

LINGUISTIQUE ET PHONETIQUE ARABES*

Abderrahmane HADJ-SALAH

Centre de recherche scientifique et technique
pour le développement de la langue arabe

Résumé

La théorie néo-khalilienne a pour point de départ la découverte, chez les plus anciens grammairiens arabes (VIII^e s.) d'une conception très originale que l'on ne retrouve chez les grammairiens arabes de basse époque que sous une forme tout à fait dénaturée (sauf exceptions).

Ces anciens grammairiens et notamment al-Ḥalīl Ibn Aḥmad (VIII^e siècle) ont bien relevé la fonction essentielle du langage à savoir la communication mais ils se sont gardés d'en faire le principe explicatif exclusif des phénomènes linguistiques et ont donc soigneusement distingué ce qui relève uniquement du communicationnel de ce qui ne concerne que la structure interne de la langue.

La théorie linguistique élaborée par ces vieux chercheurs a été d'abord analysée pendant de longues années puis reformulée dans un cadre logico-mathématique moderne et fait actuellement l'objet d'une exploitation systématique, au niveau de notre Centre, dans différents domaines.

Les grands concepts de la théorie néo-khalilienne

- 1) La notion de corpus ouvert: les données recueillies par le linguiste ne diffèrent pas de celles du biologiste ou du physicien. Dans tous les cas, la validité s'obtient par le caractère vérifiable de ces données
- 2) Distinction entre la structure grammaticale et le code, d'une part, et l'usage qu'on en fait dans des actes d'énonciation, d'autre part.
- 3) La notion de structure dans cette théorie déborde celle du structuralisme post-saussurien: la structure est ici le résultat de la synthèse de la classe et de l'ordre.
- 4) Les unités de la langue ne sont pas nécessairement des segments (ou marginalement des accents). Il existe des dénotants abstraits aussi importants que les dénotants segmentaux ou accentuels. Exemple: le schème et la racine d'un élément nominal ou verbal: chacun d'eux dénote un sens en lui-même: la synthèse des deux dénotants donne un segment dont le sens résulte également de la synthèse des deux sens abstraits (et non de leur amalgame ou de leur juxtaposition).

Cela est le résultat de l'application systématique du qiyās (bijection).

* Publié en anglais dans « Applied Arabic Linguistics and Signal and Information Processing », pp. 3-22. Hemisphere. P. C., New-York, 1987.

L'axe syntagmatique est ainsi abstrait (= ne se confond pas avec la chaîne verbale) parce que:

1° il comporte des cases vides.

2° l'ordre des éléments qui le composent n'est pas nécessairement celui de la chaîne verbale.

Cette analyse se différencie de la mise en correspondance harissienne par ces deux caractères précisément.

5) D'autre part, les transformations qui constituent ici les passages progressifs d'une séquence à d'autres plus complexes selon des règles très précises (ajouts localisés, avec ou sans alternance exclusive, combinaisons selon certains schèmes, changement de position, etc.) génèrent elles-mêmes les items de la langue contrairement à la grammaire générative (1957 et 1965) où la génération relève d'un 1^{er} système qui n'est qu'une axiomatisation simple de l'analyse en C.I.

الخلاصة

إن النظرية الخليلية الحديثة انطلقت من الاكتشاف عند أقدم النحاة العرب (في القرن الثاني الهجري) لتصور علمي أصيل جدا لا يمكن أن نجد ما يماثله عند النحاة المتأخرين إلا على صورة مشوهة (إلا من شذ منهم).

فقد عرف أولئك النحاة وخاصة الخليل بن أحمد (القرن الثاني) الوظيفة الأساسية للغة وهي البيان (communication) إلا أنهم لم يجعلوه المبدأ الوحيد لتفسير الظواهر اللغوية وعلى هذا فإنهم أقاموا التمييز الصارم بين ما هو راجع إلى الخطاب والبيان وما هو راجع إلى البنية اللفظية التي تختص بها اللغة.

إن هذه النظرية اللغوية التي وضعها هؤلاء النحاة القدامى قد نظرنا فيها طيلة أعوام ثم أعدنا صياغتها في إطار منطقي رياضي حديث وهي تقصد الآن لتعتمد وتستثمر على مستوى مركزنا في مختلف الميادين.

المفاهيم الكبرى للنظرية الخليلية الحديثة

(1) مفهوم المدونة المفتوحة: إن المعطيات التي يتحصل عليها اللغوي لا تختلف إطلاقا من المعطيات التي يتحصل عليها العالم البيولوجي أو الفيزيائي. ففي جميع الحالات تصح المعطيات بكونها قابلة للتحقيق (بالرجوع إلى الواقع).

(2) إقامة الفرق بين البنية النحوية وبين وضع اللغة (الكود) من جهة وبين استعمال المستعملين لهما في أفعال خطابية معينة من جهة أخرى.

(3) مفهوم البنية في هذه النظرية يتجاوز مفهوم البنية في المذهب البنوي الذي ظهر بعد سوسير: فإن البنية هنا هي نتيجة لترتيب بين الفئة والترتيب (بين أفراد جنس ونظائرها من أجناس أخرى حسب تعبير القدامى).

(4) إن وحدات اللغة ليست بالضرورة قطعا من الكلام (أو نبرات أحيانا قليلة). بل قد يوجد في الكثير من اللغات عناصر دالة مجردة مساوية في الأهمية للدوال المقطعة (الكلم كالأسماء والأفعال) أو النبرات. مثال ذلك بناء الكلمة وجزئها: كل واحد منهما يدل على معنى على حدة مستقل عن الآخر وتركيب المعنيين يؤدي إلى معنى الاسم أو الفعل (ولا مزج في ذلك ولا انضمام). وهذا التركيب هو نتيجة لإجراء منتظم للقياس (كما يفهمه القدامى).

فالمحور التركيبي يصبح بذلك مجرداً (أي لا يختلط بمدرج الكلام) وذلك لأن:
1- توجد فيه مواضع خالية.

2- ترتيب العناصر التي يتألف منها ليس بالضرورة ترتيب مدرج الكلام.

يختلف هذا التحليل عن تحليل هاريس بهاتين الصفتين.

5) ثم من جهة أخرى فإن التحويلات التي تمثل الانتقال المتدرج من سلسلة من الكلام إلى أخرى تكون أعقد منها وذلك بالخضوع إلى قواعد دقيقة (زوائد لها مواضعها الخاصة بتعاقب أو من غير تعاقب وتراكيب تخضع لمثل معينة وتبديل الموضع وغير ذلك) هي التي تولد بنفسها وحدات اللغة - خلافاً للنحو التوليدي (1957 - 1965) - حيث يتوقف التوليد على نظام سابق وهو عبارة عن أكسيوماتيكية بسيطة للتحليل البلومفيلدي إلى مكونات مباشرة.

Abstract

The starting point of the Neo-Khalilian theory is the discovery, for the most ancient grammarians (8th century), of an original conception that we do not find in the works of the arab grammarians who came after them only in a distorted form (except some cases).

These ancient grammarians, and more particularly al-Ḥalīl Ibn Aḥmad (8th century), have underlined the most important function of language, i.e. communication. However, they did not adopt it as their exclusive principle in the explanation of the linguistic phenomena. Thus, they have clearly separated what belongs to the communicationnal from what is related to the internal structure of language.

The linguistic theory developed by these old researchers has been first analysed for many long years, then reformulated within a logico-mathematical framework and is actually systematically being exploited in several fields at the level of our Center.

The main concepts of the Neo-Khalilian theory:

1. The notion of open corpus: Similarly to those of the physician or biologist, the data collected by the linguist need to be validated thanks to their verifiable characteristic.
2. The distinction between the grammatical structure and the code, from one hand, and their use within the utterance from the other hand.
3. The notion of structure in this theory goes beyond that of the post-saussurian structuralism: structure here is the result of the synthesis of the class and order.
4. Language units are not necessarily segments (or marginally accents). There is abstract denoters which have the same importance as the segmental or accentual denoters. Example: the nominal or verbal element's pattern and root: each denotes a meaning in itself: the synthesis of the two denoters gives a segment whose meaning results also from the synthesis of the two abstract meanings (and not from their mixture or juxta-position). This being the result of the systematic application of the qiyās. Thus, the syn-tagmatic axis is thus abstract (not to be confused with the verbal chain) because:

1° it includes empty positions.

2° the order of the elements that constitute it is not necessarily that of the verbal chain.

It is precisely, these two characteristics that distinguish this analysis from the Harissian one.

5. From another side, the transformations that constitute here the progressive passages from one sequence to other more complex ones according to very precise rules (additions, with or without exclusive alternation, combinations according to some patterns, position's change, etc.) generate themselves the language items as opposed to generative grammar (1957 and 1965) where the generation of items is related to a first system which is only a simple axiomatisation of the I.C analysis.

La linguistique arabe est née au VIII^e siècle de l'ère chrétienne (fin du 1^{er} s. de l'ère hégirienne) c'est à dire avec le début de l'expansion de l'Islām. Cette précocité s'explique par le besoin immense qu'éprouvaient les membres de la nouvelle communauté de connaître la langue du Coran, devenue langue officielle du jeune Etat islamique (décision prise par le Calife ommeiyade 'Abd-al-Malik Ibn Marwān à cette date). L'apparition de la grammaire arabe est intimement liée à celle du droit canon (fiqh), de l'exégèse coranique et surtout de la science des *qira'āt* ou orthoépie coranique. En effet, les premiers grammairiens furent tous des spécialistes de la reproduction orale du Coran et le tout premier à qui on attribue l'idée d'une grammaire est précisément l'inventeur des signes graphiques qui permettaient de distinguer les différentes fonctions grammaticales des éléments linguistiques, à savoir Abū al-l'Aswad al-Du'alī (m. en 69 H.)⁽¹⁾ (l'écriture arabe, comme on le sait, a été d'abord consonantique).

C'est, en fait, à la troisième génération d'orthoépistes que revient le mérite, d'abord, d'une recherche sur le terrain par l'organisation de vastes enquêtes linguistiques afin de recueillir le plus grand nombre possible de données et, ensuite et surtout, d'avoir affiné et systématisé les méthodes d'analyse des générations précédentes. Cette recherche minutieuse et systématique sur la langue, appelée '*ilm al-'Arabiyya* (= la science de la 'Arabiyya) par ces premiers savants et dont le *naḥw* constitue la partie qui traite de la grammaire (au sens large puisque l'étude du système phonétique y est incluse) a été inaugurée par Abū 'Amr Ibn al-'Alā' (m. en 154 H.) pour ce qui concerne la recherche sur le terrain et la codification massive - et critique - des données et par son émule 'Abdallah Ibn Abī Ishāq (m. en 117 H.) pour ce qui est de l'analyse de la langue et de l'induction des grandes constantes.

La théorie de la 'Arabiyya fut l'œuvre de plusieurs générations de grammairiens et se trouve presque entièrement contenue dans l'imposant et remarquable ouvrage d'un grammairien du II^e siècle à savoir le *Kitāb* de Sibawayh (m. en 180 H.). Mais le plus grand de ces linguistes fut, sans conteste, al-Ḥalīl Ibn Aḥmad (m. en 170 H.), le maître de Sibawayh (ce dernier le cite dans son *Kitāb* plus de 600 fois) et l'un des élèves les plus brillants d'Abū 'Amr. On lui doit l'invention de la métrique et cette conception très originale d'un dictionnaire entièrement basé sur les possibilités combinatoires des phonèmes - et la mise au point de concepts mathématiques pouvant s'appliquer à l'analyse de la langue: arrangements, factorielle, groupe cyclique, etc. Une grande partie des explications des cas aberrants, des expressions idiomatiques lui sont dus. Enfin, le système graphique spécifique aux voyelles, la gémation etc., encore en usage de nos jours est l'une de ses inventions (il est aussi l'auteur d'un système de transcription pour les enquêtes). Le mérite de Sibawayh n'a pas été seulement de codifier la plupart des analyses et explications de ses maîtres - et en particulier d'al-Ḥalīl - mais aussi de les avoir admirablement enrichies et même dépassées en maintes occasions. Parmi les linguistes qui vinrent après cette grande période toute faite de création (plus de deux mille grammairiens jusqu'au XIV^e siècle J.C.), il y a lieu de signaler, tout d'abord, la brillante école qui se forma au III^e et IV^e siècle de l'Hégire à Basra puis à Bagdad à partir de l'enseignement d'un grammairien peu connu de nos jours Abū Bakr Ibn

(1) Année hégirienne (69 = 688 J.C.).

al-Sarrāġ: (m. en 317 H.). Ses élèves ont repris la tradition et l'ont considérablement enrichie par des commentaires et des essais personnels extrêmement denses (les ouvrages produits par cette école, encore à l'état de manuscrits, pour la plupart, ne manqueront pas de faire date très probablement, lorsqu'ils seront plus connus). Il s'agit d'abord d'al-Rummāni (m. en 368) et Sirāfi (m. en 384), auteurs l'un et l'autre de deux énormes commentaires du *Kitāb* de Sībawayh (5 et 8 volumes) al-Zaġġāġī (m. en 337) et Abū 'Alī al-Fārisī (m. en 384) et l'illustre élève de ce dernier (m. en 392) dont le génie (et celui de son maître) fut presque égal à celui d'Al-Ḥalīl et de Sībawayh.

C'est à cette époque (fin du IV^e s. H) que la spéculation philosophique, d'abord proprement musulmane, puis franchement hellénisante s'installe d'emblée à côté de la recherche scientifique opérationniste et envahira peu à peu toutes les sciences islamiques et en particulier la grammaire arabe. Cette invasion consistera en une substitution progressive - mais jamais totale ni définitive - de la logique d'Aristote (= al-Manṭiq) aux concepts logico-mathématiques des premiers linguistes arabes. Contrairement à ce qu'on a pu dire jusqu'à présent de l'école ḥalilienne du IV^e siècle, ses adeptes, bien qu'ils ne se privaient pas d'utiliser des concepts relevant du manṭiq, ont toujours pris soin (le plus souvent) de ne pas les confondre avec ceux qui relevaient de la logique mathématique d'al-Ḥalīl et Sībawayh. Mais cela n'a pas été le cas de leurs collègues - moins profonds et beaucoup moins originaux - et la plupart de leurs successeurs.

Après 470 (date de la mort d'un célèbre continuateur d'Ibn Ğinnī à savoir 'Abd al-Qāhir al-Ġurgāni) la pensée arabe dans le domaine qui nous intéresse plus particulièrement, entre dans une ère de stagnation qui durera plusieurs siècles et ce n'est guère qu'au contact de la civilisation occidentale qu'elle aura l'occasion de prendre conscience de cette léthargie et de tenter par conséquent d'en sortir. Dans le domaine de l'étude scientifique du langage, la plupart des chercheurs arabes de notre époque se sont tout simplement rattachés à telle ou telle école linguistique occidentale. Il se pourrait - et nous le souhaitons vivement - que certains d'entre eux parviennent à faire œuvre originale (tout en prolongeant l'optique qui leur a servi de point de départ). Une deuxième manière consiste à approfondir les concepts de la linguistique arabe des premiers siècles de l'Hégire (après en avoir saisi tous les contours et les principes qui les fondent) en eux-mêmes d'une part, et en les confrontant aux concepts de la linguistique contemporaine, d'autre part. Ce qui suppose bien entendu, une connaissance étendue et profonde de ces concepts, *une connaissance objective et critique* (comme celle que l'on se doit d'avoir de la grammaire arabe ancienne, de la grammaire indienne, etc...) c'est à dire sans aucun parti pris. Cette seconde manière, nous l'avons, mes collègues et moi-même, essayée depuis plus de deux décennies à l'intérieur de cette tendance à laquelle on a donné le nom, que nous acceptons volontiers, d'école néo-Ḥalīlienne.

Les méthodes de recherche du 'ilm al-'Arabiyya et leurs fondements scientifiques

Les premiers savants qui se sont donnés pour tâche de codifier la 'Arabiyya ont été amenés à effectuer des enquêtes sur le terrain de la faṣāḥa⁽²⁾ pour recueillir le plus grand nombre de données relevant de cette langue. L'ensemble de ces données a bien constitué

(2) La "Faṣāḥa", pour *Sībawayh*, est la qualité de locuteur natif n'ayant pas séjourné trop longtemps dans des localités linguistiquement cosmopolites.

une sorte de corpus (appelé par eux al-samā' ou al-masmū' : l'univers des données recueillies ex-auditu) mais leur conception du corpus est foncièrement différente de celle du structuralisme empiriste qui est apparu en Occident après Saussure. Le structuralisme empiriste préconise, comme on le sait, de ne fonder la description que sur un ensemble d'énoncés donné une fois pour toutes "afin d'éviter la tentation de susciter des réalisations qui soient conformes aux propres théories de l'enquêteur". Le samā' des Arabes était, quant à lui, toujours ouvert, en ce sens que n'importe quel linguiste pouvait à tout moment prendre acte d'une ou plusieurs réalisations. En fait, le linguiste de la 'Arabiyya ne se réfère pas seulement, au moment où il entreprend son analyse, aux informations qu'il a lui-même recueillies mais à tout *le masmū'*, toutes les données recueillies par d'autres que lui jusqu'à cet instant là. D'autre part, ce *masmū'* est, pour lui, absolument intangible: s'il lui est permis d'accroître son volume en y intégrant le produit de ses observations, ce ne sera qu'au prix d'un *iğmā'* c'est à dire d'une concordance entre ses propres observations et celles de l'ensemble de ses collègues⁽³⁾. Il en découle que rien ne peut être affirmé sur l'existence tangible d'un item, d'un schéma de réalisation, sur son extension d'emploi sans une référence précise à l'ensemble de ce *masmū'* ainsi *objectivé*. Le recours à un univers de données conçu comme la somme de toutes les observations et objectivé par un *consensus effectif* nous semble infiniment plus proche de l'attitude du savant soucieux aussi bien de *rigueur que de fécondité*.

Cet ensemble de données linguistiques qui a servi de base aux descriptions des grammairiens a été obtenu par de nombreuses enquêtes qui se sont étendues à la presque totalité de la péninsule arabe. Les *enquêtes* qui ont fourni la plus grosse partie des données sont celles qui ont été menées par le promoteur de ces recherches et ses disciples à savoir Abū 'Amr Ibn al-'Alā'.

Les informateurs attitrés ou occasionnels devaient être des *locuteurs natifs* et ne pas avoir séjourné trop longtemps dans des centres où l'on parlait d'autres idiomes que la 'Arabiyya (on les appelait: fuṣṣḥā' al-'Arab).

Les enquêteurs de la 'Arabiyya ont mis en œuvre des techniques très élaborées dont l'essentiel était constitué par une observation exclusivement réceptive (*istilgā'*) et une observation active; cette dernière pouvait se faire aussi selon deux processus: la stimulation - suggestion appelée *talqīn* et l'interrogation sur des faits de langue. Dans le premier cas, on envoie au témoin une série de stimuli pour provoquer le dégel verbal et lui suggérer aussi un thème de discours (il s'agit ici d'obtenir des variantes de réalisation à tous les niveaux). Dans le second cas, il s'agit d'obtenir des renseignements sur la langue; on pose donc des questions de fait. Ici encore deux manières apparaissent: l'interrogation se faisait soit par le moyen d'une question - alternative où l'on donnait à choisir au locuteur entre deux possibilités soit par celui d'une question - test ou une phrase stimulus qui constituait un véritable test ou d'une "excitation" sous forme de phrase qu'on lançait au sujet sans autre commentaire.

(3) Ce qui semble justifier ce point de vue, c'est qu'il n'y a aucune différence, après tout, entre les données du physicien (ou celle du biologiste) et celles du linguiste: dans les deux cas n'importe quel chercheur *doit pouvoir vérifier* à tout moment une observation obtenue par un autre chercheur.

Tout cela était naturellement consigné par écrit: le grand linguiste que fut al- Ḥalīl Ibn Aḥmad avait d'ailleurs mis au point, à cet effet, un système de transcription extrêmement pratique pour noter les variantes.

La systématisation rationnelle des faits, dans les sciences islamiques, a, très tôt, pris la forme -chez les linguistes tout particulièrement - d'une symbiose assez remarquable entre deux tendances opposées - et apparemment contradictoires: un attachement aux faits extrêmement pointilleux et un engouement très prononcé pour les constructions abstraites. La symbiose, chez un Sibawayh par exemple, a consisté à accorder d'abord à l'expérience un primat absolu (elle doit avoir toujours le dernier mot) mais une importance capitale est aussi conférée aux constructions hypothético - déductives et à la formalisation en général.

Les grands concepts sur lesquels est basée cette démarche sont les notions de *bāb*, *naḡīr*, de *aṣl* et *far'*, d'*istimrār* ou *ittirād* et surtout celle de *qiyās*. Les Arabes appellent *bāb* tout ensemble d'objets ou de processus ayant, non pas *une simple propriété* en commun, mais *une structure commune*. D'où la possibilité qu'un *bāb* soit *vide* quand il n'existe aucun item linguistique dans l'usage qui puisse relever de la structure qui le caractérise, ou monaire quand il n'y en a qu'un seul. C'est le cas des objets que caractérise la structure lexicale:

C_1 i C_2 u C_3 (où C_i = consonne radicale en position i ; chez les Arabes: fi'ul), structure exigée par la combinatoire lexicale de la 'Arabiyya mais inexistante dans l'usage. D'où également l'impossibilité d'appliquer ce terme aux classes de réalisations concrètes allophones ou allomorphes. Le *bāb* n'est donc pas une classe simple au sens logique du terme mais un ensemble dont la constitution peut découler d'un ensemble d'hypothèses et/ou de règles combinatoires. Un *bāb* comporte donc des éléments homologues entre eux: ce sont des *naḡā'ir*, plur. de *naḡīr*. Un *bāb* plus abstrait est celui qui intègre d'autres *bāb*-s ayant une même structure mais se situant à un niveau supérieur d'abstraction. C'est la raison pour laquelle on donne le nom de *bāb* aux schèmes générateurs d'items (lexicaux ou syntaxiques) à différents niveaux d'abstraction, ainsi qu'aux séquences ordonnées de consonnes qui constituent des racines lexicales.

Les notions de *aṣl* et de *far'* relèvent d'une conception extrêmement intéressante en ce sens qu'elles permettent d'établir un ordre dans l'axe paradigmatique (qui constitue un véritable "fourre-tout" dans le fonctionnalisme contemporain). On qualifie, en effet, de *aṣl* tout élément qui peut se retrouver d'une manière constante dans d'autres formes d'éléments qui sont ses *furū'* (plur. de *far'*) et qui le contiennent et le débordent par un ajout matériel et/ou formel quelconque. Tout ce que nous dirons à propos des modèles de la linguistique arabe constituera une ample illustration de ces importantes notions.

Le *qiyās*, quant à lui, est d'abord une équivalence (au sens mathématique) qui peut s'établir entre deux ou plusieurs structures (donc de deux ou plusieurs *bāb*-s). Cela suppose une analyse des éléments en constituants et une mise en correspondance biunivoque de ces constituants. Mais cela est loin de suffire car il faut encore que cette correspondance soit compatible avec la composition interne des éléments mis ainsi en regard. Le *qiyās* peut s'appliquer à des objets - les items dans leur structure - mais il est encore plus intéressant quand il s'applique à des processus, plus exactement à des transformations (au sens large englobant la transformation harrisienne et chomskienne), l'isochémisme devient alors un véritable isomorphisme et les ensembles d'opérations mis ainsi en correspondance sont alors

nécessairement des structures fermées sur elles - mêmes c'est à dire *des structures de groupe*.

La notion qui se rapproche le plus de celle de *qiyās* est celle d'analogie mais il s'agit d'une analogie beaucoup plus élaborée (comme on vient de le voir) car elle s'applique à des *structures (au sens mathématique et non phonologique)*.

Elle doit par conséquent: 1- permettre une *simulation* de la réalité par la construction de *modèles* (les schèmes générateurs = *muṭul*, plur. de *miṭal*); 2 - faire émerger des structures beaucoup plus abstraites (pas une série de mises en équivalence).

La construction de schèmes générateurs (les *bāb-s* élaborés en *qiyās* et appelés *miṭāl*) ne doit pas nous faire oublier que ces *bāb-s* et ces *qiyās* sont d'abord appréhendés dans la réalité (ce n'est qu'après avoir été abstraits de leur contenu qu'ils deviennent de véritables instruments de simulation). Mais il existe, dans la réalité, un autre type d'uniformité qu'on appréhende non plus à l'intérieur d'un *bāb* mais entre deux entités unies de telle manière que l'apparition de l'une s'accompagne toujours ou presque de l'apparition de l'autre (exemple: sujet - marque -u). Il s'agit bien d'une loi au sens baconien du terme mais en insérant le *qiyās* dans cette relation occurrenceielle ou constante on obtient ce que les Arabes appellent: *qiyās mustamirr*, c'est à dire une uniformité (*iṭṭirād*) *congruente et occurrenceielle en même temps*.

Les théories et les modèles du *naḥw*

1- Les actes de discours et les notions connexes

Les grammairiens sont partis, comme on l'a vu, de l'observation directe du comportement langagier des locuteurs natifs de la 'Arabiyya et ont donc été amenés à concevoir, à peu près comme nos contemporains, les différentes composantes du circuit de communication. Dans tout échange de messages (*ḥiṭāb*), il y a un *mutakallim*, émetteur du message, nécessairement unique, et un *muḥāṭab*, cible du message qui peut être multiple: entre eux circule le *ṣawṭ* ou son qui véhicule le message (ou un substitut tel qu'un écrit ou tout autre chose). Le *mutakallim* ou locuteur détermine le *ḥāl al-ḥadīṭ*: il s'agit de la situation présente eu égard au message, situation qui sert de référence au *muḥāṭab* en ce sens qu'elle constitue le point zéro des coordonnées spatio-temporelles de la communication. Mais la communication n'est assurée que si les participants possèdent en commun la connaissance préalable, donnée et non déduite, du *waḍ' al-luġa* (ou *muwāḍa'a* qui est le code de la langue (plustard: *iṣṭilāḥ al-taḥāṭub*). Ce *waḍ'* est une véritable convention sociale: est considéré comme tel tout son articulé qui est posé par institution et qui peut être opposé à tout *lafḍ* qui n'aura pas été posé comme signifiant ou dénotant (*dalīl*) d'un signifié déterminé (*madlūl*). C'est le cas des séquences de phonèmes: *saṣ*, *dat*, *qāġ* etc. qui n'ont pas été retenues par le "*wāḍi'*" ou instituteur de la langue (pour des raisons d'incompatibilité phonétique). Mais le code ne se réduit pas à des correspondances entre signifiants et signifiés, il y a aussi tous les arrangements de signifiants qui sont aussi conventionnels, mais qui peuvent s'inférer les uns des autres grâce à des règles qui sont les *maqāyīs* (ici plur. de *qiyās*) du *luġawī* ou linguiste. C'est là que réside la différence entre la *luġa* (ou donné de la langue) et le *naḥw*, d'où l'opposition entre le *luġawī* (= qui a pour tâche de recueillir et de vérifier sur le terrain le plus grand nombre de données linguistiques) et le *naḥwī* (= qui s'est spécialisé dans l'analyse et la systématisation de ces données). Le *waḍ'* est opposé à l'*isti'māl* qui est la

mise en œuvre et l'usage effectif de la langue dans des actes de discours. Un principe retenu par les linguistes arabes à ce propos, est que tout ce qu'admet le *qiyās* ou système congruentiel de la langue, ne s'actualise pas nécessairement dans la réalité du discours.

"Le langage a été institué, affirme Ibn al-Sarrāğ (le maître de Abū 'Alī al-Fārisī) en principe, pour la transmission de l'information (*fā'ida*)... mais si l'on s'avisait de dire: "le feu est chaud", "la neige est froide", on aurait là un discours dont la *fā'ida* serait nulle" (Cf. ses *Uṣūl*, f. 43). La *fā'ida* apparaît donc comme une quantité d'information qui peut être positive ou nulle: le message apprend quelque chose à l'allocutaire ou il ne lui apprend rien. Il est dit alors "*mufīd*" ou "*ğayr mufīd*". D'où l'importance accordée, d'une part, à l'ambi-güité ou *labs* et, d'autre part, au savoir ou données que possède l'allocutaire (*'ilm al-muḥāṭab*) et à la fréquence d'emploi (*katra*) dans l'explication des phénomènes d'omission (*ḥadḥ*), de redondance (*ziyāda*) et ceux qui relèvent du contexte. Ainsi le *labs* ou ambiguïté de certains énoncés tels: "kāna insānun ḥalīman" (= un homme était magnanime) (exemple analysé par Sībawayh, *Kitāb*, I, 22) peut rendre le message totalement prévisible et probable (l'information porte dans cet exemple, sur un élément au moins de la classe des humains). On remarquera aussi que seuls les éléments qui peuvent être suppléés par le contexte sont susceptibles d'être omis: ceux qui ne le peuvent pas donc sont les éléments qui véhiculent à eux seuls la donnée qui fait défaut à l'allocutaire. Sībawayh dit aussi: "Les locuteurs implicites et suppriment souvent ce qui est fréquent dans le discours car ils ont grand besoin d'alléger ce qu'ils utilisent d'une manière très fréquente" (*Kitāb*, I, 294). Quant à la redondance, les grammairiens arabes considèrent qu'il en est de deux sortes: une redondance relevant du système et due à un excès de la fonction distinctive: *ziyāda li-l-farq* (désinences casuelles dans les énoncés non ambigus par exemple) et une redondance accidentelle due aux vicissitudes de la communication: il y a surplus de signifiants pour pallier les déficiences de la communication (*ziyāda*). Dans l'un et l'autre cas, il y a *prévention* (*iḥtiyāṭ*. V. Ibn ğinni, *Ḥaṣā'is*, III, 101-111). On dira de cette *ziyāda* qu'elle est *mufīda* c'est à dire fonctionnelle puisqu'elle assure l'efficacité et la sécurité de la communication. Cela dit, les Arabes n'ont pas entièrement basé leur théorie de la langue sur cette notion de *mufīd* comme le font les fonctionnalistes à l'heure actuelle.

2 - Acceptabilité, grammaticalité et interprétation sémantique

Rien ne se trouve dans le *naḥw* qui n'ait été appréhendé dans sa substance première dans les actes de discours (ou *kalām*) des locuteurs arabes: "Le *naḥw* nous dit Ibn al-Sarrāğ, à ce propos, est une science que les anciens ont établie à partir de l'observation répétée et systématique (*istiqrā'*) du *kalām* des Arabes" (*Uṣūl*, f. 1R). Bien qu'étant essentiellement basée sur l'observation des faits, le *naḥw* ne se réduit pourtant pas à une simple description de la langue, car comme le précise Abū 'Alī, l'élève d'Ibn al-Sarrāğ, il s'agit de "la science des *maqāyīs* (plur. de *qiyās*) abstraits (ou induits = *mustanbaṭa*) par *istiqrā'*"⁽⁴⁾ des actes de discours des Arabes" (*al-Takmila*, 1). Or le *qiyās*, comme on l'a vu, est en même temps une constante (une loi tirée de l'observation) mais c'est aussi un schème générateur ou un modèle qui permet d'engendrer un nombre infini d'énoncés grammaticaux (comme

⁽⁴⁾ observation systématique

l'imagine N. Chomsky mais avec un type de récursivité totalement différent⁽⁵⁾. Tirés ou confirmés par l'expérience, les *maqāyīs* permettent d'abord de prédire d'autres actes de discours et de rendre compte, par leur mise en correspondance structurelle d'un grand nombre de phénomènes relevés dans la réalité du discours.

C'est pour cette raison que l'idée de grammaticalité - dans le sens exclusif de "conforme à ce qu'exige le *qiyās*" - n'est pas étrangère aux grammairiens arabes.

"Le *naḥw*, affirme l'un des grammairiens du IV^es. H, est une discipline scientifique qui permet de connaître les différentes manières d'être (*aḥwāl* = descriptions) du langage des Arabes du point de vue de la validité de ses arrangements et de distinguer, par ce moyen, les énoncés corrects des énoncés incorrects" (cité dans *Iqtirāḥ*, 6) (par référence aux comportements langagiers des Arabes d'après al-Rummāni) (Cf. ses *Hudūd*, 38). Autrement dit, de distinguer les énoncés qui appartiennent à la 'Arabiyya (et non à une norme relevant d'un groupe social restreint et privilégié) de ceux qui n'appartiennent pas à cette langue, et c'est dans ce sens seulement qu'on parlait d'acceptabilité (*ḥusn, istiḥsān*): il s'agissait d'acceptabilité par rapport au plus grand nombre de locuteurs natifs (d'où des degrés d'acceptabilité que traduisaient les expressions: *faṣīḥ / afṣaḥ, ḥasan / aḥsan, qabiḥ / aqbaḥ*: ce sont là des réactions de locuteurs natifs que les grammairiens ont enregistrées au cours de leurs enquêtes et qu'ils ont totalisées au moyen d'inventaires statistiques.

Sibawayh a été l'un des premiers, parmi ces linguistes, à relever les rapports qui existent entre la grammaticalité, l'acceptabilité des énoncés et leur interprétation sémantique. "Le *kalām*, dit-il, peut être: *mustaqīm ḥasan* (grammatical et acceptable), *muḥāl* (asémantique) (al-Sirāfi son commentateur précise: *mustaqīm muḥāl* = grammatical et asémantique), *mustaqīm kāḍib* (grammatical mais invraisemblable), *mustaqīm qabiḥ* (grammatical mais non acceptable, sauf en poésie pour certaines formes), *muḥāl kāḍib* (asémantique et invraisemblable) (Cf. son *Kitāb*, I, 8).

3 - Sémiologico-grammatical Vs sémantique

L'une des plus importantes acquisitions de la linguistique arabe (d'al-Ḥalīl et de ses disciples) a été la distinction très nette qu'elle établit entre l'analyse proprement sémiologique et grammaticale de la langue et celle des effets de sens qui relèvent de l'acte d'énonciation. Il ne s'agit pas ici de donner plus de lumière à l'une aux dépens de l'autre ni même de les séparer totalement mais seulement d'éviter la confusion fâcheuse (dans laquelle sont tombés maintes approches anciennes et modernes) entre ce qui relève du *waḍ'* (système sémiologico-grammatical = signifiants/signifiés + arrangements) et ce qui appartient à l'*isti'māl* (usage effectif du *waḍ'*) avec tout ce que cela implique comme cadre communicationnel, logico-sémantique, etc. Cette distinction se fonde d'ailleurs sur une constatation (faite d'abord par Sibawayh dans son *Kitāb*, I, p. 15-16) relative au *lafẓ* (sons articulés) et au *ma'nā* (sens) lorsqu'ils sont unis par le *waḍ'*. Le premier est ainsi déphonétisé (dans cette union dans le *waḍ'*) à des degrés divers: voyelles et affixes au niveau des schèmes de lexèmes sont des variables. De même, le *ma'nā* est désémantisé à des degrés divers: les unités lexicales nominales et verbales ont, au niveau du code, des contenus génériques et, au niveau

(5) et non basé sur la formalisation du type d'analyse en C.I

des structures (racines, schèmes lexicaux et syntaxiques), le contenu est encore plus abstrait⁽⁶⁾. Enfin les phénomènes de synonymie et d'homonymie - qui sont essentiels et non accidentels - font du *wadʿ*-code et du *wadʿ*-structure (sémiologie et grammaire) **une entité totalement disponible** et capable de répondre à toutes sortes de besoins. D'autre part, le sens que peut inférer l'allocutaire de la situation ou de données antérieurement acquises ou fournies par l'intuition sensible ou intellectuelle sont des *maʿāni* (plur. de *maʿnā*) qui ne relèvent pas du *wadʿ* de la langue, mais de l'infini domaine du sémantique (V. à ce sujet les remarques de Sibawayh à propos de l'indication (*dalāla*) que peut fournir le *lafḍ*-signifiant qu'il oppose à celle que l'on infère de tout autre chose que le *lafḍ*. *Kitāb*, I, 15-16 et aussi Ibn Ğinnī, *Haṣāʾiṣ*, II, 184, 321 et III, 98 sqq.) De même, le langage ne se définit pas par sa fonction de communication (ni par aucune autre fonction). C'est ce que nous dit expressément un philosophe du langage qui a bien saisi l'esprit de l'école ḥalilienne: "Par notre propos: le discours est constitué par ce qui est *mufīd* (qui apporte une information), nous ne voulons pas dire qu'il doit comporter nécessairement une "fāʾida" (information) chaque fois qu'il se réalise mais seulement qu'il est ce par quoi la *fāʾida* peut valablement se réaliser bien que cette propriété puisse parfois lui faire défaut en raison d'une certaine manière d'être du locuteur" (*Muġnī*, VII, 11).

4- Le modèle syntaxique

- Différences essentielles avec l'approche structuraliste ou générativiste:

1- la synthèse de la classe et de l'ordre

Forme et substance sont des notions que les grammairiens arabes ont bien connues (en arabe: "ṣūra" et "mādda" qui sont la traduction de termes employés par Aristote) mais ils ne se contentaient pas de voir dans la forme un simple système d'oppositions; la forme pour eux est, en plus, un arrangement de ces oppositions. Autrement dit, une structure où les éléments d'une classe sont mis en relation directe avec ceux d'une ou plusieurs autres classes. L'emploi des termes aristotéliens pour ces grammairiens ne signifiait pas autre chose qu'à toute structure s'opposent à un niveau inférieur d'abstraction (pas nécessairement par emboîtement) l'ensemble des éléments qu'elle "informe" et, au niveau supérieur, une structure plus large qui l'intègre comme élément. Ainsi la simple opposition du type aristotélico-saussurien n'est pas à elle seule à l'œuvre dans le système immanent de la langue, même si on l'agrèment d'une dimension accessoire appelée "contraste". Il faut encore que les éléments du système soient mis en correspondance avec d'autres éléments appartenant à d'autres classes et c'est alors seulement qu'ils acquièrent par leur position (= leur *mawḍiʿ*) dans ce système extensif le statut de variable absolue (non liée à une classe).

(6) Rappelons que pour nous la notion de structure ne se réduit pas à des relations intensives d'opposition et de co-appartenance (dont l'importance a été extrêmement exagérée chez les linguistes structuralistes post-saussuriens). Ces relations ou les classes qui en découlent doivent encore être *composées entre elles* pour qu'on puisse parler de structure.

2 - Les quatre types de dénotants

Les linguistes arabes ont identifié dans la 'Arabiyya quatre sortes d'unités ou, plus exactement, quatre types de dénotants: par la racine et le schème, le segment signifiant et la marque zéro.

Le *segment signifiant* est ce que les Arabes appellent *kalima*. On serait tenté d'y voir le résultat d'une simple segmentation (avec commutation) qui constitue une analyse simple du texte et qui aboutit dans la linguistique occidentale au "morphème". Nous verrons ci-après que la *kalima* ne résulte pas d'une analyse aussi simpliste. Les grammairiens arabes opposent à cette unité deux autres dénotants: la *racine* et le *schème* qu'ils obtiennent par une analyse verticale, autrement dit, par une induction de *variables ordonnées* à partir de deux opérations qui s'effectuent simultanément: 1° une mise en équivalence de segments signifiants et, 2° la sériation (ou mise en ordre) de cette classe d'équivalence. Voici comment on peut représenter ces opérations pour l'obtention d'un schème:

sériation	f	C ₁	k	ḍ	‘	ḥ	...	} Schème (constantes + variables ordonnées)
	a	a	a	a	a	a		
	‘	C ₂	t	r	m	s	...	
	a	a	a	a	i	u		
	l	C ₃	b	b	l	n	...	

(7)

équivalence (classe de verbes trilitères à l'accompli).

On agira de même pour obtenir les racines: il suffit de changer de classe d'équivalence.

Il existe, cependant, des segments signifiants qui ne sont pas analysables en racine et en schème. Il s'agit, d'abord, de tous les morphèmes grammaticaux - appelés *hurūf al-ma'ānī* ou *adawāt* (mots outils) - auxquels s'ajoutent tous les segments qui fonctionnent comme marques de seconde puissance (substituables aux noms pleins) à savoir les déictiques ou embrayeurs (*al-asmā' al-mubhama*).

Ainsi, **tout n'est pas segment ou accent dans la langue**: chacune des entités issues de l'analyse verticale de la *kalima* possède à elle seule une *dalāla* ou dénotation. Dans "kataba", "kātib", "maktūb" la séquence virtuelle | K T B | dénote, par elle même, dans le code de la 'Arabiyya, le signifié "écrire"⁽⁸⁾ et les schèmes: *fī'āla*, *fā'ala*, *fā'il*, *maf'ūl*, délimitent les signifiés: "action", "accompli" (procès), "agent", "patient". Ainsi donc la *kalima*, lorsqu'elle est analysable en une racine et un schème, n'est pas un signifiant minimal (elle ne l'est que sur le plan du texte: aucun de ses éléments ne dénote à lui seul son signifié).

(7) à lire verticalement: katab, ḍarab, 'amil, ḥasun (il a écrit, a frappé, a agi, est beau). F, ' , l sont des symboles représentant, comme on va le voir, les consonnes radicales, dans le schème séquentiel du segment signifiant.

(8) Cette notion dénotée par la racine est abstraite d'une manière absolue et donc dépouillée de tout incrément sémilogique structurant (même de l'action d'écrire et son résultat). Il s'agit d'un schème absolument primitif (au degré 0).

Comme on le voit, les dénotants ne sont pas nécessairement des segments. Vouloir à tout prix analyser un texte en se contentant seulement de le segmenter à la manière des linguistes post-saussuriens, aboutit à des impasses ou à des solutions aberrantes. Tel ce concept de morphème "discontinu" par lequel on veut absolument rendre compte des variations de schème. L'exemple le plus frappant à cet égard est le pluriel interne de la 'Arabiyya: comment pourrait-on expliquer dans cette optique segmentaliste à outrance, le passage de "ṣāḥib" (compagnon) à "'aṣḥāb": où donc se trouve le dénotant segmental qui marque le pluriel? c'est la séquence des variantes consonantiques (f, ' , l) et des constantes vocaliques et consonantiques qui dénote le pluriel et non le prétendu morphème discontinu: [' a-voyelle Ø- (a):] dont les éléments 2 et 3 sont imprononçables !

Un dernier dénotant relevé par les Arabes - et par nos contemporains - est la marque zéro que Sibawayh appelait "'alāma ḡayr ḡāhira" (marque non exprimée) ou "tark al-'alāma" (*Kitāb*, I, 123, 220). Il s'agit de l'absence, significative, d'une expression sur le plan phonique. Elle caractérise, par exemple, le pronom affixe de la 3^e pers. du verbe, les marques nominales du masculin et du singulier opposées au féminin et au duel/pluriel qui en dérivent par l'adjonction d'une marque pleine. On l'applique également au niveau syntaxique: le régissant zéro, par ex, qui correspond à l'*'ibtidā'* est la marque de l'unité syntaxique qui sous-tend la phrase nominale simple.

Quelques auteurs, bien postérieurs à Sibawayh, influencés par la philosophie, ne comprenaient pas que l'absence d'une marque puisse avoir un effet de rection. Incapables de penser les phénomènes dans une optique opératoire, ces grammairiens ne pouvaient faire la différence entre une absence considérée en elle-même et une absence considérée dans un ensemble structuré: c'est alors qu'apparaît le rôle de cette absence par le fait même qu'elle affecte une position dans cet ensemble (Cf. le rôle du concept zéro en math.)

3 - La notion de *mawḡi'*

C'est précisément cette notion de position dans un ensemble structuré (non pas un système d'oppositions simples mais un ensemble dans lequel sont mis en correspondance biunivoque tous les éléments qui y sont contenus) qui est à la base, avec quelques autres concepts, de l'analyse du *naḥw*. Un autre terme, celui de *mawḡi'*, est utilisé quelquefois en lieu et place du mot *mawḡi'* il s'agit du "lieu d'occurrence" d'un élément dans le discours. Al-Rummānī parle, à ce propos, de la "*qismat al-mawḡi'*" c'est à dire de la combinatoire distributionnelle des éléments du discours. L'analyse du *naḥw* serait-elle "distributionnaliste"? Nous allons voir qu'il s'agit d'une "distribution" beaucoup plus élaborée. En fait, le *mawḡi'* ne se confond pas toujours avec le lieu matériel que peut occuper un élément dans le discours (ainsi l'antéposition du complément ne change pas son *mawḡi'*). Le *mawḡi'* ne se définit ni par la seule distribution ni par la fonction des éléments qui l'occupent: il s'agit d'une position contenue virtuellement dans un schème opératoire, abstrait à partir des deux axes syntagmatique et paradigmatic en même temps. Cette notion apparaîtra sous un jour plus lumineux dans les analyses des modèles du *naḥw* (qui l'impliquent).

5 - La lexie minimale comme point de départ de l'analyse ou le niveau intermédiaire entre le segment signifiant et l'unité syntaxique (un type d'unité extensible dans le système)

"Le *ism muḍhar* (= nom explicite, opp., au *muḍmar* = pronom pers.), affirme Sībawayh, ne se réduit jamais à un seul *ḥarf* (segment phonique) car il doit pouvoir être *suivi d'un silence* et ne pas être en même temps, *précédé* ou suivi d'un autre élément..." (*Kitāb*, II, 304). Toute séquence verbale pouvant être coupée de ce qui la précède (*ibtidā'* = attaque) et de ce qui la suit (*infiṣāl*) et considérée comme minimale du point de vue de cette isolabilité (*infirād*) constitue une unité formelle, puisqu'abstraite à partir du seul signifiant, que ce même auteur appelle "*kalima mufrada*" et "*bi-manzilati kalimatīn wāhida*"⁽⁹⁾. La première de ces dénominations s'applique aux séquences isolables (et insécables) qui ne contiennent aucune sous-séquence qui, dans d'autres occurrences, sont isolables. Il en est ainsi du *ism*, explicite et implicite (nom et pronom personnel isolable). La seconde s'applique aux séquences qui en dérivent par un ou plusieurs ajouts qui lui sont alors liés de telle façon que la séquence reste encore insécable.

L'*infirād* ou isolabilité ainsi comprise est d'une importance capitale car en délimitant dans le discours des tranches potentiellement autonomes, il nous permet d'accéder à l'unité véritablement fondamentale qui se situe à l'intersection de la syntaxe et du lexique - et même à l'intersection du syntaxique et du communicationnel puisque toute tranche ainsi définie peut fonctionner aussi comme message minimal (comme une phrase⁽¹⁰⁾, quelle que soit la quantité d'éléments qu'elle peut contenir, l'essentiel étant la présence de deux coupures pausales comme limites et le caractère insécable). A partir de ce niveau - central - l'analyse du linguiste s'orientera donc soit vers le bas pour rechercher les segments signifiants et leurs composantes soit vers le haut pour voir comment les minima séquentiels s'intègrent dans les structures syntaxiques.

Cette façon de procéder permet d'éviter l'arbitraire de beaucoup d'approches actuelles ou anciennes qui prennent comme point de départ la phrase ou la proposition (laquelle phrase n'est pas une donnée purement formelle). Certains auteurs contemporains se sont préoccupés de définir un point de départ formel: c'est le cas de Harris (V. Lyons, *Introduction*, 5, 1, 2) et de Revzin (V. ses *Modèles*, trad. franç., 15) mais aucun de ces auteurs, nous semble-t-il, n'a essayé d'exploiter à fond les notions qu'ils ont établies.

C'est précisément cette possibilité de génération (*tafri'*) de séquences dérivées (*furū'*) à partir d'une séquence primitive (*aṣl*) réellement et formellement par l'adjonction à la droite ou à la gauche de cette dernière, d'incrément significatifs (*zawā'id*) que les Arabes appellent *tamakkun* (capacité)⁽¹¹⁾ et *taṣarruf* (variabilité). Cela permet d'établir des distinctions (qui ont l'avantage d'être uniquement formelles) entre séquences isolables: certaines séquences ont, en effet, un *tamakkun* parfait: elles peuvent recevoir n'importe quel type d'incrément. C'est le cas du *ism ḡins* (ou nom commun), d'autres le sont moins ou

(9) Des auteurs postérieurs à Sībawayh ont donné à cette entité le nom très significatif de "*lafḍa*" (= *unité de lafḍ*) (V. Ibn Ya'īs, *Šarḥ al-Mufaṣṣal*, I, 19 et Raḍī, *Šarḥ al-Kāfiya*, I, 5).

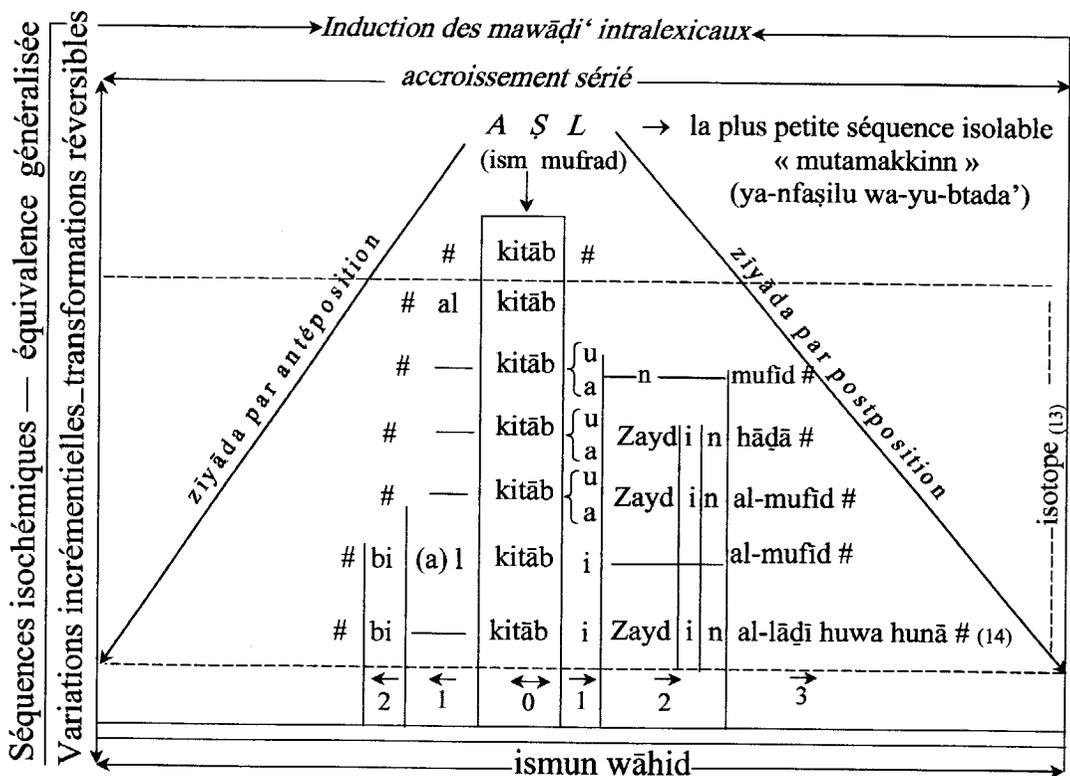
(10) Nous avons suggéré ailleurs de l'appeler 'lexie'. V. pour plus de détails (sur l'ensemble de cette étude) notre ouvrage: "Linguistique arabe et linguistique générale" (V. Références).

(11) de supporter des incréments: chaque catégorie grammaticale a sa propre capacité qui peut être égale à zéro.

beaucoup moins: le *'alam* ou nom propre, par exemple, qui ne peut recevoir l'article, ni un complément adnominal; certains adverbes, en plus, ne peuvent recevoir qu'une ou deux marques relevant de l'*i'rāb* (déclinaison) ou pas de marque du tout, etc.

Le *tašarrufou* variation (incrémentielle) à partir de la séquence initiale est soumis à des règles dont l'application (ordonnée) peut se représenter sous cette forme:

LE ḤADD DU ISM ⁽¹²⁾



(12) *Ḥadd* comme modèle de production et de reconnaissance ou dans l'optique formelle du *naḥw*: structure capable de *caractériser* (au sens mathématique) les items de la langue.

(13) *Furū'* dérivés du *aṣl* et ayant les mêmes *isotopies* ou *mawāḍi'* dans le discours.

(14) Traduction: (un) livre / le livre / un livre intéressant / le livre de Zayd que voici / le livre intéressant de Zayd / avec le livre intéressant / avec le livre de Zayd qui est ici.

Comme on le voit, les incréments sont des ajouts qui peuvent apparaître en une position et alterner aussi entre eux. Les désinences casuelles, le *tanwīn* (qui alterne horizontalement avec l'article et verticalement avec le complément adnominal), le qualificatif à droite du noyau d'une part, et l'article puis la préposition à gauche sont tous des incréments par rapport au noyau qu'ils peuvent affecter et dans la séquence duquel ils peuvent apparaître et disparaître. On remarquera, d'autre part, la présence de zones de récursivité (*iṭāla* par 'atf, *takrīr* ou *taṭniya*) à ce niveau, ce sont les *mawḍi'*-s du complément adnominal $\vec{\text{ز}}$ et celui du qualificatif $\vec{\text{ز}}$. Nous verrons que c'est en ces positions (entre autres) que peuvent se réaliser les enchâssements et emboîtements de séquences.

Ainsi les critères de l'*infirād* (*ibtidā'*, *infisāl*) et du *tamakkun* permettent de reconnaître dans le *lafḍ* (la forme signifiante) et rien que dans le *lafḍ* une première unité, qui est ici le *ism 'āmm* (ou *ḡins* = nom commun); l'élément qui aura le plus grand *tamakkun* sera celui qui est absolument non marqué dans son actualisation minimale (son *aṣl*). Nous avons donc là une définition formelle de ce type d'unité: toute séquence qui admet à sa gauche et à sa droite tous les ajouts possibles sans que cela fasse perdre à l'ensemble le caractère de séquence insécable du point de vue de sa réalisation. A partir de là, on pourra définir formellement les autres unités qui sont contenues dans les incréments (réglementaires). Cela est possible parce qu'elles apparaissent dans des positions spécifiques que l'on infère à partir de l'examen des apparitions possibles d'un même élément. Ces *mawḍi'*-s déterminent à leur tour, toujours formellement, les fonctions grammaticales des éléments qui les occupent.

C'est aussi dans cette optique que la *kalima* peut recevoir une définition formelle. En effet, toute tranche phonique pouvant apparaître dans l'un de ces *mawḍi'*-s est considérée comme *kalima*. Il s'ensuit que la *kalima* ne se confond pas nécessairement avec la notion de "morphème": en effet, la *kalima* est bien un segment signifiant qui s'obtient comme on le voit par l'élaboration du schème générateur de lexie ou d'une unité syntaxique mais ce segment n'est minimal qu'en égard au schème de lexie qui le contient. Il s'agit d'un élément de lexie alors que le morphème est une unité signifiante (presque toujours assimilée à un segment) quel que soit le niveau où l'on se situe. D'où la confusion que commettent les descriptivistes entre les éléments signifiants qui entrent en composition dans le système générateur des lexèmes (les *miṭāl-s*: *fā'āla*, *iftā'āla*, *istif'āl*, etc.) qui sont des éléments internes au lexème et les éléments signifiants externes à ce schème, ces derniers se différenciant formellement des premiers par le fait qu'ils sont détachables par suppression dans le même schème.

D'autre part, les grammairiens arabes considèrent que l'élément le plus capable de supporter des incréments est aussi le moins dispendieux à actualiser (*'ahaff*). Tel est le cas du nom commun. Le verbe est le moins économique (*atqal*) parce qu'il ne peut s'actualiser sans le sujet, les désinences, les actualisateurs temporels, etc. Le *tiqal* désigne précisément la charge incrémentielle et se situe, par conséquent, aussi bien sur le plan physiologique que sur le plan psychique.

Comme on le constate, l'axe paradigmatique est nécessairement structuré, en ce sens qu'il est le siège de transformations et que ces dernières sont nécessairement hiérarchisées en *aṣl - furū'*. Les transformations incrémentielles ou *ziyāda* déterminent bien des dis-

inctions, des rapports paradigmatiques mais ces derniers ne doivent pas être considérés quand on se réfère à l'axe syntagmatique à l'intérieur d'une seule classe morpho-syntaxique. Il faut, au contraire, les considérer dans cet ensemble structuré qui résulte de la combinaison des deux axes. Il est donc nécessaire de considérer ensemble et sans les isoler toutes les colonnes où se situent les rapports paradigmatiques: 1° dans l'ordre syntagmatique qui les caractérise, 2° dans le mouvement qui dynamise les deux axes par le passage gradué et croissant de la séquence minimale à ses dérivés et inversement, autrement dit, dans une perspective de totalité ou de structure fermée sur elle-même et ayant ses propres propriétés (nous verrons que tous les schèmes de la 'Arabiyya constituent de telles totalités qui s'avèrent être des structures de groupe). Il y a ici, de nouveau, synthèse de la classe et de l'ordre. Ce schème générateur de lexies ainsi que tous les autres schèmes, à quelque niveau qu'ils se situent, constituent, d'un point de vue un *bāb* puisqu'il s'agit d'un ensemble structuré par une mise en équivalence ordonnée (= généralisée) du *asl* et de ses *furū* (toutes les séquences dérivées constituent des lexies au même titre que le point de départ: elles sont ses isotopes) mais aussi un *qiyas* et un *ḥadd* (définition opératoire) puisque cette mise en correspondance biunivoque réversible constitue un modèle de dérivation et donc de caractérisation et de génération - réglée - d'items linguistiques.

A côté du schème générateur de la lexie nominale, il existe un autre schème qui génère les lexies verbales (en fait trois sous-schémes différents qui correspondent, en arabe, aux trois modes verbaux: l'accompli, l'inaccompli et l'impératif).

- Ainsi la lexie constitue, en langue (dans le potentiel sémiologico-grammatical), une véritable unité extensible (à l'infini) contrairement à la kalima qui n'admet qu'un nombre très limité d'incrément.

Avant de passer aux autres niveaux - situés au dessus et au-dessous de ce niveau central - il nous paraît opportun d'évoquer deux autres notions qui ont trait à la cohésion concaténatoire des composants d'une même unité. Il s'agit du *waṣl* (ou *ittiṣāl*) et du *binā*'. Les grammairiens arabes ont constaté que cette cohésion se relâche de plus en plus lorsque l'on passe d'un niveau d'analyse à celui qui lui est supérieur (V. entre autres le *Ṣarḥ* d'al-Rummānī, II. 86 R et 59 V). Ainsi pour ce qui concerne le niveau de la lexie, il ne peut être question que de *waṣl* qui est une concaténation simple (juxtaposition sans aucun effet): les incréments qui apparaissent dans le schème sont simplement concaténés; leur disparition n'affecte en aucun cas la lexie dans laquelle ils apparaissent, qui garde son statut de lexie tant que le noyau subsiste. On retrouvera cette alternance avec zéro, sans dommage pour l'unité, au niveau des périphériques syntaxiques qui sont aussi des incréments. Ce type de concaténation simple ne se réalise pas au niveau lexématique (interne à la *kalima* analysable en racine et en schème) ni au niveau du noyau syntaxique (comme on le verra ci-après): on a dans ces deux cas une cohésion très forte qu'on appelle *binā*'. Nous sommes en présence ici d'une construction, donc de l'intégration ou composition de deux éléments et plus dans une même structure: il en est ainsi des éléments de la *kalima* qui sont intégrés dans le schème qui la génère. Ce qui le prouve, c'est que la suppression de l'une quelconque des composantes détruit l'unité entière: *mukrim*-**krim* (ou provoque tout au moins-mais très rarement - le changement de schème: *tafā*'*ala* - *fā*'*ala*). Il existe aussi des niveaux intermédiaires où le *waṣl* n'est pas aussi lâche mais où le *binā*' n'est pas également aussi

strict. C'est le cas de l'agglutination des *kalim* (on parle dans ce cas de *ḍamm*): les marques du féminin, du duel et du pluriel se lient aux noms comme dans un *binā'* mais elles peuvent disparaître sans dommage pour l'unité.

6- La structure sémiologico-grammaticale de l'unité syntaxique (tectonie)

Pour ce qui concerne le niveau supérieur à celui de la lexie, on constatera tout d'abord que les unités qui s'y situent ne sont pas le résultat d'une simple combinatoire de lexies. La lexie - ni la *kalima* - ne constituent pas l'unité minimale de ce niveau. De plus, les relations qu'entretiennent les éléments relevant de ce niveau sont d'une tout autre nature.

A propos des séquences de lexies attestées dans le discours: [1] # 'Abdullāhi qā'imun# et [2] # qā'imun 'Abdullāhi #, Sibawayh affirme qu'il y a *binā'* entre ces deux lexies et non pas un simple *waṣl*, ce qui est tout à fait exact puisque la suppression de l'une des deux lexies (sans référence au contexte) fait disparaître l'unité. On remarquera, d'autre part, que ces mêmes séquences peuvent se retrouver dans des séquences plus larges: [3] # 'inna 'Abdullāhi qā'imun #, [4] # kāna 'Abdullāhi qā'imān #. Il est clair que [3] et [4] dérivent de [1] par *ziyāda* ou adjonction de "inna" et de "kāna". Etant donné qu'il y a une même relation de *binā'* entre ces lexies et puisque [3] et [4] dérivent de [1], il est possible de les faire correspondre terme à terme:

[1]	#	∅	'Abdullahi	qā'imun	#
[3]	#	inna	'Abdullahi	qā'imun	#
[4]	#	kāna	'Abdullahi	qā'imān	# (15)

On remarquera que les éléments occupant la colonne de gauche (à l'initiale des séquences) semblent avoir un rapport avec les désinences contenues dans les lexies. Ce rapport est justement considéré par les grammairiens arabes comme une réaction (*'ama*). Les éléments régissants déterminent, en effet, la marque désinentielle des éléments régis. Ceci nous permet alors de rapprocher de cet ensemble cette autre séquence de lexies qui comporte une lexie verbale: [5] # ḍaraba | 'Abdullahi | 'Amran # où "ḍaraba" agit également comme un régissant (et où 'Abdullah est considéré comme sujet (*fā'il*) du verbe "ḍaraba"). On remarquera d'autre part que [1] comporte vis à vis des autres séquences l'expression zéro du régissant et c'est cette expression zéro que les grammairiens appellent *ibtidā'*.

D'autre part, ces mêmes grammairiens ont relevé un fait important à savoir qu'il existe un élément parmi ceux que le régissant (ou *'āmil*) gouverne qui ne peut jamais être antéposé à son régissant: il s'agit de l'élément régi au *naṣb* (marque *a*) par les éléments de la classe de "inna", et du *raf'* (marque *u*) par ceux de la classe de "kāna" et "ḍaraba" autrement dit de tous les items qui figurent dans cet ensemble en seconde position. En ce qui concerne le *binā'* qui comporte l'expression zéro du régissant c'est le *mubtada'* c'est à dire l'item régi par zéro (= al-'ārī 'anī l-'awāmil al-lafdiyya) qui correspond au terme non antéposable. L'item régi obligatoirement en postposition est celui que Sibawayh appelle "awwal mā tašgalu bihi al-'āmil" (*Kitāb*, I, 245 et aussi 41) = le terme qui "absorbe" en premier le régissant. Cette

(15) Traduction: 'Abdullah est debout
'Abdullah est bien debout
'Abdullah était debout

subordination (ordre + mise en dépendance avec ce qui précède) Sibawayh la simule par les séquences que nous venons d'examiner en y visant l'ordre virtuel:

R (régissant synt.) – T₁ (terme régi en 1^{er}),

T₂ (terme régi en 2^e)

où seul T₁ doit se trouver dans la chaîne subséquente à la position de R. On peut donc avoir dans le discours les variantes suivantes: (R, T₁, T₂), # (R, T₂, T₁), (T₂, R, T₁). Nous devons cependant bien comprendre que le *binā'* ou intégration structurelle d'un élément à un autre n'est pas entre R et T₁, mais entre le couple ordonné (R – T₁) et T₂. On relèvera que ce couple ordonné peut se trouver seul - sans T₂ - dans le discours (comme # qāma 'Abdullahi # ou # qumtu # [je me suis levé]). Enfin le contenu de ces entités s'interprète sur le plan "casuel" ainsi: dans R on doit avoir soit zéro, soit un verbe que nous appelons "exponentiel" tel que "kāna" qui est un véritable exposant temporel, soit un exposant non verbal de la classe de "inna" (particule de corroboration): "layta"(souhait), "la'alla" (attente), "ka'anna" (comparaison), etc. Soit aussi un verbe non exponentiel, tel que "ḍaraba". Le contenu de R détermine, en fait, le contenu casuel des termes régis. Ainsi, si R = ∅, T₁ a nécessairement comme contenu un *mubtada'* qui est l'appellation sur le plan formel du contenu de T₁ dans cette situation mais qui peut s'interpréter sur le plan casuel comme sujet d'un *ḥabar*, ce dernier - qui est le contenu de T₂ dans ce type de structure - pouvant s'interpréter comme l'item véhiculant une information à propos du terme posé qui est le contenu de T₁. Si R = exposant (verbal ou non), le noyau de la séquence ne change pas, puisque ces exposants lui sont affectés en tant que tels. On parle seulement de *ism* et *ḥabar* de *kāna* ou *inna* = nom et *ḥabar* régis par ces exposants. Enfin si R = verbe non exponentiel, on obtient alors une séquence qui, bien que homologue à la précédente n'en a pas moins ses propriétés propres. T₁ doit avoir, alors, un sujet (*fā'il*) et T₂ un complément d'objet (*maf'ūl*), T₂ étant alors susceptible de suppression.

On a pu objecter que puisque # qāma 'Abdullāhi # (= 'A. s'est levé) et # 'Abdullāhi qāma # ont le même sens, il serait donc admis d'antéposer T₁ à R (dont les contenus sont ici: sujet et verbe *fā'il*, *fi'l*). On a répondu à cette objection par plusieurs arguments dont l'essentiel se résume ainsi (Mubarrad, *Muqtaḍab*, IV, 128):

"Il existe une position ou *mawḍi'* après le verbe qui ne peut être occupée que par son sujet, or un *mawḍi'* ne peut pas se supprimer: il peut être simplement vide. Cela apparaît clairement dans cet appariement:

| 'Abdullahi qāma | ∅ | # = 'A., | il | s'est levé.

| 'Abdullahi qāma | aḥuhu | # = 'A., | son frère | s'est levé.

Ce même prétendu sujet sur le plan grammatical est susceptible d'être régi par un autre élément, il existe donc un *mawḍi'* précédant ce type de séquence qui est ici vide:

| ∅ | 'Abdullahi qāma | # = 'A., il s'est levé.

| ra'aytu | 'Abdullahi qāma | # = j'ai aperçu 'A. se lever.

'Abdullah est donc *mubtada'* et non *fā'il* dans la 1^{ère} séquence". La formule (R → T₁) T₂ (où la flèche qui relie T₂ au couple ordonné indique la dépendance de *binā'*) constitue, en fait, un véritable schème générateur capable de caractériser tous les types de noyaux synta-

xiques (nous verrons ci-dessous qu'il existe d'autres éléments syntaxiques qui sont extérieurs à ce noyau). Ainsi il existe au niveau supérieur à la lexie, un schème générateur d'items où toutes les constantes des niveaux inférieurs sont transformées en variables: abstraction du contenu des éléments et abstraction de l'ordre à l'inter-lexical à l'exception de l'ordre: régissant de *binā'*-élément subordonné (sans lequel l'indétermination serait totale sur le plan formel). Aussi la formule que l'on vient d'examiner permet de limiter considérablement la combinatoire syntaxique. Prenons la séquence: # *ḍaraba* 'Abdullahi 'Amran #. Il nous semble aberrant d'examiner toutes les combinaisons possibles des trois segments (étant donné que nous sommes à un niveau situé au-dessus de celui des combinaisons de *kalim*). On est contraint donc de prendre en considération ce fait important relevé par les grammairiens arabes à savoir qu'un **régissant syntaxique subordonne toujours un terme** et de ne considérer dans ces constituants que leur qualité de régissant et de termes régis en premier et en second.

Comme on le verra également ci-dessous ces entités syntaxiques sont aussi susceptibles de recevoir, comme contenu du niveau inférieur, non seulement des lexies mais aussi des segments signifiants et même des unités syntaxiques de leur propre niveau à savoir la formule $(R \rightarrow T_1) T_2$ elle-même. Toutes ces unités constituent alors une substance eu égard à la forme qui les intègre.

Nous nous devons aussi de signaler qu'une telle formule qui relève exclusivement d'un niveau supérieur d'abstraction à celui de la lexie et du segment signifiant n'est pas nécessairement liée à un palier qui serait matériellement supérieur à ceux des autres unités, en fait, **il y a de la syntaxe même à l'intérieur des lexies et jusque dans le noyau de la lexie**: ainsi # *ḍarabtuḥū* # (je l'ai battu) est bien une lexie (verbale) mais elle est analysable en $(R-T_1) = \text{ḍarabtu}$ et $T_2 = \text{ḥū}$ et constitue à ce niveau d'abstraction une structure purement syntaxique.

Autre constatation importante: les variations positionnelles du contenu de T_1 et T_2 sont très grandes - et c'est là une différence fondamentale entre ce niveau et ceux qui lui sont inférieurs - mais elles sont soumises à certaines contraintes. Ainsi, quand les désinences casuelles apparaissent dans la chaîne verbale (pour des raisons phonétiques), la séquence est figée $(R \rightarrow T_1, T_2)$. D'autre part, la nature du contenu de R peut également fixer les positions de T_1 et T_2 : c'est le cas des items appartenants à la classe de "inna". Il s'agit d'exposants invariables du point de vue de leur structure interne (*ḡayr mutamakkin* ou *mutasarriḥ*): cela implique, d'après les grammairiens arabes, le figement de la position de T_1 qui ne doit pas être séparée de R, sauf si T_2 contient un complément de lieu ou de temps (*ḍarf*).

Sur ce même plan relatif au rôle joué par les contenus de ces entités syntaxiques, mais d'un point de vue non plus casuel **seulement mais aussi communicationnel**, il convient de relever cette observation de Sibawayh: "Saches que le *mubtada'* " a absolument besoin que le [contenu du] terme qui lui est relié par le *binā'* coïncide avec le [sien] ou qu'il constitue une localisation spatiale" (*Kitāb*, I, 278). Cet auteur veut dire que lorsque dans la formule précédente $R = \emptyset$ ou exposant (formule de l'*ibtidā'*) le contenu de T_2 doit avoir nécessairement le même référent que T_1 ou référer à un lieu dans lequel doit se trouver le référent de ce contenu. Exemple: # 'Abdullah qā'imun # et # 'Abdullahi fi dārihi # ('A. est chez lui). Mais ceci ne s'applique pas à la séquence où $R =$ verbe non exponentiel.

Dans ce même ordre d'idées, on relèvera également que Sībawayh et ceux qui l'ont bien compris ont soigneusement évité de mettre en parallèle cette structure avec son interprétation sur le plan communicationnel. Comme on le verra, Sībawayh traite de **ce qu'il appelle *isnād* qui est la relation minimale qui doit exister entre un *musnad* (topic ou sujet) et un *musnad ilayh* (comment ou prédicat) pour que se réalise un "kalām mustağni" (phrase complète)**. Or la seule possibilité qu'on a de mettre en parallèle ces deux plans se situe au niveau de la séquence $\mathfrak{R}_{\emptyset, e} - T_1, T_2$ (où e = exposant). Il arrive, en effet, que T_1 et T_2 dans cette structure puissent s'interpréter comme "musnad" et "musnad ilayh" mais cette correspondance ne se fait pas toujours et nécessairement terme à terme: le *mubtada'* et le *mabnī 'alayh* (T_1 et T_2 dans la formule précédente) **éléments formels de la syntaxe**, ne s'interprètent pas toujours comme sujet et prédicat (cas très courant: # 'alā Zaydin daynun #: Z. a une dette). D'autre part de nombreuses lexies peuvent fonctionner comme phrases (#ḍarabtu #, #Zaydan #, # ṣah # (chut!), etc.) et à l'inverse de nombreux *binā'*-s débordent l'énoncé minimal par la présence du *maf'ūl* qui est nécessaire pour qu'il y ait *binā'* mais non nécessaire pour qu'il y ait *isnād* (relation de sujet à prédicat).

On ne cessera de répéter, à ce propos, qu'il n'y a pas d'isomorphisme nécessaire entre les plans sémiologico-grammatical et le plan communicationnel, le premier ne concernant que la structure grammaticale et le code sémique exclusivement ⁽¹⁶⁾.

A suivre

(16) Ces structures et ces schèmes codiques s'actualisent dans le discours et ne reçoivent qu'un seul contenu (sauf ambigüité visée) phonétique et sémantique (historique et contingent). Tout cela a été déjà bien relevé par Chomsky et d'autres auteurs, mais sans en apporter la preuve (sauf pour ce qui concerne Jean Gagnepain et son école)

Références

- ‘Abd al-Ġabbār (al-Qādī), *al-Muġni*, vol. VII, Le Caire, 1961.
- Abū ‘Alī al-Fārisī, al- *Takmila*, Bibliothèque Nationale, Le Caire, manuscrit n° 100
- Abderrahmane Hadj-Salah, *Linguistique arabe et linguistique générale*. Essai de méthodologie et d'épistémologie du ‘ilm al- ‘Arrabiyya. Thèse dactylographiée. Paris-Sorbonne, 1979.
- Al-Ĥalīl (Ibn Aĥmad al-Farāhīdī), *Kitāb al- ‘Ayn*, t. I, éd. A. Darwīš, Baghdad, 1967.
- Al-Ġāhīd, *al-Bayān wa- l-tabyīn*, 4 vol. en 2 tomes, éd. A. Harūn, Le Caire, 1960.
- Al-Ġurgānī (‘Abd al-Qāhir al-), *Dalā‘il al- ‘i ‘ġāz*, éd. R. Rida, Le Caire, 1335 H.
- Al-Mubarrad (Abu l- ‘Abbās), *al-Muqtaḍab*, 4 vol., Le Caire, 1385 -1388 H.
- Al-Raḍī (al-Astrābādī), *Šarĥ al-Kāfiya*, Istamboul, 1275 H.
- Al-Rummānī (‘Ali b. ‘isā), *Šarĥ Kitāb Sibawayh*, 4 vol., microfilm n° 88, de l'Inst des Manuscrits arabes, Ligue Arabe.
- Al-Suyūṭī, *al-Iqtirāḥ*, Hayderabad, 2e éd., 1359 H.
- Ibn al-Sarrāġ (Abū Bakr), *Uṣūl al-naĥw*, Bibliothèque Générale, Rabat, man. n° 3469.
- Ibn Ġinnī, *al-Ĥašā’ iṣ*, éd. M. al-Naġār, 3 vol., Le Caire, 1952-1956.
- Ibn Ya‘iš, *Šarĥ al-Mufaṣṣal*, 10 vol., Le Caire, s.d.
- Kitāb al-Ĥudūd*, éd. M. Ġawād, Baghdad, 1969.
- Lyons (J.), *Introduction to theoretical Linguistics*, Cambridge, 1968.
- Revzin (I, -I), *Les modèles linguistiques*, trad. franç., Paris, 1968.
- Sibawayh, *al-Kitāb*, Būlaq, 2 vol., 1317 H.
- Sirr Šinā‘at al- ‘i ‘rāb*, t. I, Le Caire, 1954.

