

SÉMASIOLOGIE / ONOMASIOLOGIE : LA VOIE ARABE DE LA LEXICOGRAPHIE¹

Daniel Reig

Résumé

Après un bref aperçu lié à la relation entre le mot et l'idée, l'auteur aborde deux démarches principales dans le domaine de la lexicographie arabe : la sémasiologie et l'onomasiologie en s'appuyant principalement sur la première démarche.

Cette étude résume les différents types de présentation du lexique (d'une manière totale ou partielle), de l'ordre (phonétique ou alphabétique) et la nature de la classification (directe ou inverse) et ce à la lumière des données diachroniques et des commentaires synchroniques.

Mots clés

Lexicographie - sémasiologie - onomasiologie - dictionnaires arabes - dérivation.

¹ Cet article a été remis à la rédaction de la revue au début des années 90.

الملخص

إثر تلميح مقتضب للعلاقة بين الكلمة والفكرة، تساجل هذه الدراسة الوجهة العربية للصناعة المعجمية، متعرّضة لمنهجين أساسيين : السيماسيولوجيا والأونوماسيولوجيا. مركّزة على المنهج الأوّل.

وتلخص الدراسة صور عرض المادة في المعاجم العربية، التي تراوح الطبيعة (إجمالية، تصنيفية)، وأساس الترتيب (صوتي، هجائي) وطبيعة الرصف (مطرّد، عكسي)، مع تلميحات تاريخية وتضمينات نقدية.

الكلمات المفاتيح

المعجمية - المعاجم العربية - اشتقاق - السيماسيولوجيا - الأونوماسيولوجيا.

Abstract

After a brief survey of the relationship between the word and the idea, this study deals with "Arabic lexicography", by presenting two main procedures : the semasiology and the onomasiology; with special emphasis on the former.

This research paper presents the different ways of presenting the Arabic lexicon (in a total or partial way), the order base (phonetic or alphabetical) and the nature of classification (direct or indirect) ; all this in the light of diachronic details and synchronic commentaries.

Keywords

Lexicography - semasiology - onomasiology - Arabic dictionaries - derivation.

Il n'y a de communication possible entre deux interlocuteurs que lorsqu'il y a langage commun. On ne peut « s'entendre », en effet, avec quelqu'un que si l'on « entend » ce qu'il nous dit et s'il « entend » ce que nous lui disons.

Telle phrase, telle suite de phrases est produite par un Destinateur qui, pour exprimer une idée se sert d'unités de langue, c'est à dire des propositions, des groupes de mots, des mots. En vérité c'est l'idée elle-même qui bat le rappel de ces unités au fur et à mesure de ses besoins, qui les fait se présenter, qui éventuellement les repousse, en choisit d'autres, et s'élabore elle-même en même temps qu'elle constitue le discours qui lui donne forme.

Cette même phrase, cette suite de phrases, est reçue alors par un Destinataire qui trouve dans cette forme la substance de l'idée exprimée. J'ajouterai que jamais un mot n'est perçu tout à fait de la même manière par deux personnes différentes, bien que, il faut le dire, ces personnes « s'entendent » toujours sur un minimum de sens, même s'il reste vague, imprécis, une sorte de nébuleuse sémantique. C'est sur ce minimum-là que prend appui toute « entente » linguistique, et c'est à ce minimum-là que je m'en tiendrai ici : quelqu'un exprime avec des mots de la langue une idée que comprend quelqu'un à travers ces mots.

Le mot est donc une interface située entre un destinateur et un destinataire.

Mais le mot est aussi, par voie de conséquence, à double face : ce que signifie le destinateur et ce qu'entend le destinataire.

+ + +

Dans toutes les langues s'est développé un jeu sur les mots qui illustre parfaitement des allées et venues de l'idée au mot, du mot à l'idée. Je pense à l'institution, car c'en est une, des mots-croisés. On ne pense généralement quand on évoque ce jeu sur les mots qu'à la personne qui joue et qui dit qu'elle fait des mots-croisés. Mais le mot « faire » est ici abusif puisque, en réalité, ce n'est pas elle qui les a « faits » ; elle, les « défait » plutôt, avec le sens de : elle en donne la solution.

On oublie donc que ces mots-croisés ont un auteur, un Destinateur pour reprendre les termes déjà utilisés, et le cruciverbiste-joueur en est le Destinataire.

Or qu'a fait le Destinateur? Il est parti du « mot » de la langue et s'est efforcé d'en donner une définition, laquelle définition doit mettre le Destinataire sur la voie du mot considéré, étant bien entendu que jamais une définition ne peut épuiser la totalité de l'inventaire de signification qu'un mot transporte avec lui. Mais les manques sont comblés, dans la communication par le contexte, et dans les mots croisés, par certaines marques que l'on peut qualifier de contextuelles : nombre de lettres, lettres figurant déjà dans la ligne, etc. Bref, le cruciverbiste-destinateur cherche une définition pour un mot. Le cruciverbiste-destinataire cherche un mot pour une définition.

Eh bien nous avons là, schématisé, une représentation figurée de ce qu'est la lexicographie, activité du lexicographe qui peut se situer soit du côté de ce que j'ai appelé le cruciverbiste-destinataire soit du côté du destinateur. En effet le lexicographe peut comme ce dernier partir des mots pour en donner une définition : il fera alors de la « sémasiologie » (de sémasia = signification), ou bien il partira comme le premier de la définition pour donner le mot qui lui correspond : il fera alors de l'onomasologie (de onomasia = désignation).

La première démarche, la sémasiologie, est celle qui est la plus universellement représentée dans la constitution des dictionnaires de langue.

La seconde démarche : l'onomasiologie, outre les lexicographes intéressent les linguistes, les rhétoriciens, les écrivains car elle débouche sur la constitution de dictionnaires de synonymes, de dictionnaires analogiques, de dictionnaires spécialisés, mais elle peut intéresser aussi les anthropologues, et d'une façon générale, tous les spécialistes qui désirent étudier, non pas la langue en soi, mais à travers elle, les manifestations des différentes cultures humaines.

+ + +

Je n'aborderai ici pour l'instant que la démarche « sémasiologique »

LA SÉMASIOLOGIE

Le lexicographe établit la liste des signes qui font l'objet de sa recherche et procède à l'analyse de chacun de ses signes en dressant l'inventaire sémique, c'est à dire en en donnant une définition.

1. La définition de type sémasiologique

1.1. L'unité de lexique dont on désire préciser la signification est rangée, dans l'ordre alphabétique, à la place qu'elle doit occuper, en tant qu'unité, et elle est accompagnée d'une définition chargée de l'expliquer. Mais le terme « définition » est lui-même à définir. En effet il pourra s'agir parfois :

1.1.1. d'une définition qui aura un statut scientifique, mathématique pour tout dire = principe, construit a priori, sur lequel repose l'élaboration et le fonctionnement d'une science.

1.1.2. Mais son statut pourra être aussi logique = proposition ayant pour objet de faire connaître exactement l'extension et la compréhension d'un concept.

1.1.3. Ce type même de définition pourra avoir une portée moins ambitieuse et une destinée plus générale, partant moins rigoureuse, pour signifier plus couramment : explication.

Explication du sens par l'énonciation des propriétés essentielles de l'être ou de la chose que le mot désigne. Dans le dictionnaire de langue la définition ne puise pas dans la cohérence d'un système raisonné ses propres raisons de fonctionner. Elle n'est plus un axiome, elle est un discours sur le mot. Et ce discours peut se présenter sous la forme d'une séquence, ou de plusieurs séquences, selon le caractère polysémique ou non du mot considéré.

1.2. Cette séquence peut, elle-même, être composée d'une simple unité ou de simples unités juxtaposées, comme lorsqu'il s'agit de synonymes ou de contraires.

Voici un exemple de ce type de définition que nous limitons volontairement pour rester dans le droit fil de notre réflexion :

ḍarb = naw' wa šakl ضرب = نوع وشكل

Sont posées là, en face du lexème-vedette (le signifiant en quelque sorte) des unités représentant le signifié, senties comme équivalentes, non pas en tant qu'unités mais en tant que totalité produite par addition de ces deux unités puisqu'on a éprouvé le besoin de donner deux termes en équivalence. Ces deux termes ne sont pas simplement jux-

Pour en avoir le cœur net je me suis adressé à un autre dictionnaire : voici ce que j'ai trouvé pour le même lexème : *daḥala* = دخل (je simplifie transformant en simple lexème ce qui est parfois un syntagme construit)

daḥala = ...*ḏidd* : *ḥaraġa* دخل ... ضد، خرج

ḥaraġa = *baraza* خرج = برز

baraza = *ḥaraġa* برز = خرج، ضد : دخل

On voit qu'il s'agit là de définitions tout à fait approximatives utilisant la synonymie, c'est à dire l'équivalence totale, ou partielle quand deux ou plusieurs termes sont choisis, pour tenter de cerner mieux le sens du lexème-vedette, chacun de ces termes étant sensé apporter quelque chose en plus ; on a vu quel cas il fallait faire de ce type de définition par genus proximum et différentiam specificam. Mais la définition peut aussi utiliser les contraires : on en a vu quelques exemples car la contrariété est le comble de la différence, dans une procédure syllogistique raccourcie, par effacement de la deuxième proposition :

daḥala est x
 x est le contraire de *ḥaraġa*
 donc *daḥala* est le contraire de *ḥaraġa*

Tous ces types de définition sont approximatifs mais, aussi il convient de le dire, libres. Car d'un dictionnaire à un autre, d'un lexicographe à un autre, on perçoit des différences. Dans le cas de la sémasiologie, la définition n'établit pas, et c'est à cela qu'il faut s'arrêter, elle n'établit pas d'identité mais une simple équivalence. Celle-ci étant fondée sur l'existence d'une relation sémique provisoire et plus ou moins étroite entre les deux plans mis en parallèle.

On voit bien que dans les cas que nous avons observés, la définition de la signification procède par expansion élastique : à un mot correspond un mot, deux ou plus, un syntagme, une phrase. L'expansion est fondée sur le principe de l'équivalence d'unités inégales. Une unité pouvant avoir comme équivalent toute une suite, en expansion, qui a elle-même, par symétrie, l'unité de départ comme équivalent. On imagine donc déjà ce que pourra être un autre type de classement où à un contenu en expansion correspondra une seule unité lexicale.

Il s'agit là d'un fonctionnement que l'on peut qualifier de métalinguistique et qui est inhérent au discours. Toute proposition puise en elle-même sa propre définition, sans recourir à des éléments extérieurs, se contentant chaque fois d'intervertir les plans devisés. Ce qui est signifiant ici, peut devenir signifié là et réciproquement.

1.4. Mais il est une autre sorte de définition. C'est celle qui tente d'expliquer le terme-vedette par un signifié qui, au moment du renversement métalinguiste dont je viens de parler, va puiser ses éléments constitutifs en dehors de la langue, dans le référent, qui est ici le monde naturel, tout en utilisant les propres mots de la langue.

al-ḥaḡar = ḡism ḡāmid tabī' ī ṣalb الحجر = جسم جامد طبيعي صلب
yusta' mal fi l-binā' يستعمل في البناء
al-ḡism = *kullu mā lahu ṭūl wa 'arḍ wa-'umq* الجسم = كل ما له طول وعرض وعمق

Dans ce cas, on voit que nous avons affaire à une définition construite un peu différemment de celles que nous avons vues précédemment : à un terme particulier (*ḥaḡar*) on substitue un terme général (*ḡism*) que l'on va tenter de définir par restrictions qualificatives successives. Ces restrictions sont toutes de l'ordre référentiel. Les premières comme catégories intrinsèques : (*ḡamād* = minéralité, *tabī'a* = naturité, *ṣalāba* = dureté) et la dernière comme une catégorie extrinsèque découlant des précédentes : la faculté d'être utilisée dans la société d'une certaine manière et dans un certain but, conditionnés ou impliqués par ses qualités intrinsèques : (*al-binā'* = la construction).

On entre là dans le domaine de la définition encyclopédique à laquelle aucun dictionnaire de langue ne peut échapper à cause de la nature même de la langue qui, en se parlant elle-même, ne peut que parler du monde.

Ce que l'on pourrait dire ici, pour conclure ce paragraphe, c'est que la meilleure définition que l'on puisse donner d'une chose, c'est la chose elle-même, car aucune définition n'est complète et n'épuise jamais l'inventaire sémique d'un mot. Comment mieux faire comprendre ce qu'est une pierre qu'en montrant une pierre ? D'où l'utilisation d'une iconographie souvent très riche. Mais cette utilisation devient, dans les dictionnaires bilingues, parfaitement injustifiée sur le plan scientifique car le retour vers la langue maternelle équivaut au recours au référent ; l'iconographie n'a pas d'autre motivation que décorative dans les bilingues où, en face du mot arabe *kalb*, par exemple, on a le mot français chien, dont le signifié est encore explicité par une photographie, généralement en couleurs, représentant un chien, qui n'est jamais, de fait, n'importe quel chien, mais toujours un chien particulier, alors que le concept « chien », lui, est général et abstrait.

On voit ici, au passage que de nombreux auteurs de dictionnaires bilingues arabes auraient souvent intérêt à réfléchir à ce qu'ils font, qui n'est pas seulement de la traduction.

2. Les dictionnaires arabes

J'évoquais à l'instant les dictionnaires arabes en leur empruntant des exemples de définition. Il s'agissait bien entendu de dictionnaires modernes. Il convient de voir maintenant comment la réflexion arabe s'est appliquée à appréhender le problème de la lexicographie et tout particulièrement de la lexicographie en tant que recherche sémasiologique c'est à dire dans un classement du lexique qui part de l'unité lexicale pour présenter sa / ses signification (s).

Nous retrouverons là bien entendu le classement alphabétique dont je disais qu'il était quasi universel, mais nous rencontrons aussi, et d'abord, dans l'histoire de la lexicographie arabe, une forme de classement beaucoup plus originale : le classement phonétique.

2.1. Classement phonétique

Il s'agit là de la première entreprise lexicographique arabe, due à un linguiste et grammairien du 8^e siècle, mort entre 786 et 791, al-Ḥalīl ibn Aḥmad. Je veux parler du *kitāb al-ʿayn* dont nous ne possédons plus qu'un abrégé et qui a la particularité d'être classé selon les points d'articulation des phonèmes arabes et commençant par les phonèmes les plus « rentrés » pour finir par ceux qui sont articulés, au bord extérieur de l'appareil phonatoire, par les lèvres.

La première lettre par laquelle commence le classement est une pharyngale sonore : le *ʿayn* d'où le nom donné au dictionnaire. Viennent ensuite le *ḥā'*, pharyngale sourde, puis le *hā'*, glottale sourde (il manque la *hamza*, glottale sonore) le *ḥā'*, vélaire sourde, le *ḡayn*, vélaire sonore, le *qāf*, post palatale sourde emphatique, le *kāf*, postpalatale non-emphatique. Viennent ensuite les palatales dont les points d'articulation se situent sur la voile du palais. Les dentales, les labio-dentales et enfin les labiales.

De ce classement, on peut tout de suite dire qu'il n'est pas, au moins pour les deux premiers phonèmes, tout à fait conforme à l'ordre phonétique, puisque normalement c'est par les phonèmes articulés par la glotte que l'on aurait dû commencer. Al-Ḥalīl en avait parfaitement conscience et s'il a choisi d'écarter les glottales, c'est en raison de l'instabilité phonétique de la *hamza* et du *hā'*.

À la suite du *kitāb al-ʿayn* et selon le même principe de classement phonétique sont élaborés des ouvrages destinés à compléter le travail du premier - ou considéré comme tel par la tradition arabe - maître lexicographe. Ce sont *Tahqīb al-luḡa* d'Al-Azhari, mort en 980, puis *al-Muḥkam* d'Ibn Sīda, mort en 1006, parmi les plus importants.

Parallèlement on se rend compte cependant qu'il n'était pas facile de manipuler des ouvrages importants, aux nombreux fascicules, et dans lesquels le classement du lexique obéissait à un ordre, certes très rigoureux et scientifiquement justifié, mais dont les principes organisateurs étaient extérieurs, de plus en plus si je puis dire et je m'en expliquerai, à l'ordre même tel qu'il apparaissait graphiquement manifesté dans cet objet concret qu'était le livre. Les clés de cette organisation à base phonétique résidaient dans l'oralité, alors que la société arabo-islamique évoluait - au moins parmi les classes de lettrés - de plus en plus dans le domaine de l'écrit. Période de transition donc où, la tradition audio-orale perd ses marques où, par contre, se met en place une véritable civilisation du livre. Cette période est comparable, toutes choses étant égales d'ailleurs et de manière symétrique et inverse, à celle que nous vivons dans ce dernier quart du xx^e siècle, et dans laquelle nous passons lentement d'une culture à tradition écrite à une culture de l'audio-visuel où, par un phénomène de bascule, réapparaît, en partie, la primauté de l'oral. Transition qui, comme toutes les transitions, est en même temps de tradition et de transformation.

La fin de ce huitième siècle au cours duquel on assiste à la conception de ce premier dictionnaire à base phonétique est marquée en effet par un extraordinaire phénomène d'acculturation.

D'une part la tradition culturelle, spécifiquement arabe (poésie, récits sapientiaux ou historiques, connaissances de type encyclopédique portant sur l'environnement minéral, végétal, animal ou cosmique de la bédouinité) et islamique (Coran, hadith, droit), s'étend, répandue par les tribus disséminées aux quatre coins de l'Empire, et se voit

fixer dans des recueils fort nombreux et divers, en tout cas extrêmement spécialisés et détaillés.

D'autre part, les peuples islamisés et qui s'arabisent tant bien que mal, introduisent dans le creuset de cette nouvelle civilisation, qui a commencé à s'élaborer dès les conquêtes, leur propre culture et leurs propres préoccupations intellectuelles.

Le patrimoine culturel arabe se fond donc là dans un amalgame dont la langue arabe devient le véhicule. Langue officielle d'une communauté aux multiples entités, l'arabe transmet aussi bien une culture d'origine hellénique, que persane ou indienne et arabe bien entendu, et elle prend forme écrite. Forme écrite dans la traduction, l'explication, le commentaire, la réflexion et dont le contenu est autant la religion que le droit, l'astronomie que la langue elle-même. Mais elle est aussi devenue parallèlement la langue d'une immense administration dont la vie quotidienne permet de soulever une multitude de questions pratiques qu'il faut résoudre de manière cohérente. Les nouvelles conditions de son existence entraînent, cela va de soi, de profondes modifications de son essence. Elle est en effet soumise à codification.

Le processus, déjà engagé probablement avant même l'avènement de l'islam, dans le contexte des joutes oratoires auxquelles se livraient depuis toujours les tribus, s'est considérablement amplifié depuis un siècle pour des raisons qu'il est facile d'imaginer et va se prolonger jusqu'au x^e siècle. Cette codification a pour aboutissement dans le domaine de la morphologie, qui nous intéresse ici indirectement, la constitution d'une doctrine cohérente et remarquablement pertinente et, dans le domaine de la transcription graphique qui nous intéresse directement, une fixation écrite qui va peu à peu amener les savants lexicographes à jeter les bases de nouveaux dictionnaires.

Ainsi y a-t-il concurrence à cette époque entre tradition orale tout particulièrement du patrimoine arabe, et codification écrite, en arabe, d'une culture multiforme. A la première appartient *le kitāb al-‘ayn* et les ouvrages qui lui ont succédé, dans la mesure où leur conception nous renvoie à un rapport à la langue essentiellement oral. A la seconde appartiennent tous les dictionnaires dont la conception s'est définitivement imposée à partir du x^e siècle et qui prennent l'ordre alphabétique comme base du rangement du vocabulaire étudié.

2.2. L'ordre alphabétique

Nous avons vu comment al-Ḥalil et d'autres lexicographes arabes ont marqué la présentation du vocabulaire par une réflexion métalinguistique extrêmement rigoureuse et volontariste, si je puis dire, puisqu'elle introduisait un ordre intellectuel et raisonné dans l'univers de la langue. Cet univers paraît tout à fait naturel, tant la pratique de sa langue maternelle, est intuitive et fait croire que « les mots sont les choses ». Et donc, si à l'époque extraordinaire de la collecte des matériaux lexicographiques, en gros les deux premiers siècles de l'Islam, le naturel était « audio-oral », de plus en plus, à partir du troisième siècle de l'Hégire (IX^e), ce naturel se fait graphique et visuel.

C'est cette utilisation naturelle et intuitive de la langue, permanente, et à laquelle le linguiste lui-même n'échappe pas, qui a permis à l'ordre alphabétique de s'imposer, chez les Arabes dès les X^e et XI^e siècles et en Europe à partir du XVI^e siècle au moment de la naissance de la lexicographie européenne.

Mais cet ordre alphabétique peut se présenter de différentes manières. Il peut être direct, le classement s'effectuant sur la base de la première lettre du mot, puis de la deuxième et ainsi de suite. Il peut aussi être inverse, le classement s'effectuant à partir de la dernière lettre ou de la dernière syllabe du mot.

Qu'il soit direct ou inverse ce classement peut avoir une caractéristique supplémentaire. Il est soit total c'est à dire qu'il prend en compte tout les mots de la langue, selon l'ordre direct ou inverse dans lequel ils se présentent, soit partiel. Dans ce cas, tous les mots de la langue sont regroupés à partir de la racine, comme dans une famille, et chaque famille est rangée ensuite selon l'ordre alphabétique des lettres qui composent la racine.

La lexicographie arabe a tout tenté et nous retrouvons à toutes les époques depuis le X^e siècle les différents modes de classement alphabétique que je viens d'évoquer.

2.2.1. Classement alphabétique total et direct (= à partir de la première lettre du mot)

Ce type de dictionnaire existe dans le domaine de la lexicographie arabe et il est inspiré par deux soucis différents :

a- Dans le premier cas on a affaire à des ouvrages composés dès l'époque la plus reculée de l'histoire culturelle arabe. C'est l'œuvre, généralement de grammairiens désireux de corriger ce qu'ils considéraient comme des fautes de langage chez leurs contemporains de telle ou telle métropole. Mais, et c'est cela qui est un phénomène intéressant, on a pu voir aussi des hommes qui n'étaient pas des grammairiens, des spécialistes de la réflexion métalinguistique, s'intéresser à ce type de problèmes. C'est le cas de abū Faraġ Ibn al-Ġawzī, théologien et moraliste, d'idéologie hanbalite, auteur du *Sayd al-hāṭir*, du *Damm al-hawā* et du célèbre *Talbīs Iblīs*, et qui a composé, en 1173, un ouvrage intitulé : *Taqwīm al-Lisān*.

Ibn al-Ġawzī se situe dans la lignée d'auteurs comme Ibn Qutayba, qui lui sert d'ailleurs de référence, dont le souci était d'opposer aux pratiques linguistiques fautives (phonétiques, morphologiques ou syntaxiques) du parler quotidien des gens cultivés de la Bagdad du VI^e siècle (XII^e siècle), le « bon usage » de la 'Arabiya. Il ne s'agit du peuple, de la 'amma, mais bien d'intellectuels (*al-muntasibūn ila l-'ilm*) dont la pratique orale - c'était des orateurs, des sermonnaires - était entachée de scories, celles-là mêmes qui caractérisaient la langue parlée à Bagdad à cette époque et que l'on qualifiait de « populaire », parce que le peuple ne connaissait qu'elle.

Ibn al-Ġawzī classe donc les corrections qu'il propose selon l'ordre alphabétique total c'est à dire en commençant par la première lettre du mot. Ainsi le premier mot dont il s'occupe se situe-t-il au chapitre du *alif* (: *bāb al-alif*) :

« On doit dire « *ustuhtira fulānun bi-kadā* » avec voyelle u accompagnant le premier *tā'*, et i accompagnant le deuxième, car il s'agit d'un passif (*mā lam yusamma fā'iluhu*). Le peuple prononce ce mot, avec le même sens, en faisant accompagner les deux *tā'* de la voyelle a, ce qui est fautive. »

Ainsi de suite...Le classement alphabétique total est ici à vocation pédagogique immédiate (= la correction de fautes) et fonctionne un peu comme un dictionnaire bilingue puisque pour le même sens nous disposons de deux formes : la forme « classique », correcte, et la forme « dialectale », fautive.

b- Mais il est un autre type de dictionnaires dans lesquels le vocabulaire est rangé selon l'ordre alphabétique total, c'est celui que l'on a pu voir se publier ici ou là à l'époque contemporaine. L'objectif visé par les auteurs est d'éviter aux élèves de l'enseignement primaire et secondaire, ainsi qu'aux étudiants et, en fait, à tout les usagers, les fatigues et les difficultés d'une recherche qui doit, effectivement, s'appuyer sur un minimum de connaissances « linguistiques ». Ces auteurs ont sacrifié le sentiment linguistique à la facilité et à l'ignorance. Mais je reviendrai plus loin, en détails, sur ce problème.

2.2.2. Classement alphabétique inverse : (à partir de la dernière lettre, ou de la dernière syllabe de mot)

Tous les mots se terminant par la même lettre se trouvent rassemblés dans le même chapitre.

Mais il y a deux approches. L'approche arabe telle qu'elle se manifeste dès le X^e siècle et l'approche européenne avec le premier dictionnaire inverse connu en Europe, le dictionnaire védique publié en 1873.

a- Les dictionnaires inverses arabes

Le premier grand dictionnaire de ce type est l'oeuvre d'al-Ġawhari mort en 1006 ; il s'agit du célèbre *al-Šihāh*. Le classement y est établi à partir des racines, les entrées-bases étant les mots dépouillés de tout leurs augments ; on arrive ainsi au classement alphabétique partiel déjà rencontré, mais inversé, puisque les racines y sont ordonnées à partir de la dernière radical : d'abord tous les mots terminés par *alif*, puis *bā'*, puis *tā'* etc... Le classement interne s'effectuant à nouveau selon l'ordre normal, à partir de la première lettre du mot rapportée à l'ordre alphabétique puis de la deuxième.

Ainsi *qara'a* se trouvera classé à *alif* (*hamza*) puisque c'est la dernière lettre du mot, puis, à l'intérieur du *alif* il faudra chercher *qāf* puis *rā'*.

Par contre *qaraba* se cherchera sous le *bā'*. Les dérivés de toutes ces entrées sont rangés alors selon un ordre logique et non pas par ordre alphabétique. Obéissant à une logique que l'on peut qualifier souvent d'intuitive car les principes qui la guident ne sont pas, généralement, définis.

D'autres dictionnaires viendront, sur ce modèle, compléter le *Šihāh*. Le plus célèbre, à juste titre, d'entre eux est le *Lisān al-'arab* composé en 1311 par Ibn al-Manḍūr et qui emprunte beaucoup à ses devanciers, d'où l'immense intérêt de cette somme.

b- Les dictionnaires inverses européens

Européens de conception, en réalité ils s'occupent de langues non européennes. C'est le cas d'un Dictionnaire védique publié à Leipzig en 1873, puis d'un dictionnaire de latin et de vieil iranien. Les entreprises les plus remarquables ont été réalisées soit sur des langues mortes, soit sur des langues slaves.

c- L'intérêt du classement inverse

La première utilité de ce type de classement est, tout naturellement, de faire office de dictionnaire de rimes sans parler des secours qu'il peut apporter aux décrypteurs dans leur travail de décodage de messages secrets qu'il s'agit de reconstituer. A condition

bien entendu, quand il s'agit du classement inverse des racines, que les règles de la dérivation morphologique soient relativement simples et rigoureuses comme c'est effectivement le cas en arabe. En effet, le classement alphabétique partiel - classement par ordre alphabétique des racines- suppose connus toutes les possibilités offertes par le système déverbatif et dénominatif de la langue. Comme dans le cas du classement direct des racines, il faut ici aussi procéder à la soustraction des segments pour retrouver dans l'adresse convenable le(s) mot(s) que l'en recherche. Et, inversement, il faut imaginer, en partant d'une racine donnée toutes les possibilités offertes par le système dérivationnel pour recréer par insertion d'augment un lexique actualisable. D'autre part, il semble bien que la recherche dans un dictionnaire inverse soit simplifiée en arabe car la suffixation est relativement peu diversifiée dans le système de l'arabe, alors que la dérivation par préfixation ou infixation ou les deux simultanément est beaucoup plus développée. On est à peu près toujours sûr, en cherchant dans un dictionnaire inverse arabe, de tomber sur la bonne racine. En tout cas, c'est cette simplicité de la recherche qui a animé l'école lexicographique pratiquant la méthode inverse et dont le chef de file a été al-Ġawhari. Mais en fait, les dictionnaires par classement inverse des racines n'offrent pas beaucoup d'intérêt lexicologique car ils ne sont pas fondés, justement, sur des oppositions morphologiques produites par le système dérivationnel. La dérivation arabe s'effectue par préfixation, infixation, et aussi, mais moins, par suffixation. Dans le classement inverse total, ce sont, évidemment les suffixes qui apparaissent. Pour l'arabe les suffixes sont surtout des marques du féminin (at, ā' ou ā), ou de pluriels (ūn et āt) et généralement les termes de ce type ne sont pas notés dans les dictionnaires, sauf exception, de même qu'en français on ne notera pas les formes du féminin ou du pluriel, celles-ci étant supposées couler de source.

Par contre dans le cas des langues slaves ou romanes, la suffixation joue un rôle extrêmement important, et utile pour l'étude comparatiste des langues de même famille : il est parfaitement possible de suivre les processus de transformation qui se sont effectués sur certains suffixes à travers différentes langues issues de la même langue-mère. On voit immédiatement l'intérêt que cela peut présenter, sur le plan lexicologique pour les langues romanes par exemple : le roumain, l'italien, le romanche, le français, le portugais et l'espagnol et tous les dialectes parallèles à chacune de ces langues qui ont un système suffixal important, peuvent faire l'objet d'études extrêmement riches et fondées sur la comparaison des suffixes.

Donc pas nécessairement destinés à l'utilisation quotidienne, les dictionnaires inverses peuvent cependant être d'une grande utilité dans le domaine de la recherche linguistique (phonétique et phonologique), lexicologique et probablement aussi anthropologique.

La mécanisation du travail, il va sans dire, peut rendre ici de grands services et même directement, à partir du fichier général, introduit dans la mémoire de machine par ordre alphabétique normal, on peut imaginer des procédures de reclassement et redistribution des données qui éviteraient toute manipulation.

2.3. Avantages du classement alphabétique total

Quels sont les avantages et les inconvénients du classement alphabétique total normal ou inverse. Ce classement est visuel essentiellement. Le fait que l'on soit obligé

de partir du signifiant sous sa forme graphique et selon son orthographe complète, pour arriver au signifié, va conditionner l'ordre dans lequel les signes doivent apparaître. C'est à partir de ceux-ci que s'effectue la recherche de l'utilisateur : il faut donc qu'ils aient une position stable, acceptée de manière générale sinon universelle. C'est la position par rapport à l'ordre alphabétique qui correspond le mieux à cette définition. En effet, les mots sont pris là comme des objets en dehors de toute autre considération que leur forme même. La recherche s'effectue visuellement, c'est la forme graphique qui est donc ici concernée, et telle qu'elle est définie par l'usage traditionnel.

En dernière analyse, cette forme est arbitraire, et c'est ce qui fait sa force : on a affaire à de véritables objets orthographiques, car leur réification rend possible l'installation des mots dans une grille de relations positionnelles, à partir de la relation à l'alphabet de chacune des lettres qui les composent sans dépendre en aucune façon de considérations théoriques à priori, et subjectives, au tout cas abstraites, qui pourraient ne pas être acceptées de manière universelle. En effet, dans tout les alphabets, les lettres occupent des positions stables : ainsi dans l'alphabet latin A est avant B, B est avant C et après A, C..., Z est après Y. Ensuite, conformément à cet ordre, les mots trouveront leur place, les uns par rapport aux autres, chacun se situant avant ou après tel et tel autre. Le classement alphabétique est une sorte de métaphore du classement numérique qui ressortit à une vision totalement abstraite et arbitraire de l'organisation, c'est pourquoi ce type de classement est universel. C'est pourquoi aussi, il peut rendre des services considérables à l'ère des ordinateurs.

En effet, l'introduction de dictionnaires alphabétiques dans des machines et les possibilités immenses de stockage, de classement, de repérage et de sélection dont celles-ci disposent, permettant de procéder dans le fichier général à des opérations de tri qui sont pratiquement illimitées et qu'il était difficile d'envisager quand ces opérations étaient manuelles... Mais pour parvenir à faire donner à la machine les aptitudes de réponses que l'on attend d'elle il a fallu, au préalable, que les données qui y sont entreposées soient établies sur une invariance de classement, et sur sa simplicité, peu importe que tel ou tel type de classement soit au départ de cette invariance.

Il est vraisemblable que n'importe quel classement imaginable, normal ou inverse, pourvu qu'il fût rigoureux, aurait pu être utilisé, les machines s'arrangeant, grâce à leur programmation, des données qui leur sont fournies.

2.4. Inconvénients du classement alphabétique total

Les avantages, considérables, que nous venons d'évoquer, d'un tel classement alphabétique pour une utilisation mécanique, ont conduit les chercheurs à rêver d'utiliser ces mêmes données dans le cadre de la traduction. Nous étudions par ailleurs le problème de la traduction automatique : ce que je veux dire ici, pour être bref, c'est que la traduction n'est pas un transfert de forme, de mot à mot, de locution à locution, et ainsi de suite, mais un transfert de substance. C'est à dire que partir du classement alphabétique des mots pour essayer de traduire des sens c'est se heurter à des difficultés que les chercheurs sérieux et honnêtes ne se cachent pas. Et s'imaginer que l'on pourra, à partir d'un classement sémasiologique (dans quelqu'ordre qu'il se présente), résoudre les problèmes de la traduction, relève de l'utopie.

Le classement alphabétique, d'une façon générale, est spécifique à chaque langue puisque l'ordre dans lequel les alphabets se présentent n'est pas nécessairement le même. Et dans le cas où ces alphabets tirent leur origine des mêmes sources (langues européennes qui utilisent l'alphabet latin, l'arabe et le persan, etc.). Le traitement phonologique de certaines lettres dans chacun des systèmes est tel que l'on a ou bien des ressemblances abusives ou bien des dissemblances trompeuses. Un classement à base phonétique pourrait peut-être faire ressortir davantage les parentés entre les langues d'une même aire linguistique. En effet, le seul principe organisateur d'un tel ordre est la forme graphique du mot qui n'a aucune réalité véritablement linguistique puisqu'elle est fondée sur une convention et que, dans la plupart des cas, surtout dans les langues européennes, elle ne correspond pas à l'exacte prononciation du mot.

Mais c'est d'ailleurs le cas en arabe. Pour prendre un exemple, le mot *infa'ala* (انفعل) qui se prononce avec le phonème *i* à l'initiale est à chercher, dans ces dictionnaires alphabétiques total, à la lettre *alif* (a) et non à y, parce que malgré les apparences, la lettre *alif* n'est pas prise en compte dans le classement alphabétique arabe ; c'est la *hamza*, en réalité première lettre de l'alphabet qui sert de repère. Cela a l'air bien confus. En effet, il faut savoir que, à l'initiale, la *hamza*, quelle que soit la voyelle brève dont elle permet l'articulation, s'orthographie toujours avec un *alif* pour support. Dans l'exemple que j'ai pris de *infa'ala*, ce que le classement alphabétique retient ce n'est pas la *hamza*, pharyngale sonore, c'est son support graphique, elle-même étant le support phonique de la voyelle *i*.

Je dis bien support phonique car les grammairiens arabes ont bien pris soin de distinguer entre *hamzat al-waṣl* et *hamza al-qat'*. Ce que l'on appelle *hamzat al-waṣl* a pour vocation de permettre l'articulation, à l'initiale, d'une syllabe longue fermée dont il n'existe que la fermeture. Pour pouvoir ouvrir cette syllabe, avant sa fermeture, il faut une voyelle, et aucune voyelle ne peut s'articuler sans une attaque glottale dans les conditions phonatoires de l'articulation de la *hamza*.

On voit que les chemins qui nous conduisent au déchiffrement d'un dictionnaire à classement alphabétique total sont très détournés et qu'il faut pouvoir déjouer les pièges de l'orthographe, c'est à dire des conventions qui ne trouvent leur justifications que dans la longue tradition qui continue de les imposer.

Les avantages offerts par le classement sémasiologique, sont considérables tant qu'il s'agit d'effectuer des recherches scientifiques dans les domaines linguistiques, lexicologiques, statistiques, etc. Mais sur le plan pédagogique général, c'est à dire quand il s'agit d'affiner et d'approfondir le sentiment métalinguistique des usagers d'une langue, sentiment métalinguistique tout à fait naturel, ce classement alphabétique total des dictionnaires devient, comme le disait déjà Vaugelas le dernier de tous les ordres car il agit dans l'esprit des usagers comme le pire des désordres.

En effet, le reproche majeur que l'on peut adresser au classement alphabétique total est que ce dernier fait éclater l'unité de la langue considérée. En effet, toute langue s'enrichit, se développe bref, vit, grâce à ce que l'on peut qualifier de permanences génériques autrement dit, de régularités d'ordre morphosémantique et morphosyntaxique. Ces régularités conditionnent la génération, toujours possible, de mots nouveaux, en assurant la cohérence de l'ensemble, grâce au système dérivationnel stable qui

permet ainsi à la langue de s'adapter aux différents besoins socio-culturels qu'elle éprouve au fur et à mesure de son développement.

L'extraordinaire sentiment de pérennité, de cohérence d'unité que l'on éprouve quand on observe la langue arabe malgré ses quinze siècles d'existence et la relation que les Arabes entretiennent avec leur langue, malgré des destinées socio-culturelles souvent différentes, tient sans aucun doute davantage à la nature du système morpho-sémantique de l'arabe qu'à la volonté des hommes politiques ou religieux. Plus qu'à ceux-ci, en effet, le mérite d'avoir maintenu cette cohérence, cette unité de la *'arabiya* en revient aux linguistes : lexicographes et grammairiens qui ont œuvré considérablement, à toutes les époques, pour conserver à la langue sa cohérence, et mettre en lumière ses régularités.

Or le classement alphabétique total semble poser comme axiome que tout, dans la langue, est irrégularité. Les unités lexicales qui constituent le vocabulaire ne sont plus perçues à travers des convergences, des cohérences, des régularités et le lexicographe semble être impuissant à maîtriser la formation des entités linguistiques bref, à rendre compte d'un système organisateur.

Il existe, on l'a vu, en arabe deux types de dictionnaires à classement alphabétique total.

Le premier, historiquement parlant, puisqu'il remonte au XII^e siècle a une forte justification linguistique au sens scientifique, même s'il est plus l'œuvre d'un théologien et moraliste que celle d'un linguiste, comme on l'a vu (2-2-1-a).

Le second type de dictionnaire alphabétique total, tout à fait contemporain, n'a aucune justification scientifique. Les motivations sont démagogiques, pour ne pas dire commerciales, car le seul argument que les auteurs aient pu trouver pour se justifier, est celui de la simplicité de la consultation. En fait il s'agit de favoriser la paresse du lecteur même si cela doit faire éclater le lexique arabe. L'ordre alphabétique total contribue à déstabiliser les unités lexicales et, au-delà peut-être, à porter atteinte à son système de dérivation pourtant clair et simple pour lui ôter, enfin, sa faculté créatrice.

Le seul argument qui ait été avancé pour défendre l'alphabétisation totale des dictionnaires arabes est la simplification et l'immédiateté de la recherche. On n'aurait plus, ainsi, à reconnaître, par exemple, dans le verbe l'une des dix formes dérivées encore usitées puisque celles-ci sont rangées selon l'ordre alphabétique de leurs constituants : III kātaba à k,ā,t,b VI takātaba à t,k,ā,t,b VIII / (i)ktataba à a,k,t,t,b -X (i) staktaba à a,s,t,k,t,b.

Il y a une première contradiction. Ces verbes sont rangés, dans les dictionnaires à classement alphabétique total, sous cette forme que je viens de rappeler et qui est une entrée métalinguistique en même temps que la 3^e personne du masculin de l'accompli. Mais pour retrouver : X (i)staktaba à a,s,t,k,t,b à partir de yastaktibūna (y,s,t,k,t,b) par exemple il faut avoir appris les conjugaisons arabes et être capables d'écarter les préfixes et suffixes personnels et de réintroduire le *alif* formant graphique mais non phonique (puisque'il est prononcé /i/ ou supprimé par aphérèse). Le rangement alphabétique ne produit donc pas une économie totale des connaissances morphologiques.

Il y a une deuxième contradiction. Dans les entrées de tous les dictionnaires, le nom désignant l'action effectuée par le verbe (*maṣdar*) est toujours exprimé à la suite du verbe dans un ensemble morphosyntaxique clairement identifié. Or la première lettre de

ce nom d'action n'est pas toujours la même que celle du verbe et l'on sait par ailleurs qu'il est aussi normalement productif que le verbe sans pour autant se figer dans une lexicalisation définitive. Il faudrait donc que le dictionnaire à rangement alphabétique total, pour être conséquent, donne à leur place alphabétique normale tous les noms d'action dont la lettre initiale diffère de celle du verbe : ainsi *taf'īl* devrait se chercher à : *t, f, 'ī, l* et non à *f, 'ī, l* et *mufā'ala* devrait se chercher à *m, f, ā, 'l* et non à *f, ā, 'l* comme c'est le cas. En effet, si l'on a des chances de trouver quelques noms d'action de ces types à leur place alphabétique c'est dans une proportion infime et presque uniquement pour les noms d'action lexicalisés avec le sens de « résultat de l'action exprimée par le nom ». Les auteurs de ces dictionnaires considèrent donc le verbe comme la seule entrée métalinguistique et pensent, à juste raison, que les utilisateurs connaissent suffisamment bien la morphologie de l'arabe pour s'y adresser directement.

Ce que je viens de dire des noms d'action des formes dérivées ou de la forme simple peut l'être de tous les participes présents et passés de ces mêmes formes puisque leur initiale est différente de celle du verbe.

Mais il y a plus encore. Les formes de pluriel, dont certaines sont extrêmement productives, sont, elles aussi, systématiquement écartées de l'ordre alphabétique au bénéfice du singulier. On voit donc que pour une méthode qui désire simplifier les choses, voilà effectivement beaucoup de laissés pour compte.

Ce qu'il faut retenir ici, c'est que l'on ne peut de toute façon échapper aux contraintes morphologiques d'une langue et que ce n'est pas en les fuyant qu'on les élimine. Il faut tout simplement en donner une vision claire. Mais nous touchons là au problème de la didactique des langues, de celle de l'arabe en particulier. Et c'est, presque, un autre problème.

A une époque où toutes les réflexions scientifiques comme les sensibilités nationales aspirent à reconstituer des systèmes linguistiques cohérents et organisés à partir de quelques règles morphosyntaxiques et sémantiques simples, la position adoptée par les rédacteurs de dictionnaires arabes à ordre alphabétique total apparaît comme totalement rétrograde sinon réactionnaire.

Les premières discussions à ce sujet ont débuté vers les années 1930 alors que certains arabes s'imaginaient que la seule chance de réussite était l'imitation servile de l'Occident et de ses manifestations culturelles. Ce sont d'ailleurs les mêmes qui préconisaient la transcription de la langue arabe en caractères latins... Combien plus imaginatifs, plus créateurs, plus progressistes ont été leurs ancêtres des VIII^e, IX^e, X, et XI^e siècle qui nous ont légué un instrument théorique unique en son genre et toujours actuel. De plus en plus actuel, ajouterai-je.

3. Classement alphabétique partiel (direct ou inverse)

Il s'agit d'un classement alphabétique des racines que les premiers linguistes arabes ont mis en place dès le VIII^e siècle et qui est fondé sur des principes et des analyses que la linguistique scientifique contemporaine, américaine et européenne, rencontre seulement depuis quelques dizaines d'années par un cheminement tout à fait différent.

Dans la description qu'ils ont faite de leur langue au Moyen Age, les linguistes arabes ont donc découvert une loi de cohérence morphosémantique extrêmement forte à laquelle ils ont donné le nom de *istiḳāq* (اشتقاق) et que l'on résumera ainsi : la dérivation

à partir d'un mot-souche est toujours prévisible, les schémas organisateurs des possibilités dérivationnelles étant en nombre restreint. D'où une motivation très puissante qui tient à ce qu'est posé, avec la racine, un champ morpho-sémantique informel que toute la dérivation va ensuite découper selon un certain type de régularités et ordonner selon la règle, fondamentale en linguistique, qui prétend qu'à tout changement de forme correspond un changement de sens. L'avantage de cette manière d'envisager les choses est qu'elle se prête aussi au classement alphabétique des racines.

Mais il est peut être bon de voir très vite ce qu'est l'*iṣṭiqāq*.

De nombreuses définitions en ont été données. Je reproduirai ici celle que rapporte as-Ṣuyūṭī² dans son *Muzhir* (p. 346, T I) et qu'il attribue à Ibn Diḥḥya.

« (L'*iṣṭiqāq* est) le fait de tirer un mot d'un autre, de façon que le mot nouveau, tout en ayant la même base formelle et sémantique que le premier, produise un sens augmenté par rapport à celui de la base, cet augmentation sémantique étant consécutive à la modification intervenue dans le nombre des « lettres » et leur disposition (...) La procédure de reconnaissance de l'*iṣṭiqāq* consiste à soustraire les augments introduits dans les formes ainsi obtenues pour revenir à celle qui leur a servi de base. »

أخذ صيغة من أخرى مع اتفاقهما معنى ومادة أصلية ليبدل بالثانية على معنى الأصل بزيادة مفيدة لأجلها اختلافا حروفا وهئية.

طريقة معرفته تقلب تصاريف الكلمة حتى يرجع منها إلى صيغة هي أصل الصيغ.

Nous pouvons retenir de cette définition que l'*iṣṭiqāq* permet la saisie simultanée de deux termes-objets (les lexèmes), ou plus, et de la relation qui les unit ; il s'agit donc bien là d'un concept métalinguistique qui rend compte d'une situation linguistique diffuse, traduite généralement par l'expression « famille de mots » et dégage un certain nombre de constantes dans le développement du lexique.

Nous dirons tout d'abord que les lexèmes sont articulés par un même axe défini par des caractéristiques ressortissant à l'expression (*mādda*) et au contenu (*ma'nā*).

La relation qui unit sur cet axe chacun des lexèmes considérés aux autres, est double :

Elle est conjonctive, car elle met en évidence ce qu'ils ont de commun (= la même base formelle et sémantique)

Elle est disjonctive, car elle révèle par là-même ce par quoi chacun est différent des autres (= les augments de l'expression et du contenu).

Cette relation de conjonction-disjonction est démontrée par les opérations de réduction effectuées sur toutes les formes lexématiques articulées par un même axe et

² 'Abd ar-Rahmān as-Ṣuyūṭī, ou al-Asyūṭī (de la ville d'Assiout en Egypte où il naquit en 849/1445) fut l'un des polygraphes les plus féconds de la littérature arabe. Certains de ses ouvrages font encore autorité et, en particulier, sa somme linguistique « al-Muzhir fī 'ulūm al-luġa wa anwā'ihā » dans laquelle nous retrouvons toutes les questions qui ont été débattues au cours des siècles antérieurs à l'auteur par les linguistes arabes. Suyūṭī, qui sait aussi, à l'occasion faire preuve d'originalité, rapporte le plus souvent les propos tenus par ses devanciers et extraits d'ouvrages aujourd'hui disparus. Il est mort au Caire en 911/1505.

qui font apparaître, par soustraction, d'une part un élément ultime, intangible : la racine (*asl*) et, d'autre part, les divers augments.

Ainsi, chaque « famille de mots » est-elle définie par une racine commune, sorte d'éponyme lexical auquel sont affiliés par *'intisāb* tous les individus d'un même groupe. Elle apparaît en outre comme se réalisant dans la concaténation de la racine, invariant spécifique et définissable sur les plans de l'expression et du contenu, avec les augments, variants par rapport à la racine et eux-mêmes étant parfaitement définissables par la relation qu'il est possible d'établir entre le plan de leur expression et celui de leur contenu.

Et d'autant plus légitimement que ces augments se retrouvent dans les autres « familles lexicales » affiliées à des racines différentes.

Ainsi, l'ensemble du champ lexématique arabe peut-il être articulé selon deux axes orthogonaux, que l'on pourrait qualifier l'un de générique (*'intisāb*) puisqu'il subsume les actualisations différentes de la même base et l'autre d'homologique (*tanāsub*) puisqu'il subsume les actualisations semblables de bases différentes. C'est dire que les augments pourront, dans une vision inversée, devenir, dans le système de l'arabe, des invariants, dont les différentes racines avec lesquelles ils sont concaténés constitueront les variants.

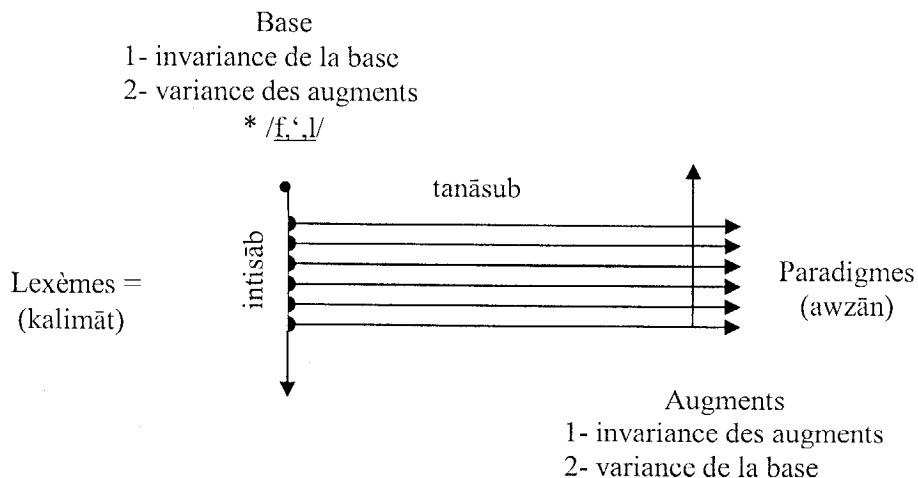
Pour représenter cette dualité de l'analyse morpho-sémantique : à partir de la base vers les augments ou à partir des augments vers la base, les linguistes arabes ont choisi l'une des bases susceptibles de se rencontrer dans le système phonétique de leur langue, à savoir /f,',l/, à laquelle ils ont attribué une valeur de symbole. Ainsi, /f,',l/ a-t-elle deux réalités : la première est d'ordre linguistique puisque cette base est manifestée au niveau lexématique comme l'une quelconque de ses sœurs. La seconde, d'ordre méta-linguistique, a pour fonction de symboliser toutes celles-ci : les phonèmes f (fricative labio-dentale sourde), ' (fricative pharyngale sonore), l (fricative latérale sonore), représentant respectivement

- f toute première radicale
- ' toute deuxième radicale
- l toute troisième radicale

Les linguistes arabes utilisent ce symbolisme pour formaliser leurs analyses : ils parleront

- du fā' (dénominateur du phonème f) de /k,t,b/ pour y désigner le k.
- du 'ayn (dénominateur du phonème') de /m,k,t/ pour y désigner le k.
- du lam (dénominateur du phonème l) de /m,s,k/ pour y désigner le k.

En un deuxième niveau /f,',l/ est affectée de toutes les augmentations admises dans le système de la langue et reconnues par l'analyse, la concaténation « super-base (f,'l) + augment(s) » ayant pour résultat l'établissement d'un paradigme (*wazn*). Ainsi est mis en évidence le caractère modifiant de chacun des augments, tant sur le plan de l'expression que sur le plan du contenu.



Dès lors, l'*Istiqāq* se confirme comme un concept opératoire qui rend compte simultanément de l'ensemble des phénomènes de la diversification lexématique en arabe. Soit une « famille de mots » :

maktab (مكتب), *kātib* (كاتب), *kitābat* (كتابة), *mukātabat* (مكاتبة), *yastaktibu* (يستكتب), *maktūb* (مكتوب),... etc.

Tous les lexèmes de cette famille sont articulés par un même axe défini, sur le plan de l'expression (*mādda*), par une certaine séquence invariante tant dans sa composition (=les phonèmes-graphèmes k,t,b) que dans l'ordre de ses composants qui peuvent être discontinus (1=k, 2=t, 3=b), et sur le plan du contenu (*ma'nā*), souvent par un trait sémantique de valeur générale : « écrire ».

Chacun de ces lexèmes diffère des autres par des traits spécifiques : sur le plan de l'expression, augments consonantiques (préfixes, infixes, suffixes) et/ou vocaliques ; et sur le plan du contenu, augmentations sémémiques corrélatives aux transformations morphologiques et diversifiant l'inventaire sémique de chacun des lexèmes obtenus :

endroit, agent, action, réciprocité,..., de l'écrire.

Inversement, la vision paradigmaticque, à partir du « *wazn* »

<i>maf'al</i> = مَفْعَل	<i>fā'il</i> = فاعِل	<i>fi'ālat</i> = فِعالَة	<i>mufā'alat</i> = مُفاعلة
↓	↓	↓	↓
<i>maktab</i> = مَكْتَب	<i>kātib</i> = كاتِب	<i>kitābat</i> = كِتابَة	<i>mukātabat</i> = مُكاتِبَة
<i>maṣna'</i> = مَصْنَع	<i>ǧālis</i> = جالِس	<i>dirāsat</i> = دِراسة	<i>muḥādāt</i> = مُحادِثَة
<i>ma'mal</i> = مَعْمَل	<i>qātil</i> = قاتِل	<i>qirā'at</i> = قِراءة	<i>mubālaǧat</i> = مُبالِغة
...

qui renvoie chacun de ces lexèmes à une réserve de lexèmes analogues, rend évidente la concaténation d'un augment, invariant chaque fois, avec une base qui, elle, est variable. Ainsi l'augment qui n'a, est-il besoin de le préciser, aucune existence lexicale autonome, non plus que la base d'ailleurs, prend-il l'allure d'un formant phono-sémantique

spécifique et, ce faisant, tend-il à limiter l'arbitraire, ou l'immotivé du signe arabe, en introduisant un principe de régularité dans la diversification lexématique : tout lexème peut donc faire l'objet d'une analyse syntagmatique en un composant nucléaire immotivé (mais nous verrons à l'occasion que certains linguistes ont tenté de réduire cette immotivation) associé à un ou des augments, lesquels étant parfaitement motivés dans le système de la langue, renvoient simultanément, à une analyse paradigmatic dans laquelle ces augments apparaissent, en effet, comme des formants de lexèmes. Cela favorise, moyennant certains aménagements morpho-phonologiques, l'emprunt (l'arabisation) des termes étrangers ou la création directe de mots nouveaux à partir de bases anciennes.

Soit le mot « nickel »³ qui s'est introduit en arabe sous la forme *nikal* sans modification morpho-phonologique, car, d'une part : les phonèmes consonantiques n, k, l sont parfaitement reçus par le système phonétique de l'arabe et la suite qu'ils forment ne présentent aucune incompatibilité, les oppositions vocalique i/a et syllabique longue ouverte/longue fermée, (s'il y a apocope de la voyelle finale = *ni/kal*) ou longue/brève s'il n'y a pas apocope de la voyelle finale : *ni/ka/lu*) sont acceptables en arabe. Et d'autre part, le signifié ne provoque aucune remotivation car il s'intègre sans difficulté à la série, elle-même immotivée des dénominatifs de métaux : *nuḥās* = cuivre, *ḥadīd* = fer, *fiḍḍa* = argent, *dahab* = or ... etc.

Cependant, ce nouveau « mot » arabe a permis de générer le verbe *nakkala* = nickeler et son nom d'action *tankīl*, dont les paradigmes sont respectivement, pour le premier : *fa'āla* et pour le second : *taf'īl*, l'ensemble dénotant la tensivité ou l'extensivité d'une qualité.

Mais il faut bien comprendre qu'il ne s'agit point là d'un simple processus additionnel, mais bien d'une formation par génération, c'est à dire de la réécriture d'un substantif : *nikal*, concaténé avec un affixe (ici, en l'occurrence, il s'agirait exactement, d'un infix), marquant d'un paradigme et obéissant à certaines lois morphosémantiques comme cela s'est produit par exemple en français (nickel + er = nickeler, addition sur le plan graphique, mais la transformation phonique est plus complexe puisqu'elle est rendue nécessaire par un recours à une sorte de *as/* pour assurer le passage d'une syllabe fermée : *Kel*, où une syllabe ferme *Kc* + une syllabe ouverte */ler/*).

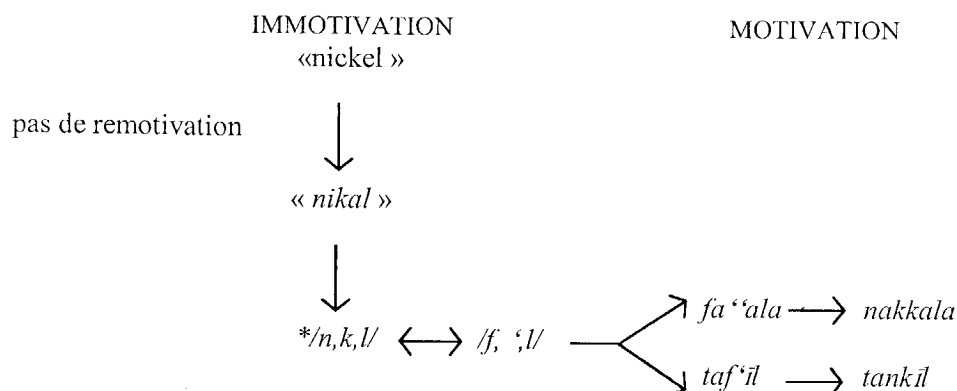
La génération en arabe n'est rendue possible que par le recours au *as/*. Ainsi, à partir de *nikal* (نیکال), l'arabe redescend à n, k, l (ن, ك, ل) en faisant abstraction des phonèmes vocaliques, ceux-ci n'étant pas sentis conformément au système morphologique arabe, comme faisant partie de l'invariant, lequel est généralement réduit à une séquence tri-consonantique « la racine, qui s'avère ainsi douée d'une véritable existence linguistique »⁴

Parallèlement, le contenu du mot « nickel » se vide de toute sa référence par trop concrète et spécifique pour n'être plus qu'une substance virtuellement apte à être in-formée, chaque fois, selon les traits particuliers des paradigmes.

³ Cité par David COHEN, in : Etudes de linguistique sémitique et arabe, p.32.

⁴ Ibid. J'ajouterai que cette existence me paraît être plus métalinguistique, même si elle reste inconsciente, que linguistique, car elle ressortit à la référence à l'étymologie.

Nous montrons dans la figure ci-dessous le cheminement qui, à partir d'un mot non arabe permet de générer toute famille de mots arabes.



La motivation, qui joue un rôle considérable en arabe, ne peut se comprendre que dans une perspective de parcours génératif c'est à dire dans le recours à un *asl* totalement immotivé et relayé dans son accession à la réalité linguistique (actualisation) par des composants morphosémantiques, virtuels (les paradigmes en situation d'attente d'un investissement). L'exemple suivant illustrera encore ce jeu complexe.

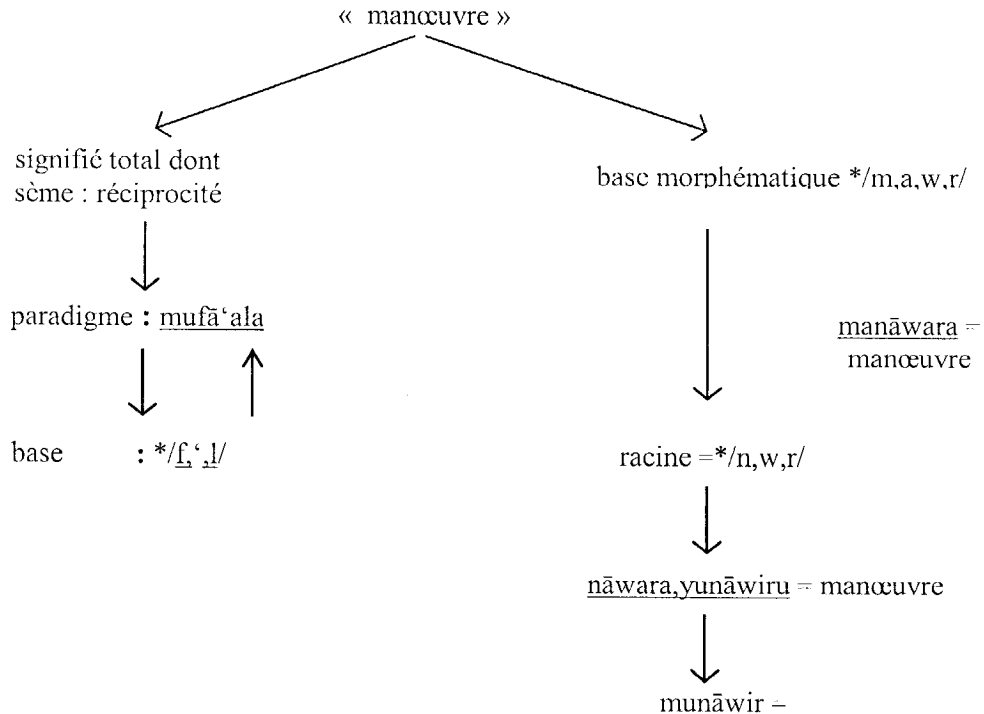
Le mot «manoeuvre»⁵ du vocabulaire militaire est entré en arabe sous la forme *munāwarat*. Ici, c'est le signifié qui a remodelé le terme d'origine en le motivant par référence à un paradigme arabe. En effet, le signifié «manoeuvre» est compris avec une valeur dynamique, impliquant un jeu d'actions réciproques entre deux camps militaires ou politiques, le sème «action réciproque» ayant la propriété d'être symbolisé par le paradigme arabe : *mufā'ala*. Compte tenu de ce signifié reconnu et en conservant l'armature consonantique du signifiant (m,n,v,r) après neutralisation des différenciations vocaliques et moyennant une transformation du phonème «v» en «w»⁶ l'arabe génère le nom d'action *munāwara* :

a- en recourant, comme dans l'exemple précédent, à un invariant triconsonantique à signifié général, indéterminé dans le cas présent : $*/n,w,r/$

b- en motivant le sème «réciprocité» par utilisation d'un paradigme adéquat (*mufā'ala*), ce qui permet de rendre compte aussi de la transformation vocalique de la première syllabe de «/ma/noeuvre» en $/mu/$, cette syllabe $/ma/$ étant généralement reconnue, en arabe, comme préfixe d'un formant discontinu avec valeur de nom de lieu (*ism al-makān*). La présence de $/mu/$ va commencer à donner une réalité morphologique à une virtualité sémantique reconnue dans le sens du mot emprunté.

⁵ Exemple cité par D. Cohen, *ibid.*

⁶ Que l'emprunt ait été fait au français ou à l'anglais et parce que l'arabe ne possède pas de fricative labio-dentale sonore de type \underline{v} .



On voit donc bien que le mot arabe se définit par sa structure. Et celle-ci est permanente. C'est la raison pour laquelle un « mot arabe ne meurt pas » sur le plan morphologique. Et c'est la raison pour laquelle l'arabe peut générer n'importe quelle suite conforme au système à partir de n'importe quelle suite considérée comme une racine, arabe ou non. L'époque classique d'abord mais surtout l'époque contemporaine en donnent de multiples exemples. On a vu celui de l'emprunt au mot : nickel. Je citerai aussi l'arabe *'uksīd* emprunté de (oxyde), et qui a, tout à fait régulièrement engendré : *'aksada yu'aksidu* (oxyder), *'aksadat* (oxydation provoquée), *ta'aksada yata'aksadu* (s'oxyder), *ta'aksud* (oxydation subie), *mu'aksid* (oxydant) *mu'aksad* ou *muta'aksid* (oxydé).

Contrairement à cet état de choses, les lexicographes qui se sont intéressés à la langue française ont longtemps cru que le mot français ne pouvait se définir que dans une diachronie renvoyant à d'autres systèmes antérieurs (le latin) ou à d'autres époques de la langue (le vieux français).

Mais récemment, en France certaines tentatives ont été faites sur la langue française pour « regrouper les mots selon des critères à la fois sémantiques et morphologiques, ce qui doit permettre au lecteur de comprendre le processus de formation du lexique français » le but étant « de présenter une unité sémantique cohérente ».⁸

La réorganisation entraînée par de tels regroupements ne met pas fondamentalement en cause la consultation de ce dictionnaire. Celui-ci, en effet, rend compte, par ces

⁷ Larousse de la langue française, LEXIS, sous la direction de Jean DUBOIS, Librairie Larousse, Paris 1979, p.VI.

⁸ Ibid. p.VIII.

regroupements et en même temps par tout un système de renvois, de certaines régularités que l'on peut dégager dans la dérivation du français.

Mais l'école, déjà, palliait les inconvénients du classement alphabétique grâce aux exercices portant, d'une part, sur la reconstitution des familles de mots et, d'autre part sur les possibilités offertes par l'utilisation de suffixes comme able, ible, ou de préfixes comme in. Ainsi s'agissait il de constituer des ensembles autour d'une unité considérée comme « souche », ou « racine », ou « base » et dont les membres pouvaient avoir subi des traitements différenciateurs très forts, de même qu'inversement, on observait les possibilités de génération offertes par un préfixe ou un suffixe avec les limites que l'on ne pouvait pas rencontrer. L'école remettait de l'ordre, dans ce que les dictionnaires présentaient de manière non ordonnées, et chaque petit français pouvait introduire dans son rapport à sa langue le sentiment d'une certaine régularités morphologiques (comme il y a des régularités syntaxiques). Le sentiment qu'il existe des régularités dans la langue est tout à fait indispensable pour assurer la survie linguistique.

Le lexis donc en systématisant les résultats d'une réflexion métalinguistique qui ne transparaissait pas jusqu'à lui dans les dictionnaires de langue, fait œuvre pédagogique.

Le dictionnaire *as-Sabīl*

Mon entreprise lexicographique qui s'est concrétisée dans la confection du dictionnaire *as-Sabīl* se situe donc au carrefour de la tradition et de la modernité.

La tradition, c'est celle que les linguistes de l'âge d'or de la civilisation arabe ont puissamment contribué à installer par une réflexion extrêmement riche et nuancée, mais ignorée, tout à fait, par les fabricants arabes contemporains de dictionnaires.

La modernité, c'est celle de l'analyse métalinguistique actuelle qui va puiser ses sources dans un structuralisme authentique et dans les découvertes de la sémiotique.

Et c'est parce que cette modernité à la pointe de la recherche scientifique est vécue profondément qu'elle n'a pas honte de renouer avec la tradition.

J'ai donc voulu conserver à l'organisation d'*as-Sabīl* les avantages du classement alphabétique en exploitant au maximum les caractéristiques de l'*ištiqāq*.

Avantages du classement alphabétique en multipliant les renvois aux racines, à partir de l'ordre alphabétique, pour certaines formes particulièrement difficiles à repérer : ces renvois sont symbolisés par des flèches (→ ou ←), par des formules : v. aussi (tel numéro) ou v. ordre alphabétique.

Avantages du classement par racines en systématisant l'ordre interne introduit par l'*ištiqāq*. Je me contente simplement d'évoquer ici cet ordre car j'en ai déjà parlé dans l'introduction d'*as-Sabīl*. (p.8 sqq. de la partie français et 12 sqq. pour la partie arabe).

Mais il y a plus. Le classement par racines semble devoir ouvrir une troisième voie située entre la semasiologie dont j'ai beaucoup parlé et l'onomasiologie que je me contenterai simplement d'évoquer en conclusion car elle doit faire l'objet d'une vaste réflexion.

Conclusion

L'onomasiologie, ou lexicographie conceptuelle, notionnelle ou idéologique part du sentiment implicite, ou clairement explicite comme chez les disciples de W. Von Humboldt,

- que la langue est un ensemble organique, un système cohérent où tout se tient et qui organise le monde comme une totalité intellectuelle qui n'est plus reliée au monde réel que par les mots simples interfaces et non point principes organisateurs.

- que la langue n'est pas un système de communication entre les hommes. Mais un système de communication entre les hommes et le monde.

Dans ces conditions, le classement alphabétique des mots :

* détruit cette cohérence par éclatement des racines et dispersion des unités lexicales.

* ne peut qu'être spécifique d'une langue particulière sans avoir de portée universaliste.

Alors que tous les systèmes intellectuels produits par les hommes (qui prennent la forme des doctrines philosophiques, religieuses ou politiques) sont tous :

cohérents

et universels

L'onomasiologie se propose donc d'abandonner le mot comme point de départ pour partir des principes intellectuels ou concepts.

1-La méthode

Contrairement à la sémasiologie qui procédait par expansion prenant la forme de la définition, l'onomasiologie procède par condensation prenant la forme de la dénomination.

Or la dénomination est le résultat d'un certain nombre de procédures parfaitement repérables et dont rend compte l'*istiqāq* et d'une façon générale tout le système dérivatif qu'il s'agisse de dérivation morpho-sémantique ou purement sémantique comme dans ce cas de la dénomination figurative ou de la dénomination translative.

2. Le classement.

Quelle que soit la méthode d'analyse des relations se substituant ainsi à l'analyse des termes jalonnant ces relations, il faut bien qu'à un certain moment, on procède à un classement des concepts qui donnent lieu à ces analyses.

Les concepts doivent ils être rangés :

par ordre alphabétique

logique

scientifique ?

2.1. Classement alphabétique des concepts

C'est le cas des dictionnaires spécialisés et des techniques des dictionnaires analogiques et de synonymes.

2.2. Classement logique

Il relève d'une réflexion théorique apriorique. Et là, on peut se demander quels sont les critères qui peuvent faire juger tel ordre plus logique que tel autre.

2.3. Classement scientifique

Tombe sous le coup de la critique précédente adressée au type de classement logique.

En effet, si l'on choisit les concepts scientifiques que manipule la civilisation contemporaine on ne pourra pas rendre compte des découpages du monde naturel, découpages pré-scientifiques héritées d'époques antérieures et qui coexistent avec les concepts scientifiques.

3. Onomasiologie et typologie des cultures

Cela dit une réflexion de ce type aurait le mérite de pousser les chercheurs à analyser les cultures dans leurs détails les plus aigus, ce qui est le fondement même de toute ambition typologique, c'est à dire généralisatrice à un deuxième stade. Ces chercheurs ne sont pas seulement des lexicographes, mais des anthropologues, des ethnologues, des sociologues, bref des culturologues qui, de proche, par inductions successives, arriveraient à constituer un fichier extrêmement détaillé et pourraient ensuite généraliser, dans un deuxième temps, leurs observations pour parvenir à définir cette typologie des cultures.

On devrait ensuite établir une multitude de rubriques conceptuelles, comportant des contraintes extrêmement précises facilitant la reconnaissance d'abord puis l'exploitation et de telle sorte que l'ensemble des données puisse être installé dans un fichier général. Il est facile d'imaginer, à partir de là, l'établissement de programmes de saisies mécanisées où justement les contraintes évoquées faciliteraient la traduction automatique et permettraient à cette dernière de réaliser des performances plus fines que celles qu'elle a jusqu'ici réalisées.

4. Onomasiologie et traduction automatique

C'est là une application possible de la recherche onomasiologique et de la volonté d'arriver à une typologie des/de certaines cultures. En effet, la démarche onomasiologique semble coller au naturel dans la mesure où elle part du concept pour rechercher les signes linguistiques qui lui correspondent, en d'autres termes, elle part de l'idée à exprimer pour arriver aux mots qui vont permettre d'exprimer cette idée.

On pourra toujours objecter que dans le fonctionnement habituel de la pensée et du discours, l'idée n'est pas séparable de l'expression de cette idée et que la pensée sans les mots ne serait qu'une nébuleuse (Saussure). Si, effectivement, la substance du contenu est indissociable de la forme du contenu, dans le fil du discours et dans le processus de création linguistique, au moins peut-on imaginer que l'analyse, pour les besoins de la cause, se donne la commodité de séparer ce qui ressortit à l'un et à l'autre plan.

Ce que l'on peut dire, c'est que pour exprimer une idée, quelle qu'elle soit, on ne part pas des mots. Un discours ne se constitue pas par concaténation d'unités lexicales dont les unes seraient les mots (substantifs, adjectifs, verbes) porteurs de sens et les autres les mots-outils assurant la liaison entre les mots ou les groupes de mots. On pense d'abord à ce que l'on veut dire, les mots surgissent alors dans la dynamique ainsi créée, prenant leur place comme pour répondre à l'appel de la pensée.

C'est peut être en partant de cette idée, toute simple, que l'on peut imaginer faire progresser la recherche dans le domaine de la traduction automatique qui s'est contentée longtemps de partir avec obstination des mots.

