277-278. Il s'agit d'une étude du vocabulaire utilisé dans les définitions de dictionnaire, et de la proposition d'une méthode pour l'évaluation de la «circularité». Le vocabulaire de base ne coïncide pas avec le vocabulaire utilisé dans les explications données. L'auteur entrevoit la possibilité de rassembler un vocabulaire limité mais fonctionnel et suffisamment clair, que différents dictionnaires pourraient accepter. Il se pose aussi la question de savoir si tous les dictionnaires utilisent les mêmes hypéronymes. Enfin, il traite de dictionnaires spécialisés, de registres techniques, en discutant surtout le fait de savoir dans quelle mesure il convient d'utiliser dans les définitions des mots de la langue générale ou des termes techniques.

-VI- «SHORT REVIEWS»

Cette section présente des comptes rendus par H. E. WIEGAND (pp. 279-282), W. WOLSKY (pp. 282-285), et Ladislav ZGUSTA (pp. 285-286), sur des dictionnaires et des travaux de lexicographie parus entre 1982 et 1984.

Ouvrage de référence bibliographique à plus d'un titre, *Lexicographica* entame avec ce premier numéro une bibliographie systématique des articles de lexicographie de 122 revues. Chaque numéro de *Lexicographica* exploitera ainsi les revues de l'année passée. C'est sur l'année 1983 que s'ouvre cette bibliographie (pp. 287-295).

Enfin, les derniers dictionnaires et les dernières publications en lexicographie et dictionnaire sont indiqués pp. 296-300 et un index des noms facilite la consultation de ce volumineux ouvrage (pp. 300-307).

J'ai souligné à plusieurs reprises combien les planches et tableaux sont utiles et efficaces dans les études présentées. Mais ils gagneraient encore en efficacité à être régulièrement accompagnés de titres, de légendes.

Cette nouvelle publication est la bienvenue. Elle sera saluée non seulement par les lexicographes et «dictionnaristes», mais aussi par tous les linguistes amenés à consulter des dictionnaires de toutes langues.

Indispensable dans toutes les bibliothèques spécialisées.

D. CANDEL
C.N.R.S., Paris.