

1988  
Stages  
10

0883

Année universitaire 1987-1988  
24ème promotion ENSB

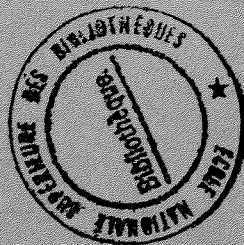
DESS "Informatique Documentaire"  
Université LYON I

Option ENSB "Conception et Gestion des Systèmes et des Réseaux  
d'Information"

Jean-François GIRARDOT

RAPPORT DE STAGE

"Problèmes posés par la documentation informatisée  
et le projet de thésaurus sectoriel de l'esclavage antique  
au Centre de Recherches d'Histoire Ancienne de Besançon  
(LIA 338 du CNRS)".



1988  
Stages  
10

Directeur de stage : Pierre LEVEQUE

Année universitaire 1987-1988  
24ème promotion ENSB

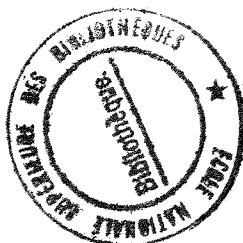
DESS "Informatique Documentaire"  
Université LYON I

Option ENSB "Conception et Gestion des Systèmes et des Réseaux  
d'Information"

Jean-François GIRARDOT

RAPPORT DE STAGE

"Problèmes posés par la documentation informatisée  
et le projet de thésaurus sectoriel de l'esclavage antique  
au Centre de Recherches d'Histoire Ancienne de Besançon  
(UA 338 du CNRS)".



1988  
Stages  
10

Directeur de stage : Pierre LEVEQUE

SOMMAIRE

1. PRESENTATION DU STAGE

- 1.1. Stage et lieu de travail
- 1.2. Stage et résultats du travail

2. MOTS-CLES, THESAURUS, INFORMATISATION: L'EXPERIENCE DE LA MAISON DE L'ORIENT

- 2.1. La Maison de l'Orient Méditerranéen
- 2.2. Activités documentaires
- 2.3. Réalisation des notices
- 2.4. Le vocabulaire d'indexation
- 2.5. L'élaboration du thésaurus

3. LE CENTRE DE RECHERCHE D'HISTOIRE ANCIENNE ET LA QUESTION DU THESAURUS SECTORIEL DE L'ESCLAVAGE ANTIQUE

- 3.1 Le Centre de recherches d'Histoire ancienne de Besançon
- 3.2 Bases documentaires d'un thésaurus sectoriel
  - 3.2.1 Le fichier GIREA
  - 3.2.2 Structure de l'index du GIREA
    - 3.2.3.1 Caractéristiques liées à l'écriture en général
    - 3.2.3.2 Caractéristiques quantitatives

4. METHODOLOGIE ADOPTEE POUR L'ELABORATION DU THESAURUS

- 4.1. Règles adoptées pour le tri du vocabulaire de l'index
  - 4.1.1. Préfixe spécificateur de champ
  - 4.1.2. Nombre du descripteur
  - 4.1.3. Guillemets
  - 4.1.4. Tiret
  - 4.1.5. Déterminant spécifiant le sens d'un terme
  - 4.1.6. Rapports entre notions
  - 4.1.7. Mention de point de vue
  - 4.1.8. Permutation des mots-clés
  - 4.1.9. Caractères employés
  - 4.1.10. Règles d'ordre général
- 4.2. Détermination des relations entre descripteurs
- 4.3. Produits documentaires

5. UTILISATION DU THESAURUS ET CHOIX DOCUMENTAIRES.

- 5.1. Solution centralisé
- 5.2. Solution intégrée
- 5.3. Conclusion et remerciements

## 1. PRESENTATION DU STAGE

### 1.1. Stage et lieu de travail

Le stage proposé par le Centre de Recherches d'Histoire Ancienne de Besançon concernait l'étude des possibilités de réalisation d'un thésaurus sectoriel de l'esclavage antique. Il s'est déroulé en deux lieux correspondant à peu près à deux étapes dans notre travail:

- la phase d'information et de prise de contact s'est effectuée à la Maison de l'Orient à Lyon, qui sert de structure d'accueil au réseau FRANTIQ. La présence d'une base locale hébergée sur place, d'une équipe d'informaticiens, d'un poste de coordination documentaire et de plusieurs bibliothèques travaillant en collaboration représentait pour notre propos un intérêt non-négligeable.

- l'essentiel du stage s'est effectué à Besançon, dans un local du Centre d'Histoire Ancienne, où deux Ingénieurs Techniques gèrent et exploitent la base de données du Centre; celle-ci est hébergée au Centre de Calcul de Grenoble, et consultée localement à l'aide d'un terminal. Le stage a consisté principalement en un travail sur le vocabulaire d'indexation produit par le Centre, et en une prise de connaissance des moyens informatiques utilisables.

### 1.2. Stage et résultats du travail

Deux contraintes ont limité le présent travail:

- le manque de temps : le vocabulaire de départ comptait 4500 termes, la partie à structurer environ 2900. Or il est admis qu'un thésaurus de 2000 à 3000 termes demande 6 à 8 mois de travail à une personne (cf. VAN SLYPE G. - Conception et gestion des systèmes documentaires, p.115).

- la période de vacances universitaires, qui a réduit considérablement le contact avec les usagers de la base locale, chercheurs ou étudiants, ainsi que les occasions d'observer le comportement réel des interrogateurs.

Nous avons réagi à ces contraintes tout d'abord en concentrant notre effort sur une seule facette du thésaurus qui nous était plus accessible parce que plus familière et mieux connue: cela nous a obligé à avancer petit à petit, micro-facette par micro-facette, alors qu'il eût été plus intéressant et plus payant d'avoir une vue globale des problèmes. Par ailleurs, nous avons dû simuler une demande documentaire: en effet, pour un ancien étudiant en lettres classiques et en linguistique ancienne, étranger à la problématique propre à l'histoire de l'Antiquité, il n'était pas difficile d'imaginer une demande documentaire venant de l'extérieur, car elle correspondait à une situation réellement vécue. Dans les deux cas, de ce corpus de notions référant plus spécifiquement à l'esclavage antique, c'est donc la partie linguistique qui a retenu notre attention.

D'autres facteurs ont également donné à ce travail une résonance et une signification particulières:

- le fait que le thésaurus était d'emblée présenté comme nécessaire, et comme une étape inéluctable dans les productions documentaires du Centre de recherches d'Histoire Ancienne.
- l'élaboration au niveau de la base FRANTIQ d'un macro-thésaurus consacré aux sciences de l'Antiquité en général, et qui pouvait constituer une base pour notre travail.
- l'existence d'un vocabulaire riche et déjà constitué, indexant un ensemble de textes fondamentaux pour notre sujet.
- la présence d'une équipe de recherche soudée autour d'un vocabulaire théorique commun.

Le fragment de thésaurus déjà constitué ne peut être tenu pour immédiatement opérationnel, puisqu'il n'a pas subi l'épreuve du feu représentée par un essai d'indexation sur des documents réels; il n'a pas non plus été présenté à l'ensemble de l'équipe, ce qui ne manquerait pas de faire apparaître des faiblesses cachées. Notre ambition se limite à le voir jouer, pour le Centre de recherches d'Histoire Ancienne, le rôle d'un document de travail qui permettrait:

- de situer l'originalité du thésaurus en regard des autres produits documentaires: index, fichiers, lexiques, etc...
- d'expliquer l'intérêt d'un thésaurus, mais aussi de rendre sensible la quantité et la qualité du travail que nécessite sa réalisation.
- de présenter le fonctionnement du thésaurus, son élaboration, et de constituer une typologie des écueils documentaires à éviter pour rendre possible son élaboration.
- d'amener à une réflexion sur la spécificité du travail documentaire et sur la place à lui accorder dans une structure de recherche.

## 2. MOTS-CLES, THESAURUS, INFORMATISATION: L'EXPERIENCE DE LA MAISON DE L'ORIENT

### 2.1. La Maison de l'Orient Méditerranéen

L'étude préliminaire de cet établissement se révèle riche d'enseignements pour qui veut analyser les problèmes parallèles qui se posent au Centre de Recherches d'Histoire Ancienne de Besançon : la Maison de l'Orient étudie les mêmes domaines ( sciences de l'Antiquité ), collabore à la même base informatisée ( FRANTIQ ), et se trouve confrontée aux mêmes questions concernant l'indexation des documents et l'interrogation automatisée.

On peut définir très rapidement la Maison de l'Orient comme un groupement fédératif d'unités de recherche et des centres documentaires qui en dépendent. Ces centres sont tous producteurs de notices pour FRANTIQ, mais des liens plus étroits les unissent à l'intérieur de la Maison : proximité géographique, collaboration, thèmes de recherche communs, coordination des activités documentaires. On y trouve 8 bibliothèques spécialisées :

- Victor Loret ( Egyptologie ) : 6500 titres.
- Salomon Reinach ( Archéologie classique ) : 30000 volumes, 52 périodiques.
- Institut d'Archéologie Orientale : 2000 ouvrages.
- Fernand Courby ( Antiquité classique ) : 9000 ouvrages, 68 périodiques.
- Antoine Bon ( Arabo-byzantin ) : 3000 volumes.
- Sources Chrétiennes : 10000 volumes.
- Céramologie : 100 ouvrages, 70 revues.
- Zumoffen ( préhistoire et quaternaire ) : 350 volumes et 12 périodiques.

La Maison de l'Orient totalise par ailleurs un fonds commun de 28700 diapositives, 1500 cartes, 20000 estampages. Elle réunit dans ses murs une centaine de chercheurs.

## 2.2. Activités documentaires:

Les notices bibliographiques produites par ces bibliothèques sont enregistrées dans une base locale, puis reversées périodiquement dans FRANTIQ sous le SUNIST comme celles des autres producteurs. La base FRANTIQ s'est édiflée en trois phases :

- 1) Etude de faisabilité.
- 2) Concentration : dès la fin 1986 la base atteignait 46500 références avec une augmentation annuelle de 15000 unités.
- 3) Travail sur le vocabulaire et mise en cohérence : la phase la plus intéressante pour nous puisqu'elle en est à son début, attendant actuellement les moyens financiers nécessaires à sa réalisation.

La "philosophie" présidant à cette élaboration apparaît dans quelques principes directeurs:

- Respect des particularités de chaque centre : notamment des politiques d'achats, des pratiques locales, des conceptions de la recherche.
- Liaison nécessaire des réalisations documentaires avec la pratique concrète des laboratoires de recherche, à la fois consommateurs et producteurs d'information.

Ces deux règles visent à protéger la recherche et à assurer au chercheur un minimum d'autonomie et de liberté. Dans la pratique, elles se trouvent limitées par des impératifs documentaires : la communicabilité des documents, et son corollaire : la compatibilité des documents dans un système informatisé. Il est cependant primordial de ne pas se laisser entraîner par idéal intempestif de normalisation, séduisant en théorie, mais dangereux en pratique. Les activités scientifiques de recherche produisent leurs propres concepts que le secteur documentaire doit non seulement respecter, mais également favoriser et protéger. Ce souci constitue la raison d'être de la documentation au sein d'une unité de recherche et doit se retrouver dans chaque produit documentaire, où l'on privilégiera un aspect réaliste et pragmatique.



### 2.3. Réalisation des notices:

Leur confection sous le logiciel TEXTO nécessite l'emploi d'un certain nombre de champs communs et identiques:

- entre participants à FRANTIQ, cette communauté est limitée à un minimum correspondant à un moyen terme entre notice minimale et moyenne des normes AFNOR : zones auteur, titre, mot-clé, maclé, langue, ISBN/ISSN. Chaque participant peut en outre définir ses champs supplémentaires propres
- elle est poussée un peu plus loin à l'intérieur de la maison de l'Orient où l'on a tendu à définir une notice "pragmatique". Des réunions de bibliothèques ont tenté d'aboutir à une définition commune du contenu et de la forme des zones concernées.

Mais la notice elle-même est conçue comme un produit évolutif: on en est venu à créer des champs nouveaux ( par exemple : le champ "mot-libre" ), à remodeler le contenu des anciens.

### 2.4. Le vocabulaire d'indexation:

Il s'agit là essentiellement des descripteurs du champ "mot-clé". En conformité avec les règles précédemment mentionnées, ils sont choisis librement par l'indexeur et non limités en nombre. Cette liberté est contre-balancée par l'utilisation des "maclés" et par un travail de réflexion opéré sur les descripteurs.

- 1) Les "maclés" ou "mots-clés macroscopiques" sont constitués d'entités de trois chiffres obéissant à un découpage de la réalité selon trois axes : espace, temps, discipline. Ils ont été mis au point par le C.R.A. de Valbonne, et apportent une aide considérable à l'interrogation des données archéologiques. Ils évitent les écueils du vocabulaire naturel : l'étiquette de chaque maclé peut être redéfinie ou modifiée à la demande. A terme, ils peuvent devenir transparents à l'utilisateur de la base.
- 2) Le travail de réflexion est mené à partir du vocabulaire déjà engrangé dans la base : c'est donc un travail a posteriori qui puise à trois sources:
  - a) les réunions de travail.
  - b) les interrogations de la base.
  - c) les documents de travail propres à chaque producteur.

- a) La confrontation des expériences particulières a conduit à l'élaboration de listes d'autorité définissant un vocabulaire contrôlé commun dans les domaines suivants: LIEUX, PERSONNAGES, AUTEURS, ARTISTES, MYTHOLOGIE, POLITIQUE, ADMINISTRATION, ETAT, ARMEE, RELIGION, PERSONNAGES SECONDAIRES, MUSEE, COLLECTION, ADMINISTRATION ( = collectivité-auteur ). Elle a également produit une normalisation dans la formulation des dates.
- b) La pratique des interrogations a permis de définir une syntaxe plus précise des descripteurs. On a ainsi préféré la forme composée LITTERATURE LATINE aux formes multiples LITTERATURE, LATIN : l'interrogation LITTERATURE ET LATIN produirait beaucoup de bruit dû à l'accrochage de notices indexées par exemple comme suit: LITTERATURE, GREC, PHILOSOPHIE, LATIN. On utilisera donc plutôt le syntagme, la forme simple LATIN présentant le désavantage supplémentaire d'un descripteur trop large et difficile à utiliser, qu'il serait souhaitable de remplacer par LATIN (LANGUE) pour un emploi plus restrictif.
- c) La base FRANTIQ a édité des listes de mots-clés propres à chaque participant : elles ont permis une première analyse des données ( nombre de descripteurs différents utilisés, fréquences moyenne et extrême d'emploi, fréquences comparées et pourcentage d'occurrences dans la base par centre ) d'où ressortait la nécessité d'une réduction rationnelle du nombre de mots-clés employés. Elles ont également permis un premier "nettoyage" de la base par une correction des erreurs les plus grossières : fautes d'orthographe, erreurs de saisie, erreurs de champs,... Faute de temps et de moyens, on n'a pu que commencer ce travail long et fastidieux, auquel il faudrait adjoindre un ensemble de tâches plus intéressantes qui sont actuellement traitées au coup par coup:

- doublets singulier / pluriel, doublets graphiques et problèmes de translittération, synonymes.

- formes homonymiques à disjoindre, formes à vérifier ou à préciser : "Caton" à côté de "Caton d'Utique" et "Caton l'Ancien"; "mouvement (concept); quel est le sens du descripteur "horizon"? Peut-on accepter un terme aussi large que "Antiquité"?

- on a choisi dans un grand nombre de cas de laisser coexister formes simples et syntagmes ( autres qu'expressions toutes faites et locutions consacrées du type : "comédie nouvelle", "machine de guerre", "cratère à volutes" ).

On trouve ainsi, après "médecin",

"médecin grec"  
 "médecin militaire"  
 "médecin public"

après "philosophe"	"philosophe grec"
après Aristote	Aristote "Métaphysique"
	Aristote "Rhétorique"

ce qui nous contraindra à interroger avec la troncature: MEDECIN\*, PHILOSOPHE\*, ARISTOTE\*. Ceci pose éventuellement des problèmes supplémentaires, par exemple dans la mesure où la mention du titre de l'oeuvre en descripteur constitue une redondance par rapport à la liste d'autorité AUTEURS et OEUVRES. Il faudra donc penser par la suite à un chaînage terme à terme dans le sens d'une inclusion du syntagme dans le terme simple.

### 2.5. L'élaboration du thésaurus:

Elle correspond également au démarrage de la phase de mise en cohérence du vocabulaire, non plus à partir des producteurs, mais à partir du noyau de pilotage de la base qui possède une vue synthétique de l'ensemble. Le thésaurus est élaboré essentiellement sur un vocabulaire déjà existant, et ne répond pas à des normes extérieures précises : on pourra le définir comme un thésaurus a posteriori non sans ressemblance avec certains thésaurus à facettes; il respecte la multiplicité des approches de l'Antiquité et le point de vue "anthropologique" ( au sens anglo-saxon du concept ) dominant dans la Maison. C'est en fait un macro-thésaurus couvrant le domaine des sciences de l'Antiquité dans son ensemble et destiné à servir de base à tout thésaurus sectoriel.

Le thésaurus est construit sur la base d'une référence comparative à trois documents :

- une liste de tous les mots-clés de la base.
- un dictionnaire encyclopédique de type courant ( vocabulaire a priori ).
- la grille RIDA élaborée par le C.R.A. de Valbonne.

L'élaboration a fait apparaître plusieurs types de termes :

1) Le vocabulaire médian "PACTOL" ( P = Peuples, A = Anthroponymes, C = Chronologie absolue / relative, T = Toponymie, O = Oeuvres écrites, L = géographie ). Dans une première étape, les noms propres sont simplement ventilés sous ces rubriques sans liaison entre eux ; cependant, on étudie sous chaque rubrique les problèmes posés par une éventuelle hiérarchisation. Certaines questions en effet méritent une réflexion approfondie : ainsi par exemple pour le terme "Byzance" qui sort évidemment de ses limites purement géographiques.

2) Le vocabulaire médian constitué par les concepts eux-mêmes :

un premier travail a permis de dégager 31 "Thèmes" principaux hiérarchisés jusqu'au 4ème niveau :

HISTOIRE-CIVILISATION  
 ---> fait de civilisation  
       ---> expansion  
           ---> hellénisation

Entre les termes sont pris en compte 4 types de rapports : inclusion, exclusion, synonymie et homonymie. Le premier état du travail privilégie bien sûr les rapports de type simple.

Parallèlement à cette élaboration a été conçu un plan de travail sous la forme d'un retour aux unités productrices. Il sera demandé aux indexeurs de classer leurs descripteurs en 3 niveaux : termes Génériques, Spécifiques et Fins; les deux premiers niveaux sont destinés à entrer en contact avec le vocabulaire du thésaurus d'une manière qui reste encore imprévisible : de cette adéquation dépend le devenir du thésaurus et sa capacité à remplir réellement son rôle de macro-thésaurus englobant. La règle suivie en la matière consiste à ne jamais appauvrir le vocabulaire local et particulier au profit du général et du central. Quant au 3ème niveau des termes "Fins", il restera définitivement hors-thésaurus et probablement inaccessible à une interrogation "normale".

Ce travail appelle deux types de remarques :

- Il semble que, si l'on comprend bien la stratégie adoptée, beaucoup d'espoirs sont placés dans une cohérence spontanée entre indexeurs et noyau de pilotage de la base; si par hasard un raccordement "naturel" s'avère impossible entre le thésaurus et les strates supérieures des vocabulaires particuliers, ne risque-t-on pas d'aboutir à deux types de vocabulaires disjoints, l'un central et contrôlé, l'autre local et plus libre? Dans cette hypothèse, les producteurs et la base pourraient souffrir des conséquences d'une sur-indexation (redondance mots-clés / maclés / thésaurus : d'où perte de temps et de maniabilité).

- On peut se poser la même question sur une base concrète en se demandant quel serait le statut d'un thésaurus sectoriel de l'esclavage dans son rapport à ce macro-thésaurus. Une brève analyse de ce macro-thésaurus dans son état actuel suffit pour comprendre que les principales rubriques intéressant les descripteurs de l'esclavage seront totalement dispersées dans ce système. Il est toujours possible effectivement d'imaginer une possibilité automatisée pour passer de ce vocabulaire éclaté de la base à un thésaurus sectoriel cohérent; il faut cependant remarquer que dans ce cas le travail accompli au niveau central aura très peu profité au producteur local : et ceci restera d'autant plus préoccupant tant que le problème de l'adéquation des vocabulaires des deux niveaux n'aura pas été résolu. Le seul remède envisageable à cet état de fait réside probablement dans un renforcement de la fonction de coordination des activités documentaires aussi bien entre l'équipe de pilotage et les différents producteurs, qu'entre les producteurs eux-mêmes.

### 3. LE CENTRE DE RECHERCHES D'HISTOIRE ANCIENNE ET LA QUESTION DU THESAURUS SECTORIEL DE L'ESCLAVAGE ANTIQUE.

#### 3.1 Le Centre de recherches d'Histoire ancienne de Besançon

Le Centre d'Histoire Ancienne de Besançon, UA 338 du CNRS, regroupe une cinquantaine de personnes travaillant suivant trois axes de recherche:

- I. Formes de dépendance et rapports sociaux.
- II. Systèmes politiques et dynamiques sociales.
- III. Espaces, productions, échanges.

Il alimente en notices la base FRANTIQ, comme la Maison de l'Orient, le GS de Bordeaux et le CRA de Valbonne. Sa production tend à se spécialiser dans le repérage et l'indexation de la littérature concernant l'esclavage antique.

Le Centre a formé avec les Universités de Milan et de Naples le noyau fondateur du Groupe International de Recherches sur l'Esclavage Antique (GIREA) qui compte à son actif depuis 1970 l'organisation de 17 colloques internationaux, avec la participation des Universités de Camerino, de Varsovie, de Lecce, de Rome-La Sapienza, de Madrid, de Mayence, de Barcelone.

#### 3.2 Bases documentaires d'un thésaurus sectoriel

##### 3.2.1 Le fichier GIREA

Il est clair que toute la littérature recensée ou acquise par le Centre d'Histoire Ancienne ne présente pas le même intérêt pour mener une réflexion commune concernant l'utilité et l'élaboration d'un thésaurus: seule une production propre au Centre et à ses partenaires les plus proches est à même de garantir une cohérence entre cette entreprise documentaire et les visées des chercheurs travaillant sur l'esclavage antique.

Comme à Lyon, les listes de mots-clés correspondant aux notices rentrées dans la base ont subi un premier "désherbage" consistant en une élimination des erreurs les plus évidentes et faciles à repérer; ici cependant des efforts particuliers ont été déployés pour commencer la mise en cohérence d'un secteur-clé du fichier général "besbib": il s'agit du sous-fichier "GIREA" qui réunit la bibliographie des travaux organisés par le Groupe International de Recherches sur l'Esclavage Antique, textes de référence pour les chercheurs du Centre. Des 17 colloques du GIREA, les 15 premiers tenus de 1970 à 1986 ont été publiés; acquis au fur et à mesure, catalogués et indexés, ils ont été intégrés au fichier "fordep" qui réunit la littérature concernant les formes de dépendance sous l'Antiquité. Ils en ont ensuite été extraits et leurs 422 notices ont formé un fichier spécial où le vocabulaire d'indexation a été remanié. Ce corpus représentait donc, grâce à son index re-travaillé, un champ de travail de premier choix en même temps qu'une base de départ tout indiquée.

### 3.2.2 Structure de l'index du GIREA

Les 4525 descripteurs qui composent l'index sont répartis en champs de volume variable, mais souvent très voisins de ceux du vocabulaire "médié" déjà évoqué:

- (A)Anthroponymes et Théonymes: 258 références correspondant à peu près à la section A de PACTOL.
- (CG)Chronologie et Géographie: 1296 références réunissant les champs L, T, C et P de PACTOL.
- (H)Historiographie: 195 références secteur particulier au Centre de Besançon qui révèle là une de ses originalités: tenter de mener une réflexion sur l'histoire de la recherche historique concernant l'esclavage antique.
- (M)Méthodologie: 294 références secteur non pris en compte par PACTOL, spécialité de la Bibliothèque du CRA de Valbonne.
- (S)Sources: 525 références champ proche des Oeuvres de PACTOL, mais de contenu moins littéraire: le Centre privilégie la notion de sources globales (types d'oeuvres, types de documents ou de support), plutôt que celle d'oeuvre individuelle.

(MHS):765 références

ancienne catégorie en refonte qui réunissait auparavant les trois précédentes.

(T)Thèmes:6685 références

la partie de loin la plus importante, qui correspond très précisément aux Thèmes de PACTOL et fait l'objet du macro-thésaurus. Ce champ est actuellement le plus travaillé par FRANTIQ, qui demande à ses participants l'étude d'une ventilation de ses termes en sous-thèmes sur trois niveaux: Générique, Spécifique et Fin. Ce souci d'un raccordement possible au macro-thésaurus doit guider l'élaboration d'un thésaurus sectoriel. Pour l'instant, seuls les termes (T) en langue originale ont été ventilés à part dans l'index du GIREA. L'équipe de Besançon étudie actuellement la ventilation des autres termes.

L'étude du vocabulaire se fera donc principalement à partir de ce champ, qui constitue le noyau central de la terminologie technique spécifique à la recherche sur l'esclavage. Cependant, il ne faut pas sous-estimer la nécessité de relations avec les termes des autres rubriques, même si ces dernières sont moins originales et peuvent être élaborées hors du Centre ( répertoires communs, listes d'autorité, ouvrage de références ).

### 3.2.3 Caractéristiques des descripteurs

Les mots-clés figurant dans l'index possèdent un certain nombre de traits rendant impossible leur fonctionnement comme descripteurs au sein d'un thésaurus. Ces traits doivent être répertoriés et explicités:

#### 3.2.3.1 Caractéristiques liées à l'écriture en général:

\* Difficultés liées à certaines notations :

- la spécification du champ apparaît en tête du descripteur sous la forme: (T)ESCLAVAGE. Or ce type de notation s'avère peu pratique pour une consultation documentaire extérieure. Pour une interrogation de la base, il contraint à interroger en utilisant la troncature gauche, procédé long et coûteux interdit par de nombreux serveurs. Dans le classement alphabétique des descripteurs, il introduit des sous-classements rendant plus difficile la consultation de guides et de listes contrôlées, particulièrement pour un usager extérieur au Centre de recherches d'Histoire Ancienne.



- la forme adoptée pour la **transcription des mots étrangers** peut aboutir à multiplier les mots-clés différents: ainsi on a en face du terme francisé (T)CLAROTES les formes transcrites (T)"CLAROTAI" et (T)"KLAROTAI" qui renvoient à la même réalité sinon aux mêmes références. S'il est tout à fait judicieux de vouloir multiplier les entrées dans une liste alphabétique ou un index pour assurer une liaison entre ces différentes graphies, il est inutile, fastidieux et même préjudiciable dans la perspective d'un thésaurus de les traiter comme autant de descripteurs.
- la distinction entre nombres est maintenue, si bien que ESCLAVE et ESCLAVES possèdent une signification légèrement différente: le premier se rapporte à un concept, le second à des personnes réelles. Cette distinction est concevable dans le cadre d'une équipe travaillant sur un vocabulaire fin. Encore apparaît-il souvent assez artificiel de vouloir séparer, dans un travail de recherche, aspect conceptuel et mention d'objets réels qui se trouvent intellectuellement liés; on sent d'ailleurs une certaine hésitation dans ce domaine avec - par exemple - l'indexation d'un même document au moyen des deux genres: (T)VENTES-VOCABULAIRE et (T)VOCABULAIRE DE LA VENTE. Toujours est-il qu'une telle distinction ne peut être maintenue dans le cadre d'un thésaurus: elle nécessiterait en effet une abondante littérature explicative qui réduirait la maniabilité recherchée; par ailleurs et surtout, le dédoublement des descripteurs augmenterait de façon considérable le nombre des relations sémantiques à prendre en compte pour l'élaboration du thésaurus, et les structures hiérarchiques ne pourraient plus être représentées de manière graphique: d'où perte de temps, de clarté, de lisibilité, et augmentation des risques d'incohérences.

\* Ambiguïté de certains symboles graphiques:

- les **guillemets** sont utilisés pour noter les termes en langue originale: (T)ESCLAVE coexiste avec les termes latin et grec (T)"SERVUS" et (T)"DOULOS"; ces guillemets introduisent un sous-classement qui peut se révéler préjudiciable à l'interrogateur, mais ils ont surtout l'inconvénient de servir aussi à la notation d'expressions en français du type: (T)ESCLAVES "CONTRACTUELS" ou (T)ESCLAVAGE "IDEAL" dont la forme reste difficile à prévoir.
- l'utilisation du tiret "-" semble particulièrement polysémique, car il est employé pour noter:
  - a) la spécification de certaines notions: (T)ESCLAVE-PERSONNAGE DE THEATRE, (T)ACCADIEN-LANGUE. On rencontre cependant également des formes utilisant d'autres symboles, comme (T)PLEBE=CLIENTELE IMPERIALE ou (T)ESCLAVES="INSTRUMENTUM VOCALE".
  - b) les expressions employant normalement le tiret: (T)ESCLAVAGE-MARCHANDISE.

c) les rapports entre deux descripteurs: (T)ESCLAVES-MAITRES ou (T)PAUVRETE-RICHESSE. On rencontre également la barre de rapport: (T)ESCLAVAGE ANCIEN/ESCLAVAGE MODERNE.

d) le point de vue sous lequel est abordé un sujet: (T)MYCENES-ECONOMIE. Cependant cet emploi doit être examiné dans le cadre plus général de la permutation des mots-clés.

e) la permutation des mots-clés précoordonnés, pour permettre une entrée dans l'index à tous les termes jugés significatifs: on a ainsi en face de (T)NOMBRE DES ESCLAVES CORINTHIENS, (T)ESCLAVES CORINTHIENS-NOMBRE et (CG)CORINTHIENS-NOMBRE DES ESCLAVES; (T)MYCENES-ECONOMIE répond au mot-clé composé (T)ECONOMIE MYCENIENNE. Ce procédé et la manière dont il s'effectue soulèvent un certain nombre de problèmes qui seront étudiés plus loin. Pour notre propos actuel, il nous suffit de mettre en évidence les ambiguïtés d'ordre sémantique nées de l'emploi polysémique du tiret; comment le non-initié peut-il distinguer entre spécification, permutation et rapport entre notions? Ainsi, quelle est la signification exacte des mots-clés suivants:

- (T)DIEUX-ESCLAVES
- (T)ECONOMIE-SOCIETE STRUCTURES
- (T)HILOTES-POLITIQUE SPARTIATE-RAPPORTS
  - (T)ESCLAVES-BERGERS LETTRES
  - (T)ESCLAVES-BERGERS UNIONS
  - (A)PALEMON-ONOMASTIQUE ? (1)

En fait, tous ces problèmes liés à la rédaction des mots-clés peuvent être résolus de manière assez simple, par une normalisation dans la création des termes et par l'adoption de règles clairement formulées communes aux équipes d'indexation; il est important que cette normalisation aboutisse à la rédaction d'un document susceptible de servir de guide d'interrogation des bases qui l'auront adopté, de manière à rendre l'utilisateur lointain plus autonome.

---

1. Les réponses sont respectivement: relations dieux/esclaves, structures économiques et sociales, rapports entre les hilotes et la politique spartiate, bergers esclaves sachant écrire, unions entre bergers esclaves, Q. Remmius Palemon sous le rapport de l'onomastique.

### 3.2.3.2 Caractéristiques quantitatives:

#### \* Nombre des mots-clés et des références

L'index du GIREA contient un total de 4525 mots-clés différents pour 10148 références, soit une fréquence moyenne de 2,24 par mot-clé. En comparaison, les bibliothèques de la Maison de l'Orient fournissent des chiffres totaux légèrement supérieurs pour les mots-clés, mais les fonds indexés sont incomparablement plus volumineux: ainsi 6635, 4576 et 6906 mots-clés en 1986 pour des fonds s'élevant respectivement à 5267, 2355 et 7200 ouvrages; les fréquences s'étagaient de 4,36 à 5 occurrences par mot-clé, en légère augmentation par rapport à l'année précédente, ce qui était le but recherché.

Nous avons testé les 4 premières lettres du champ (T) qui comportent exactement 679 mots-clés, soit 15% de l'ensemble des mots-clés, soit encore 22% des mots-clés du champ: un échantillon réellement représentatif par conséquent; cet examen nous apprend que 76,8% des mots-clés sont employés une seule fois, 9,7% deux fois, 4,4% trois fois; qu'au-delà de la fréquence 2 il ne reste déjà plus que 13% du vocabulaire, et que ce chiffre diminue rapidement: 5,7% au-dessus de 5 et 1,9% au-dessus de 10.

#### \* Exhaustivité-spécificité de l'indexation

Chacun des 422 documents du GIREA est indexé par 24 descripteurs en moyenne, alors que les bibliothèques déjà évoquées, à la même époque, connaissaient des moyennes correspondantes allant de 4,87 à 9,47 et que l'objectif déclaré était de tenter de diminuer encore ce nombre. Il est vrai que l'on ne peut pas comparer des textes de référence qui sont l'objet d'une étude approfondie et de débats théoriques importants, dans la mesure où ils constituent l'identité-même du Centre, avec des documents d'étude d'une portée plus courante. Il faut cependant songer que ces documents ( soit 407 discussions et travaux particuliers dépouillés en 15 colloques indexés de façon globale ) réunissent 5153 pages rédigées: ils se voient donc attribuer, par page, presque 2 indexations (1,96 très exactement) et presque un mot-clé (0,87); on ne s'étonnera donc pas de l'extrême finesse de l'indexation, ni de la tendance à une pulvérisation du vocabulaire.

#### \* Taux de pré-coordination

Le taux de pré-coordination du vocabulaire (nombre moyen de mots significatifs par descripteur) a été évalué de manière rapide à partir d'un corpus réduit de 213 mots-clés: il est important (1,91%), mais reste dans la fourchette communément admise: 1,5 à 2% (2).

Ces données nous fournissent des indications utiles concernant notre marge de manoeuvre pour transformer le vocabulaire de l'index en langage documentaire compatible avec l'élaboration d'un thésaurus: on ne pourra en effet pas jouer sur une augmentation du taux de précoordination pour réduire le nombre de descripteurs et limiter son volume.

### 3.2.3.3 Caractéristiques structurelles:

La conception du thésaurus repose tout entière sur le descripteur comme entité de base du langage documentaire et sur une séparation fondamentale entre descripteurs et non-descripteurs. Tout langage d'indexation qui ne respecte pas cette distinction ou remet en cause le principe d'identité du descripteur reste alors, sous cette forme première, incompatible avec ce projet documentaire.

#### \* Descripteurs homonymes dans des rubriques différentes

Nous avons évoqué la ventilation des descripteurs sous différentes rubriques et la mention de la rubrique sous la forme de préfixes de type (A),(CG),(M), etc...Or la même forme peut se rencontrer dans deux rubriques différentes, comme (MHS)ESCLAVAGE-CHAMP LEXICAL et (T)ESCLAVAGE-CHAMP LEXICAL; ceci se produit également avec les mots-clés permutés, classés sous une autre rubrique que la forme non-permutée: ainsi (M)CONCEPT DU TRAVAIL en face de (T)TRAVAIL-CONCEPT, (MHS) RESEAU SEMIO-LEXICAL DE L'ESCLAVAGE en face de (T)ESCLAVAGE-RESEAU SEMIO-LEXICAL. Ce type de procédé pose le problème délicat de savoir s'il s'agit encore du même descripteur:

- Si oui, sa forme écrite doit être constante et ses références classées toujours sous la même rubrique ( ou les mêmes rubriques ) car un descripteur doit correspondre à une seule notion (3). Si l'on fractionne le descripteur en unités plus petites, il faudra identifier à leur tour ces unités de manière formelle pour les lier entre elles par un réseau de relations encore plus dense.
- Sinon, la similitude de graphie peut prêter à confusion et il vaut mieux utiliser deux termes différents, entre lesquels la question des relations se posera de façon plus claire (4); en effet quelles sortes de relations

---

3. AGENCE FRANCAISE DE NORMALISATION. Paris. - Règles d'établissement des thésaurus monolingues : norme française homologuée Z 47 100. - Paris : AFNOR, 1981. Cf. 3.1.1. a)

4. Au contraire des termes en langage naturel, les descripteurs d'un langage documentaire sont aussi peu polysémiques et ambigus que possible.

Cf. VAN SLYPE G. - Conception et gestion des systèmes documentaires, p 112 "résolution des cas de double emploi: [...] affectation d'un même terme dans deux champs différents, [...].

sémantiques envisager entre deux aspects de la même notion?

On voit que, quelle que soit la réponse adoptée, on a intérêt à garder intacte la notion d'identité du descripteur, seul élément capable de fournir un point d'attache solide aux relations sémantiques à élaborer.

\* La pratique de la permutation des mots-clés

Comme de nombreux mots-clés étaient composés de plusieurs termes significatifs, on a éprouvé le besoin de les faire figurer dans l'index à chaque entrée significative; au lieu de rédiger une liste alphabétique permutée, on a donc créé de nouveaux mots-clés sur le schéma suivant: à partir de (T)CONTACTS PAUSANIAS-HILOTES, on a forgé (T)PAUSANIAS-HILOTES CONTACTS et (T)HILOTES-PAUSANIAS CONTACTS. Cette pratique a eu pour conséquence une confusion entre non-descripteurs (formes destinées à être de simples renvois vers les descripteurs choisis) et descripteurs (formes "pleines" servant d'étiquette à un stock de références). En conséquence ces formes permutées se sont comportées comme de véritables descripteurs autonomes: rangées, comme nous l'avons vu, sous une rubrique différente de celle de la forme non-permutée, ou comportant des références autres que celle-ci; on peut citer comme exemples:

36 (M)ANALYSE DU DISCOURS

.00112,00113,00121,00128,00129,00131,00152,00153,00154,00156,  
.00169,00173,00175,00180,00182,00192,00196,00198,00199,00202,  
.00208,00214,00217,00232,00238,00239,00240,00241,00242,00243,  
.00244,00249,00270,00278,00290,00411.

20 (M)DISCOURS-ANALYSE

.00111,00214,00217,00232,00238,00239,00240,00241,00242,00243,  
.00244,00249,00270,00278,00290,00317,00320,00321,00403,00411.

1 (MHS)ANALYSE DU DISCOURS

.00122.

27 (MHS)DISCOURS-ANALYSE

.00022,00024,00108,00109,00110,00112,00113,00121,00122,00128,  
.00129,00131,00152,00153,00154,00156,00169,00173,00175,00180,  
.00182,00192,00196,00198,00199,00202,00208.

25(T)NOMS DES ESCLAVES

.00010,00011,00013,00033,00037,00038,00111,00140,00154,00161,

---

AGENCE FRANCAISE DE NORMALISATION. Paris. - Règles d'établissement des thésaurus monolingues : norme française homologuée Z 47 100. Cf. 5.7.2 : " Une notion ne trouve en général sa place que dans une seule facette. Quand un terme apparaît dans plus d'une facette, sa signification exacte mérite d'être redéfinie."

.00177,00182,00189,00194,00197,00200,00211,00212,00240,00241,  
00243,00294,00303,00306,00339.

16 (T)ESCLAVES-NOMS

.00010,00013,00033,00037,00038,00140,00154,00161,00182,00200,  
.00212,00240,00243,00294,00306,00339.

Ces quelques exemples permettent de mieux cerner les inconvénients de cette pratique:

- un surcroît de travail inutile: au lieu d'entrer 1 ou 2 mots-clés permutés à chaque indexation d'un terme composé, il suffirait d'établir une liste alphabétique permutée où les entrées seraient établies une fois pour toutes.
- un risque d'erreurs accru, dans la mesure où l'autonomie du mot-clé permuté rend le repérage de l'erreur incertain et la correction plus difficile.
- une incompatibilité du vocabulaire avec le langage d'un thésaurus, dans la mesure où l'identité entre mots-clés non-permutés et permutés est sujette à caution et pose un problème quant au type de relations à établir entre eux.

\* Les relations hiérarchiques entre mots-clés

Il est intéressant d'examiner comment se pose la question des relations hiérarchiques entre mots-clés de l'index à partir d'exemples:

a) Soit un article discutant le problème de l'évaluation du "nombre des esclaves de Corinthe" (référence n°204). Les mots-clés correspondant à cette question sont:

2 (CG)CORINTHE  
.00078,00272.

1 (CG)CORINTHIENS-NOMBRE DES ESCLAVES  
.00204

68 (T)ESCLAVES  
.00010,00011,00019,00022,00024,00033,00034,00037,00041,00054,  
.00059,00075,00080,00085,00090,00091,00093,00100,00101,00111,  
.00115,00127,00167,00169,00175,00176,00180,00181,00188,00190,  
.00192,00195,00198,00208,00209,00214,00226,00228,00230,00274,  
.00277,00283,00290,00294,00298,00299,00301,00302,00306,00307,  
.00309,00310,00316,00354,00356,00359,00361,00366,00369,00370,  
.00375,00376,00377,00382,00393,00396,00399,00401.

1 (T)ESCLAVES CORINTHIENS-NOMBRE  
.00204.

15 (T)NOMBRE DES ESCLAVES  
.00014,00029,00030,00032,00039,00053,00058,00059,00060,00067,  
.00140,00173,00200,00204,00394.

1 (T)NOMBRE DES ESCLAVES CORINTHIENS  
.00204.

On remarquera que les notions sont systématiquement indexées au mot-clé le plus spécifique possible, contenant une référence unique, tandis que le mot-clé générique plus fréquent reste employé plus rarement et toujours en double. Le mot-clé se comporte alors plutôt comme une clé d'accès au document que comme un recueil de références. Cette méthode d'accès direct convient à des chercheurs qui dès le départ connaissent bien le fonds et le contenu des articles recherchés. Il n'est cependant pas possible dans ce cas, pour un interrogateur extérieur, d'interroger en utilisant les possibilités d'opérations booléennes à partir de descripteurs simples (offertes par exemple par le logiciel TEXTO) comme suit:

mot-clé=(NOMBRE D'ESCLAVES)[notion précise correspondant à une question théorique souvent débattue] et (CORINTHE);  
il faudra interroger avec la double troncature par les expressions \*NOMBRE\* \*ESCLAVE\* \*CORINTH\* après consultation de la longue liste des mots-clés contenant chacun de ces termes. La raison d'être du thésaurus est de signaler ces liens générique/spécifique et de renvoyer l'interrogateur de l'un à l'autre.

b) Soit un usager recherchant toute la littérature du GIREA concernant la question des noms propres humains (anthroponymes); il devra interroger la zone mot-clé en liant les chaînes \*NOMS\* \*ANTHROPONYMIE\* \*ONOMASTIQUE\* par "ou"; il aura auparavant tâtonné et sélectionné ses termes à partir d'une somme de mots-clés non-pertinents recueillis par la demande des chaînes \*NOM\* et \*NYMIE\*; encore ne pourra-t-il pas trier dans cette masse de références en ne demandant par exemple que les noms d'origine grecque appliqués à des esclaves ou à des affranchis. Le rôle du thésaurus est de permettre ce type d'interrogation en fournissant des termes génériques structurés de façon à pouvoir "chapeauter" tout un secteur sémantique, et des possibilités d'intersections intéressantes entre secteurs.

Pour nous résumer, nous dirons que le thésaurus doit créer et signaler des liens générique/spécifique qui dans l'index du GIREA sont:

- implicites et réels comme entre NOMBRE DES ESCLAVES et NOMBRE DES ESCLAVES CORINTHIENS (mais dans ce cas il faut une double indexation à la fois au spécifique et au générique, donc un travail accru).
- implicites (grâce au voisinage alphabétique) et virtuels, comme entre CORINTHE et CORINTHIENS-NOMBRE DES ESCLAVES.
- inexistants comme entre ANTHROPONYMIE et NOMBRE DES ESCLAVES, TERMINOLOGIE DES ESCLAVES et "DOULOS", ou ESCLAVES, GUERRES SERVILES et "BELLUM SERVILE".

#### 4. METHODOLOGIE ADOPTEE POUR L'ELABORATION DU THESAURUS

##### 4.1. Règles adoptées pour le tri du vocabulaire de l'index

Les mots-clés de l'index ne pouvant être récupérés tels quels pour les raisons évoquées plus haut, nous proposons ici quelques règles de restructuration en descripteurs pour la constitution du thésaurus. Les transformations réellement opérées sur le vocabulaire serviront d'exemples.

##### 4.1.1. Préfixe spécifique de champ:

Le champ "Thèmes", qui doit être le départ de toute recherche - automatisée ou manuelle - sur le vocabulaire, ne doit pas porter de préfixe spécifique tels que (T)... ou (T)"...". Les autres champs, qui ne sont pas pour l'instant destinés à une intégration dans le thésaurus, peuvent continuer à figurer sous forme d'identificateurs non-structurés et classés selon l'ordre alphabétique (1). Ceci ne vaudra pas pour le champ "Méthodologie", dont les descripteurs sont trop fréquents et trop imbriqués aux notions des "Thèmes". On aura donc:

- Un champ "Thèmes" à accès direct où les termes seront appelés sous leur forme usuelle non-préfixée.
- Un champ "Méthodologie" à accès également direct, possédant sa hiérarchie propre et développant des relations associatives avec le champ précédent.
- Des champs d'identificateurs non-structurés, classés en suites alphabétiques, préfixés en (A),(CG),(H),(S), auxquels il faut ajouter les termes en langues originales (T)"...".

Exemple: les descripteurs ANTHROPONYMIE, TOPONYMIE, THEONYMIE, ONOMASTIQUE du champ "Thèmes" correspondent à un descripteur de l'arbre hiérarchique "Méthodologie": ANALYSE ONOMASTIQUE. De la même façon, chaque fois que la nécessité de noter l'aspect méthodologique d'un travail s'est fait sentir, on a créé un descripteur hiérarchiquement lié à METHODOLOGIE et associé au descripteur notant le thème linguistique principal.

---

1. AGENCE FRANCAISE DE NORMALISATION. Paris. - Règles d'établissement des thésaurus monolingues : norme française homologuée Z 47 100. - Paris : AFNOR, 1981. Cf. 3.1.3. Identificateurs et 6.2.7. Listes annexes.



#### 4.1.2. Nombre du descripteur:

Le descripteur est entré sous sa forme au singulier, sauf si l'usage ne connaît qu'une forme au pluriel, ou s'il existe une différence de sens entre les deux nombres; en-dehors de ces cas exceptionnels, le pluriel est traité comme un homonyme documentaire du singulier: l'entrée est limitée à la forme prise au singulier, aucun renvoi n'est prévu à partir du pluriel (2).

Pour les identificateurs qui constituent le vocabulaire fin destiné aux spécialistes d'une question, il est souvent préférable de maintenir la diversité des formes; il faut cependant noter que certains regroupements seraient utiles: normalisation des entrées nominales et verbales, de systèmes de renvois d'un déterminant à une expression complexe, etc...

#### 4.1.3. Guillemets:

Les guillemets ne sont pas employés, même dans la formulation d'expressions ou pour introduire une distance par rapport à l'emploi d'une forme; une seule exception: ils sont partout obligatoires dans le champ "vocabulaire en langue originale": (T)"...".

Exemple: (T)"DOULOS" = terme grec entré au nominatif singulier signifiant "esclave".

#### 4.1.4. Tiret:

L'emploi du tiret est réservé aux expressions qui le réclament dans l'usage courant; il est exclu dans tous les autres cas (3).

#### 4.1.5. Déterminant spécifiant le sens d'un terme:

On évite en général l'emploi des termes polysémiques. En cas de nécessité on juxtapose simplement au terme ambigu un déterminant destiné à en préciser le sens (4).

2. AGENCE FRANCAISE DE NORMALISATION. Paris. - Règles d'établissement des thésaurus monolingues : norme française homologuée Z 47 100. - Paris : AFNOR, 1981. Cf. 3.2.5.Nombre.

3. AGENCE FRANCAISE DE NORMALISATION. Paris. - Règles d'établissement des thésaurus monolingues : norme française homologuée Z 47 100. - Paris : AFNOR, 1981. Cf. 3.3.2. "C'est l'orthographe la plus largement admise qu'il convient d'utiliser. Si par suite de plusieurs usages on est arrivé à accepter plus d'une orthographe, les différentes versions doivent être introduites dans le thésaurus avec les renvois voulus pour indiquer leur synonymie."

4. AGENCE FRANCAISE DE NORMALISATION. Paris. - Règles d'établissement des thésaurus monolingues : norme française homologuée Z 47 100. - Paris : AFNOR, 1981. Cf. 3.3.4. Polysémie-Homographes.

Exemple: ACCADIEN LANGUE au lieu de: (T)ACCADIEN-LANGUE.

#### 4.1.6. Rapports entre notions:

Les descripteurs n'ayant vocation qu'à dégager des notions fondamentales simples, les rapports entre notions ne sont pas pris en compte, sauf cas particuliers:

- si le rapport en question correspond au domaine de deux descripteurs (ou plus), il doit être rentré sous chaque descripteur.

Exemple: (A)PALEMON et ONOMASTIQUE au lieu de: (A)PALEMON-ONOMASTIQUE.

- si le rapport en question ressortit à l'antithèse, il est possible de faire de ces termes d'une opposition sémantique bipolaire des synonymes documentaires (5).

#### 4.1.7. Mention de point de vue:

Les descripteurs ne peuvent subir les subdivisions courantes pour les mots-matières: on n'évoque donc pas le point de vue; si sa mention est jugée indispensable, il justifie l'indexation à un second descripteur (voir ci-dessus 4.1.6. Rapports entre notions).

#### 4.1.8. Permutation des mots-clés:

La pratique de la permutation des descripteurs subsiste exclusivement dans la réalisation d'une liste alphabétique permutée des descripteurs et des non-descripteurs: il s'agit là d'un artifice de présentation des descripteurs composés permettant des entrées pour ces descripteurs à chacun de leurs termes significatifs; ce procédé ne crée en aucun cas de nouveaux descripteurs.

Par ailleurs, les notions complexes introduites par la pratique de la permutation devront être réduites par une indexation sous plusieurs descripteurs simples. Ainsi, le mot-clé double qui apparaît simultanément dans deux champs différents (T)CLASSE-CONCEPT et (M)CONCEPT DE CLASSE devra être indexé par deux termes de sens simple:

- l'un sous la hiérarchie "LINGUISTIQUE" par le descripteur CONCEPT.

- l'autre sous la hiérarchie "METHODOLOGIE" par le descripteur ANALYSE SEMANTIQUE.

#### 4.1.9. Caractères employés:

5. AGENCE FRANCAISE DE NORMALISATION. Paris. - Règles d'établissement des thésaurus monolingues : norme française homologuée Z 47 100. - Paris : AFNOR, 1981. Cf 4.1.3. Critères de choix - rapprochement des différentes valeurs d'une même variable.

Les descripteurs seuls doivent à terme figurer dans le thésaurus en majuscules; cependant, tant que le vocabulaire de l'index n'aura pas été totalement remanié, il est bon de conserver sous cette forme les anciens mots-clés, ainsi que ceux toujours en place dans les facettes non-traitées et les identificateurs des listes annexes; les autres non-descripteurs (synonymes non retenus ne figurant ni parmi les anciens mots-clés ni dans une liste d'identificateurs) seront présents dans la liste alphabétique en minuscules. Les identificateurs constitueront des listes exactement semblables aux sections actuelles de l'index, suites alphabétiques non-structurées de termes en majuscules et préfixés.

#### 4.1.10. Règles d'ordre général:

Dans le remaniement à opérer sur le vocabulaire de l'index, il convient de suivre quelques lignes directrices touchant à sa structure générale.

- Le nombre des descripteurs, suffisamment élevé pour l'instant, ne doit pas être encore augmenté inconsidérément par la création inopportune de termes supplémentaires.

- l'exhaustivité de l'indexation ne doit pas être accrue: il ne s'agit pas, dans l'optique d'un thésaurus, de repérer la mention ou l'apparition d'un concept à une page précise (objet des index d'ouvrages), mais de rendre compte du contenu global d'un travail. En effet, plus un grand nombre de descripteurs caractérisent le même texte, et plus l'interrogateur risque de faire des rapprochements arbitraires: ainsi le spécialiste de la religion grecque à la recherche de documents évoquant la barbarie des Grandes Dionysies sera-t-il déçu par l'évocation concomittante de ces deux notions à propos d'un article dont l'intérêt se situe ailleurs (6). Le seul moyen de remédier à ce type de problème

---

#### 6. Réponse à la question: mot-cle=((T)BARBARIE et (T)GRANDES DIONYSIES):

ref .00182  
 auteur .MACTOUX (Marie-Madeleine)  
 titre .LE SYSTEME SEMANTIQUE DES TERMES DESIGNANT LES  
 ESCLAVES CHEZ ARISTOPHANE  
 local ."Ceti dipendenti e schiavitu nel mondo antico",Atti  
 del colloquio(GIREA),Camerino,11-13 gennaio  
 1978,Index,8,1978-1979,p7-47  
 donnees .- Tableau 1:La designation de l'esclave,p11.  
 - Tableau 2:Determinants,p12.  
 - Tableau 3:Actions de,p14-15.  
 - Tableau 4:Actions sur,p16.  
 - Tableau 5:Systeme antonymique,p18-19.  
 - Tableau 6:Schema associatif,p22.  
 - Tableau 7:Sens propre,p24.  
 - Tableau 8:Sens semantique de "doulos",p26.  
 - Tableau 9:"Oiketes",p26.

tout en conservant une indexation aussi riche, résiderait dans la possibilité d'augmenter encore le nombre de descripteurs et/ou le taux d'indexation.

- Or, le taux de précoordination, bien qu'acceptable, trahit une tendance à l'excès plutôt qu'à l'insuffisance: il faut donc y songer lorsqu'on rencontre des candidats-descripteurs comprenant trois termes significatifs. Il est vrai que l'élimination des mots-clés permutés de la liste des descripteurs aura comme effet une chute de ce taux, les mots-clés les plus composés étant aussi les plus répétés.

- 
- Tableau 10: "Therapon", p29.
  - Tableau 11: "Pais", p30-31.
  - Tableau 12: p32-33.
  - Tableau 13: Schème des termes designant les esclaves dans le discours comique d'Aristophane, p36.

date .1978-1979

langue .français

mot-cle .(MHS)SOURCES LITTERAIRES, (S)ARISTOPHANE, (CG)GRECE CLASSIQUE, (MHS)DISCOURS-ANALYSE, (M)ANALYSE DU DISCOURS, (T)TERMINOLOGIE DES ESCLAVES, (T)ESCLAVES-TERMINOLOGIE, (T)THEATRE GREC, (MHS)METHODOLOGIE, (T)GRANDES DIONYSIES, (T)DIONYSIES, (T)CONCOURS SCENIQUES, (T)"POLIS"-FETES, (T)FETES DE LA CITE, (T)ESCLAVES-ROLE COMIQUE, (T)ROLE COMIQUE DES ESCLAVES, (M)ANALYSE LEXICALE, (MHS)LEXICOLOGIE, (T)"DOULOS"-RESEAUX DES DETERMINANTS, (M)RESEAUX DES DETERMINANTS DE"DOULOS", (A)PANOURGOS, (M)SUBSTITUTS SEMANTIQUES, (MHS)SEMANTIQUE, (T)ESCLAVES-IMAGE PEJORANTE, (T)IMAGE PEJORANTE DES ESCLAVES, (T)"DOULOS"-RESEAU VERBAL, (M)RESEAU VERBAL DE "DOULOS", (T)"DOULOS"-RESEAU DES ACTIONS, (MHS)RESEAU DES ACTIONS DE "DOULOS", (T)"DOULOS"-VALEURS MORALES NEGATIVES, (T)VALEURS MORALES NEGATIVES DE"DOULOS", (T)"DOULOS"-SYSTEME ANTONYMIQUE, (M)SYSTEME ANTONYMIQUE DE"DOULOS", (T)PAIX PUBLIQUE, (T)PAIX DOMESTIQUE, (T)"OIKETES", (T)"THERAPON", (T)RAPPORTS ESCLAVAGISTES, (T)ESCLAVES-RAPPORTS, (T)STRUCTURES FAMILIALES, (T)FAMILLE-STRUCTURES, (T)"DIAKONOS", (T)"AKOLOUTHOS", (T)EMPLOIS-FONCTIONS, (T)FONCTIONS-EMPLOIS, (T)SERVITEURS, (T)ESCLAVES DOMESTIQUES, (T)DOMESTIQUES ESCLAVES, (T)"PAIS", (T)"DMOES", (T)VOCABULAIRE EPIQUE, (MHS)EPOPEE-VOCABULAIRE, (T)ESCLAVES-INSTRUMENTS DU POUVOIR, (T)BARBARES, (T)NOMS DES ESCLAVES, (T)ESCLAVES-NOMS, (M)ANTHROPONYMIE, (T)BARBARIE, (T)"ANDRAPODON", (T)PROCESSUS D'INTEGRATION DES ESCLAVES, (T)ESCLAVES-PROCESSUS D'INTEGRATION, (T)PROCESSUS D'EXCLUSION DES ESCLAVES, (T)ESCLAVES-PROCESSUS D'EXCLUSION, (T)"POLIS", (T)"OIKIA", (CG)GRECE-THEATRE

besbib .3851

Pour toutes les raisons énoncées ci-dessus, ainsi que pour la clarté de notre travail, nous avons délibérément choisi de réduire la spécificité de l'indexation. Il existe en effet une nette tendance à la redondance de l'information la plus spécifique au domaine étudié (7), qui nuit à la fois à l'économie du système et à l'exemplarité des modèles proposés. Nous avons donc systématiquement indexé au niveau générique lorsque nous pensions qu'il convenait de ne pas augmenter la pyramide hiérarchique, et que la spécificité devait être atteinte par l'utilisation logique des opérateurs booléens. Ainsi notre descripteur VOCABULAIRE remplace-t-il 47 mots-clés composés en (T)VOCABULAIRE... ou (T)...-VOCABULAIRE: en effet, la présence des spécificateurs ne s'avère pas indispensable, dans la mesure où l'interrogation par VOCABULAIRE pourra être, à terme, affinée par un recoupement avec les descripteurs ACHAT, AFFRANCHISSEMENT, ALIMENTATION... VIOLENCE. De la sorte, à partir d'un corpus de travail qui comptait initialement 175 mots-clés répartis sur 2 champs différents, nous aboutissons à un total de 36 descripteurs seulement, volume généralement considéré comme normal(8).

7. Nous remarquons en effet que des mots-clés comme NOMS D'ESCLAVES, ANTHROPOLOGIE DES ESCLAVES, ONOMASTIQUE DES ESCLAVES indexent fréquemment des passages déjà indexés par plusieurs mots-clés contenant le spécificateur D'ESCLAVES, sinon le mot-clé (T)ESCLAVES lui-même; il serait évidemment plus économique d'utiliser les ressources de l'interrogation automatisée en généralisant l'utilisation des opérateurs booléens: mot-clé = (ANTHROPOLOGIE) et (ESCLAVE). On peut citer comme exemple de redondance le passage n°11:

```
ref      .00011
auteur   .MASSON (Olivier)
titre    .LES NOMS DES ESCLAVES DANS LA GRECE ANTIQUE
local    .Actes du colloque 1971 sur l'esclavage, Paris, Les
        Belles Lettres, 1973, p9-23. Collection des "Annales
        littéraires de l'Université de Besançon, 140, Centre de
        Recherches d'Histoire Ancienne, vol.6
date     .1973
langue   .français
mot-cle  .(MHS)SOURCES LITTERAIRES, (MHS)SOURCES
        ARCHEOLOGIQUES, (MHS)ANTHROPOLOGIE, (T)NOMS DES
        ESCLAVES, (T)ESCLAVES-
        ANTHROPOLOGIE, (T)ESCLAVES, (MHS)SOURCES
        EPIGRAPHIQUES, (S)TABLETTES MYCENIENNES, (CG)GRECE
        MYCENIENNE, (CG)GRECE HOMERIQUE, (CG)GRECE
        CLASSIQUE, (T)ESCLAVES-TERMINOLOGIE, (T)TERMINOLOGIE
        DES ESCLAVES, (T)ESCLAVES-ORIGINE
        GEOGRAPHIQUE, (T)ORIGINE GEOGRAPHIQUE DES ESCLAVES
besbib   .6591
```

8. VAN SLYPE G. - Conception et gestion des systèmes documentaires. Cf page 118 à propos des études sur les données quantitatives usuelles: "Importance des groupes (thèmes ou

#### 4.2. Détermination des relations entre descripteurs

Le thésaurus prend en considération les trois principaux types de relations définissant un descripteur:

- relations d'équivalence entre synonymes documentaires, de la forme rejetée (non-descripteur) à la forme retenue.
- relations hiérarchiques entre descripteurs génériques et spécifiques; le cas de la relation partitive ne sera pas abordé.
- relations associatives ou "de voisinage" (le seul type de relation non-réciproque) indiquant une analogie de significations entre descripteurs non-liés par d'autres relations et appartenant ou non à une même hiérarchie; notre thésaurus prend cependant également en compte dans ce type de relation des termes dont le statut n'a pas été examiné (hors facette), ou appartenant à une liste extérieure non encore structurée avec les éléments de laquelle il est préférable de ne créer des relations d'ordre général.

#### 4.3. Produits documentaires

La réalisation expérimentale d'une première facette du thésaurus se concrétise par l'élaboration de trois documents:

- Une liste des termes intéressant cette facette, sous la forme d'une suite alphabétique permutée unique réunissant descripteurs et non-descripteurs. Chaque catégorie est ensuite détaillée dans une série de listes particulières concernant les types de relations entre éléments.
- Un recueil des schémas graphiques des divers arbres hiérarchiques, par facettes et micro-facettes (8 schémas).
- Un fichier automatisé sous le logiciel TEXTO, où chacun des 36 descripteurs correspondra à une fiche découpée en 8 champs et organisée comme suit:

---

facettes): fourchette de 30 à 40 descripteurs par groupe, en moyenne."

- \* Nom du descripteur en MAJUSCULES.
- \* Note d'application; pour la définition du sens du descripteur dans le thésaurus; particulièrement utile lorsque ce sens diverge par rapport à l'emploi habituel; par exemple, notre travail attribue au terme ANALYSE LINGUISTIQUE un sens assez particulier, puisqu'il est destiné à indexer les ouvrages traitant non pas d'analyse linguistique en général, mais de façon très spécifique de considérations méthodologiques à propos de la linguistique soit considérée de manière générale, soit simultanément sous plusieurs de ses aspects, soit sous l'aspect linguistique comparée exclusivement.
- \* Employé Pour = Non-descripteurs équivalents. Pour le rappel de tous les termes non-sélectionnés: synonymes envisagés en vocabulaire libre (minuscules) et anciens mots-clés pris dans la restructuration de la facette.
- \* Termes génériques, en MAJUSCULES: mention importante, car le rappel des termes supérieurs est indispensable pour l'indexation.
- \* Termes spécifiques, en MAJUSCULES.
- \* Termes associés, en MAJUSCULES: descripteurs au même niveau de la même hiérarchie, ou descripteurs de sens voisin dans d'autres hiérarchies. Sont notamment précisés les liens associatifs liant les hiérarchies "Linguistique" et "Méthodologie".
- \* Identificateurs associés: mention de tous les termes associés non admis comme descripteurs, c'est-à-dire identificateurs des champs (A), (H), (S), (T)"...", mais également mot-clés pas encore traités des champs (M) et (T), car extérieurs au domaine de la facette considérée.
- \* Coordonnées graphiques: mention des coordonnées du descripteur sur les schémas graphiques .

## 5. Utilisation du thésaurus et choix documentaires:

Il doit être bien évident que le travail réalisé dans le cadre du stage ne présentera pas le même intérêt selon les choix qui seront opérés par le Centre en matière de politique documentaire. Pour le domaine qui nous intéresse directement, les choix peuvent se résumer à l'alternative suivante:

- Solution "centralisée": reprise pure et simple d'un thésaurus élaboré à l'extérieur du Centre.
- Solution "intégrée": réalisation d'un thésaurus sectoriel propre compatible et adaptable au macro-thésaurus.

Le but du présent travail ne peut consister à trancher pour les responsables du Centre: il s'agit simplement de leur fournir un avis technique, des éléments de réflexion et de comparaison nécessaires à la prise de décision.

Notre étude s'est pour l'instant bornée à la structuration expérimentale d'une facette de thésaurus sectoriel sur trois ou quatre niveaux seulement, selon un schéma qui reste théorique et assez général, donc susceptible de remaniements et d'améliorations. Nous voudrions montrer ici les adaptations possibles de ce travail en fonction de la solution qui aura été choisie:

### 5.1. Solution centralisée:

La reprise d'un thésaurus extérieur implique une double indexation: à partir du vocabulaire commun aux participants de la base, et, si l'on ne veut pas perdre le "vocabulaire fin" local, à l'aide de ce second type de vocabulaire. On pourrait espérer un assez fort degré de compatibilité spontanée entre vocabulaires: or l'étude de ceux-ci nous a persuadé qu'il n'en était rien. Dans cette circonstance, notre travail propose des solutions de raccordement: vocabulaire du macro-thésaurus/vocabulaire fin, notamment en utilisant l'instauration de liaisons "floues" de type associatif. Pour faciliter l'opération d'indexation double, il est en effet possible d'associer descripteurs des derniers niveaux hiérarchique et termes du vocabulaire fin (quel que soit leur champ d'origine). Pour illustrer notre propos, nous prendrons l'exemple du descripteur ANTHROPONYMIE (fiche n°17 dans l'annexe) et du raccordement que nous proposons dans le schéma suivant:



SCHEMA N°1:vocabulaire  
extérieur:secteur du  
macro-thésaurus  
3ème niveau

ANTHROPONYMIE-----

relation associative

vocabulaire  
local:secteurs des mots-clés  
du Centre
 (T)NOMS DES ESCLAVES  
 (T)NOMS DES AFFRANCHIS  
 (T)NOMS DESAFFRANCHIES  
 (A)ACTE=affranchie  
 (A)ALEXIS=esclave  
 (A)ANTIGONUS=affranchi  
 etc...

N.B.: Nous observons que dans ce schéma, tous les champs du vocabulaire local partagent le même statut associé non-structuré.

5.2. Solution "intégrée":

L'élaboration d'un thésaurus sectoriel propre nécessite l'adaptation réciproque des vocabulaires des deux thésaurus sur les niveaux supérieurs des hiérarchies (niveaux "générique" et "spécifique" du champ "Thèmes"). Cette solution est celle qui garantit la meilleure collaboration entre équipes d'indexation, la plus grande cohérence documentaire, l'effort de réflexion le plus intéressant; il ne faut pas dissimuler non plus qu'elle réclame un surcroît de travail, un développement des échanges, diplomatie et opiniâtreté. Elle permet d'éviter la scission entre deux vocabulaires d'indexation incompatibles, mais tous deux indispensables.

Pour reprendre notre exemple, il semble très possible de mettre en cohérence les descripteurs génériques de ANTHROPONYMIE, c'est-à-dire ONOMASTIQUE et LINGUISTIQUE (Cf. tableau C de l'annexe), avec le secteur correspondant n°701 des "Thèmes" du macro-thésaurus; l'adaptation demanderait une simple révision des relations unissant trois termes que ce secteur présente sur le même plan: 728 anthroponymie, 729 toponymie, 733 onomastique, qui doivent à notre sens être hiérarchisés. Il en va de même, sans problème majeur, pour notre arbre hiérarchique METHODOLOGIE ANALYSE LINGUISTIQUE (Cf. tableau M), qu'il est possible d'harmoniser avec le secteur 1267 METHODES et le terme 1274 terminologie.

L'inconnue la plus gênante réside en fait dans le sort réservé au "vocabulaire fin" propre au Centre. Dans un tel cas de figure, le recours systématique aux relations associatives n'est plus envisageable. Pour les identificateurs, éléments contenus dans les champs autres que "Thèmes" et "Méthodologie", mieux vaudra avoir recours

à des listes communes d'autorités, utilisables pour l'ensemble des participants à la base, et issues de la refonte des listes particulières; de telles listes, et des fichiers-matières performants existent d'ailleurs déjà à l'extérieur de la base.

Quant au vocabulaire fin contenu dans les champs (T) et (S) - c'est-à-dire en général dans les champs possibles à structurer et à intégrer dans le champ "Thèmes", nous imaginerons deux possibilités de traitement:

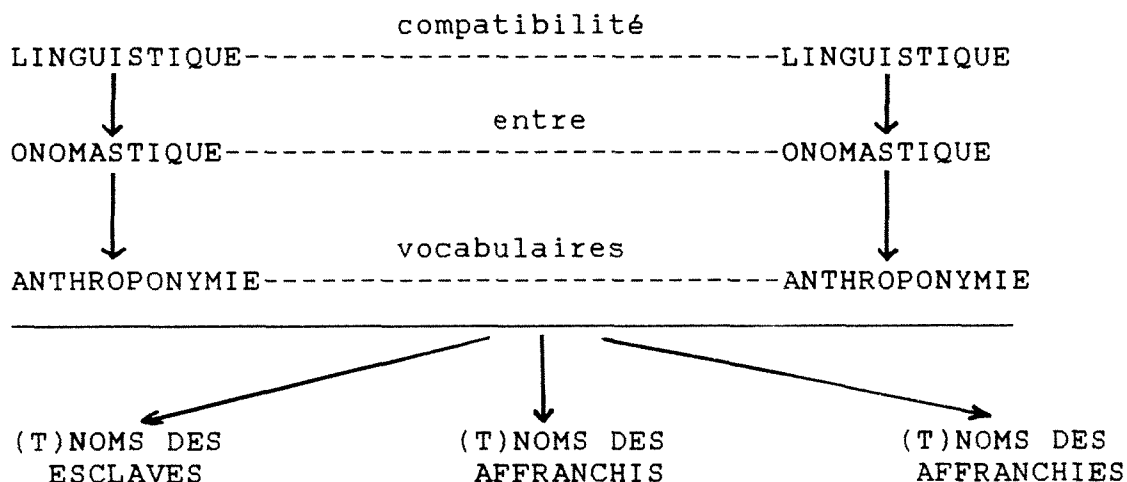
### La hiérarchisation du vocabulaire fin:

La réutilisation des données précédentes nous donne le nouvel exemple suivant:

#### SCHEMA N°2:

vocabulaire  
extérieur:

vocabulaire  
local:



(A)ALEXIS=esclave  
etc...

(A)ANTIGONUS=affranchi  
etc...

(A)ACTE=affranchie  
etc...

Le travail à réaliser pour y atteindre se résume à la perpétuation des structures hiérarchiques au-delà des troisième et quatrième niveaux; le gonflement des pyramides hiérarchiques risque de rendre le thésaurus moins maniable, car il enregistrera en conséquence une augmentation des relations de toutes sortes, et en particulier des cas de polyhiérarchies: en effet les termes fins les plus rares du dernier niveau devront dépendre de plusieurs hiérarchies (par exemple ici les noms de personnes des hiérarchies ONOMASTIQUE, mais aussi LINGUISTIQUE COMPAREE et LEXICOLOGIE. Ceci nous semble un travail énorme et fastidieux pour un intérêt limité en ce qui concerne la communication: les non-spécialistes et interrogateurs extérieurs seront gênés par cette quantité de données très "pointues", les spécialistes feront autant leur profit de listes non-structurées.

On n'évite pas non plus le danger déjà signalé de redondance de l'information: le mot-clés NOMS DES ESCLAVES aura de nombreux parallèles comme TERMINOLOGIE DES ESCLAVES, CONDITION DES ESCLAVES, DEFINITIONS DES ESCLAVES, etc... (plusieurs dizaines de formes au total).

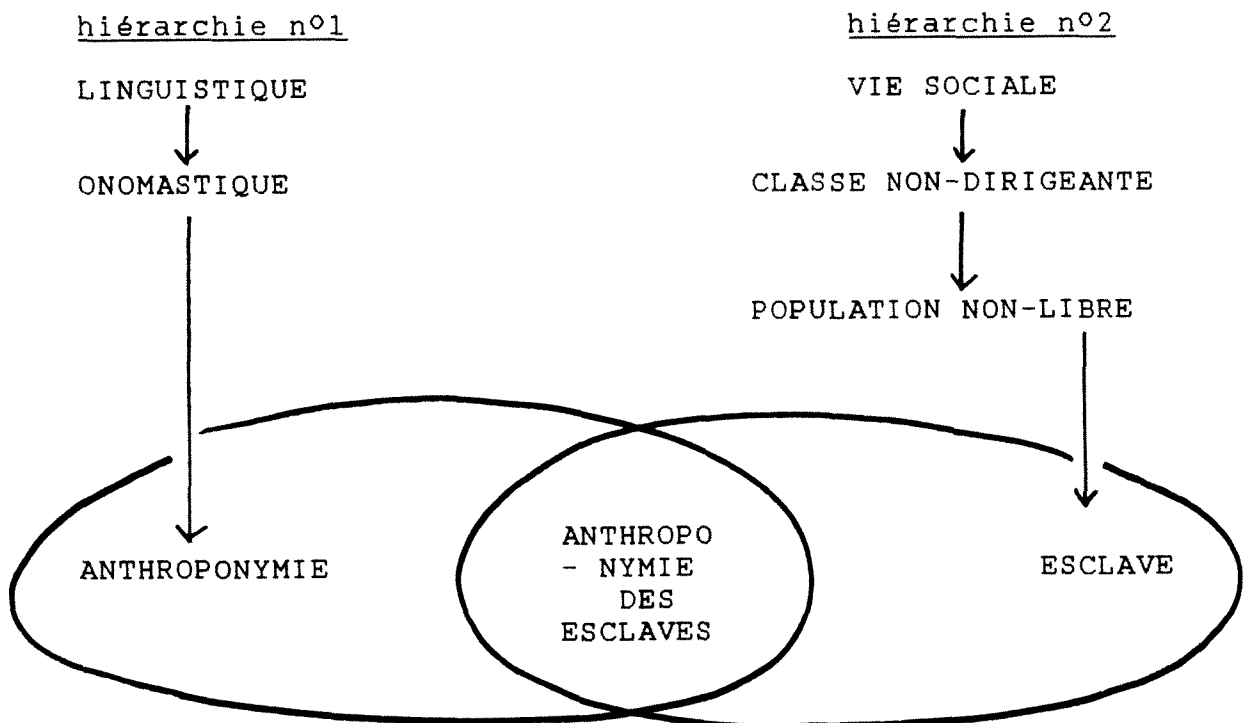
Il faut reconnaître cependant que ce système possède une précision remarquable et permet de conserver une exhaustivité et une spécificité d'indexation très poussées - mais à quel prix ?

Restructuration du niveau fin et exploitation des possibilités informatiques:

C'est l'optique que nous avons privilégiée dans ce rapport et qui a constitué notre hypothèse de travail. Le vocabulaire fin disparaît du thésaurus proprement dit, ses références sont rangées sous des rubriques plus larges qui correspondent aux descripteurs du dernier niveau et permettent des interrogations qui font systématiquement appel aux opérateurs booléens. Ainsi par exemple, les références des anciens mots-clés NOMS DES ESCLAVES et NOMS DES AFFRANCHIS seront-elles rangées respectivement sous les descripteurs: ANTHROPONYMIE et \*ESCLAVE (descripteur à créer), et ANTHROPONYMIE et \*AFFRANCHI; à terme, les relations d'équivalence correspondantes devront s'écrire: NOMS DES ESCLAVES: employer ANTHROPONYMIE et ESCLAVE; NOMS DES AFFRANCHIS: employer ANTHROPONYMIE et AFFRANCHI.

Nous expliciterons cette proposition à l'aide du schéma suivant:

SCHEMA N°3:



Cette technique se distingue évidemment par son caractère économique, mais pour la rendre applicable encore faut-il s'imposer une discipline consistant à indexer chaque passage avec un nombre limité de termes et

à réfléchir sur le niveau de précoordination souhaitable; sinon, danger déjà évoqué précédemment, l'interrogation nous fournira un "bruit" considérable, nous orientant vers des documents non-pertinents. Pour ce faire nous préconisons deux types de règles:

- une limitation impérative de la **spécificité** de l'indexation: nous avons vu qu'il est possible de décomposer le mot-clé NOMS DES ESCLAVES en deux concepts distincts; de la même façon, un descripteur NOMBRE DES ESCLAVES CORINTHIENS ou RESEAUX THEMATIQUES DES ESCLAVES ne nous semble pas viable pour une telle utilisation; par contre la question peut se poser pour un certain nombre de concepts qui risquent de perdre leur substance sémantique en subissant une telle décomposition: pour ceux-ci il convient de se demander s'il faut les inclure à notre thésaurus, ou s'il ne vaut pas mieux les introduire sur une liste annexe de vocabulaire fin.

- une limitation de l'**exhaustivité** de l'indexation: dans cette optique, il n'est pas utile, comme dans un index, de repérer tous les termes techniques, concepts historiques et sujets annexes contenus dans un article; il s'agit bien davantage de repérer le sujet central du travail et d'indexer à ce concept: ainsi ne peut-on utiliser les descripteurs ESCLAVE et ANTHROPONYMIE que pour l'indexation d'un article consacré essentiellement à la question du nom des esclaves et tentant une systématisation dans le domaine de l'anthroponymie(1).

---

1. C'est le cas par exemple pour l'article de O. Masson "Le nom des esclaves dans la Grèce antique" (fiche GIREA N°11), qui permet en outre de repérer une quantité d'anthroponymes serviles à intégrer dans la liste (A). Cependant ailleurs on évoque généralement les noms des esclaves ou des affranchis de façon rapide, comme un point parmi d'autres dans un raisonnement, sans accorder à ces remarques la valeur d'un système: il serait abusif de parler à leur propos d'étude anthroponymique. Cela vaut pour la plupart des travaux que nous avons examinés; nous pouvons citer entre autres exemples:

M.-P. CARRIERE-HERVAGULT - Esclaves et affranchis chez les orateurs attiques (fiche n°13). Un tableau relevant les noms propres, sur une quarantaine de page, une et demie consacrée à l'identification des personnages. Pas d'analyse sur la formation ou la signification de ces noms.

STAERMAN E. M. - L'étude de l'esclavage dans les provinces romaines (fiche n°140). Sur 11 pages, une remarque de deux phrases sur les noms des esclaves provinciaux.

MACTOUX M.-M. - Le système sémantique des termes désignant les esclaves chez Aristophane (fiche n°182). une vingtaine de lignes sur 40 pages.

Cette "suppression" du vocabulaire fin et ces économies en spécificité et en exhaustivité de l'indexation doivent en fait aboutir à la rédaction d'une liste relevant le vocabulaire très fin, propre à l'équipe locale et permettant, comme un index d'ouvrage, de répertorier de brèves mentions de concepts.

L'avantage le plus évident de cette solution réside dans son aspect cohérent et rapide pour une efficacité maximale: elle accélère les opérations d'indexation et permet une communication plus aisée de l'information, en particulier en direction des non-spécialistes; elle constitue également une expérience nouvelle en matière de vocabulaire d'échange dans le domaine des sciences de l'Antiquité.

### 5.3. Conclusion et remerciements:

Selon nous, le Centre doit définir sa politique documentaire en matière de thésaurus en fonction de deux considérations:

- l'importance accordée à la communication documentaire automatisée: celle-ci permettrait surtout l'exportation de l'information à l'extérieur, c'est-à-dire vers un public de non-spécialistes, d'étudiants "avancés", de chercheurs européens. Or cette perspective ne présente pas nécessairement, de l'avis général, un caractère d'urgence: il existe dans le domaine de l'esclavage antique des rencontres régulières, notamment par le biais des structures internationales du GIREA qui semblent assurer de manière satisfaisante l'échange d'information entre spécialistes; beaucoup semblent emprunter le canal personnel: personnalités-ressources, "collège invisible". L'observation des interrogations de la base locale, notamment pour la recherche bibliographique à l'intention des étudiants, nous a montré que celles-ci se réalisent dans la grande majorité des cas en présence de l'indexeur: nous sommes en face d'un système très personnalisé où l'interrogateur est à la fois indexeur et chercheur.

---

KOLENDO J. - Intellectuels et couches serviles: le cas des grammairiens chez Suétone (fiche n°194). 15 lignes sur la signification de certains noms de grammairiens affranchis ou esclaves, sur un total de 7 pages.

Dans tous ces cas de figure, il nous semblerait utile non pas d'indexer à l'aide du descripteur ANTHROPONYMIE, mais d'introduire les divers noms propres relevés dans la liste (A), ce qui permettrait un repérage fin destiné aux spécialistes.

- les disponibilités du Centre pour s'adjoindre les services d'un spécialiste en techniques de la documentation pour une période assez longue (environ 1 an): il semble en effet que cela corresponde au temps nécessaire pour la structuration complète de la section (T) des mots-clés du GIREA.

Quelle que soit la décision finalement adoptée par le Centre, nous voudrions terminer ce rapport en remerciant pour la cordialité de leur accueil et pour leur aide:

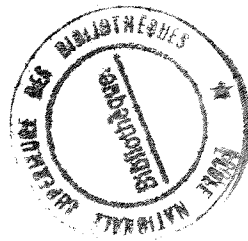
- au Centre de Recherches d'histoire Ancienne:

Marguerite GARRIDO-HORY  
Pierre LEVEQUE  
Marie-Madeleine MACTOUX  
Christine PEREZ

- à la Maison de l'Orient:

Michèle DALONGEVILLE  
Catherine DIEDRICHS  
Bruno HELLY  
Françoise LEROUGE

ainsi que: Brigitte LEQUEUX dont le travail m'a servi de base de travail, et Richard BOUCHE, dont les remarques et les conseils pratiques m'ont été indispensables.





\* 9 5 4 0 2 8 5 \*