Arabic Linguistics and Phonetics[*]
An Introduction to the Neo-Khalilian Theory

Abderrahmane Hadj-Salah
Centre de Recherche Scientifique et Technique
pour le Développement de la Langue Arabe

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-Ḫalil* lbn *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 ancient 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:

 l. The notion of open corpus: Similarly to those of the physician or biologist, the data colleted 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 syntagmatic 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.

الملخص

إن النظريـة الخليليـة الحديثـة انطلقـت مـن الاكتشـاف عنـد أقـدم النحـاة العـرب (في القـرن الثـاني الهجـري) لتصـور علمـي أصيـل جـدا لا يمكـن أن نجـد مـا يماثلـه عـدد النحـاة المتأخريـن إلا عـلى صـورة مشـوهة (إلا مـن شـذ منهـم).

فقـد عـرف أولئـك النحـاة وخاصـة الخليـل بـن أحمـد (القـرن الثـاني) الوظيفـة الأساسـية للغـة وهـي البيـان (communication) إلا أنهـم لم يجعلـوه المبـدأ الوحيـد لتفسـير الظواهـر اللغويـة وعـلى هـذا فإنهـم أقامـوا التمييـز الصـارم بـين مـا هـو راجـع إلى الخطـاب والبيـان ومـا هـو راجـع إلى البنيـة اللفظيـة التـي تختـص بهـا اللغـة.

إن هـذه النظريـة اللغويـة التـي وضعهـا هـؤلاء النحـاة القدامـى قـد نظرنـا فيهـا طيلـة أعـوام ثـم أعدنـا صياغتهـا في إطـار منطقـي ريـاضي حديـث وهـي تقصـد الآن لتعتمـد وتسـتثمر في مسـتوى مركزنـا في مختلـف الميادين.

المفاهيم الكبرى للنظرية الخليلية الحديثة

1) مفهـوم المدونـة المفتوحـة: إن المعطيـات التـي يتحصـل عليهـا اللغـوي لا تختلـف إطلاقـا عـن المعطيـات التـي يتحصـل عليهـا العـالم البيولوجـي أو الفيزيـائي. ففـي جميـع الحـالات تصـح المعطيـات بكونهـا قابلـة للتحقيـق (بالرجـوع إلى الواقـع).

2) إقامـة الفـرق بـين البنيـة النحويـة وبـين وضـع اللغـة (الكـود) مـن جهـة وبـين اسـتعمال المسـتعملين لهـما في أفعـال خطابيـة معينـة مـن جهـة أخـرى.

3) مفهـوم البنيـة في هـذه النظريـة يتجـاوز مفهـوم البنيـة في المذهـب البنـوي الـذي ظهـر بعـد سوسـير: فـإن البنيـة ههنـا هـي نتيجـة لتركيـب بـين الفئـة والترتيـب (بـين أفـراد جنـس ونظائرهـا مـن أجنـاس أخـرى حسـب تعبـير القدامـى).

4) إن وحـدات اللغـة ليسـت بالضـرورة قطعـا مـن الـكلام (أو نـبرات أحيانـا قليلـة). بـل قـد يوجـد في الكثـير مـن اللغـات عنـاصر دالـة مجـردة مسـاوية في الأهميـة للـدوال المقطعـة (الكلـم كالأسـماء والأفعـال) أو النـبرات. مثـال ذلـك بنـاء الكلمـة وجذرهـا: كل واحـد منهـما يـدل عـلى معنـى عـلى حـدة مسـتقل عـن الآخـر وتركيـب المعنيـين يـؤدي إلى معنـى الاسـم أو الفعـل (ولا مـزج في ذلـك ولا انضـمام). وهـذا التركيـب هـو نتيجـة لإجـراء منتظـم للقيـاس (كـما يفهمـه القدامـى).

فالمحور التركيبي يصبح بذلك مجرّدا (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 الصفين.

5) ثـم مـن جهـة أخـرى فـإن التحويـلات التـي تمثـل الانتقـال المتـدرج مـن سلسـلة مـن الـكلام إلى أخـرى تكـون أعقـد منهـا وذلـك بالخضـوع إلى قواعـد دقيقـة (زوائـد لهـا مواضعهـا الخاصـة بتعاقـب أو مـن غـير تعاقـب وتراكيـب تخضـع لمثـل معينـة وتبديـل الموضـع وغـير ذلـك) هـي التـي تولّـد بنفسـها وحـدات اللغـة-خلافـا للنحـو التوليـدي (1957-1965) – حيـث يتوقـف التوليـد عـلى نظـام سـابق وهـو عبـارة عـن أكسـيوماتيكية بسـيطة للتحليـل البلومفيلـدي إلى مكونـات مبـاشرة.

ا

Résumé

La théorie néo-khalilienne a pour point de départ la découverte, chez les plus anciens grammairiens arabes (VIIIe 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-Ḫalīl Ibn Aḥmad* (VIIIe 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).

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 1er système qui n'est qu'une axiomatisation simple de l'analyse en C.I.

Arabic Linguistics came into being in the eighth century A.D. (end of the 1st century of the Hegir), that is, with the beginning of the expansion of Islam. This early start can be explained in terms of the tremendous need felt by the members of the new community to know the language of the Koran, which had become the official language of the young Islamic state (a decision made by the Umayad Caliph *'Abd al-Malik Ibn Marwān* at that date). The appearance of Arabic grammar is intimately linked with that of canon law *(fiqh)*, of Koranic analysis and especially of the Science of *qirā'āt* or Koranic orthoepy. In fact, the first grammarians were all specialists in the oral presentation of the Koran, and the first to whom the idea of grammar is attributed was precisely the inventor of the graphic signs which made it possible to distinguish the different grammatical functions of linguistics elements, namely *Abū al-Aswad al-Du'alī* (d.69H) (the Arabic writing system, as we know, was originally consonantal).

It is in fact the third generation of orthoepists who are first credited with having carried out fieldwork by organising largescale linguistic surveys in order to collect the largest possible amount of data and then, most importantly, having refined and systematized the earlier generations' methods of analysis. This meticulons and systematic research on the language, called *'ilm al- 'Arabbiya* (= the science of *'Arabiyya*) by these first scholars, in which the naḥw constitutes the part dealing with grammar (in the phonetic system), was initiated by *Abū 'Amr Ibn al-'Alā'* (d. 154 H.), as far as fieldwork and largescale critical codification of data are concerned, and by his rival *'Abdallah Ibn Abī Isḥāq* (d. 117 H.) in the domain of language analysis and induction of constants.

The theory of *'Arabiyya* was the work of several generations of grammarians and is almost entirely contained in the remarkable and impressive work of a grammarian of the second century, namely the *Kitāb* of *Sībawayh* (d. 180 H.). But the greatest of these linguists was uncontestably *al-Ḥalīl Ibn Aḥmad* (d. 170 H) *Sībawayh's* teacher (*Sībawayh* quotes him more than 600 times in his *Kitāb*) and we owe to *al-Ḥalīl* the invention of metrics and the very original idea of a dictionary entirely based on the combinatory possibilities of phonemes - and the perfecting of mathematical concepts which could be applied to the analysis of language: permutations, factorials, cyclic group, etc. A large number of explanations for exceptional cases and idiomatic expressions are due to him. Finally, the graphetic system specific to vowels, gemination, etc, still in use today is one of his inventions (he is also the author of a system of transcription for surveys). *Sībwayh's* importance was not only in codifying most of his

teachers' analyses and explanations - in particular those of *al-Ḫalīl* - but also in admirably enriching them and even surpassing them in many cases. Among the linguists who came after this great period of creation (more than 2000 up to the fourteenth century A.D.), it is worth mentioning first the brilliant school which was established in the third and fourth centuries of the Hegir in Basra and then in Baghdad, based on the teaching of a grammarian little known today: *Abū Bakr ibn al-Sarrāǧ* (d. 317 H.). His pupils took up the *Ḫalilian* tradition and enriched it considerably by extremely full personal commentaries and essays (the works produced by this school, still mostly in manuscript form, will probably mark an epoch when they are better known). They include first of all *al-Sīrāfī* (d. 337) and *al- Rummānī* (d. 384), each the author of an enormous commentary on *Sībwayh's kitāb* (5 and 8 volumes), *Al-Zaǧǧāǧī* (d. 337) and *Abū 'Alī al- Fārisī* (d. 384) and the illustrious pupil of the latter *Ibn Ǧinnī* (d. 392) whose genius (like that of his teacher) was almost equal to that of *Ḫalīl* and *Sībawayh*.

It was at this period (end of the 4 th century H.) that philosophical speculation, at first truly Muslim, later frankly Hellenist, was straight away established alongside operationist scientific research and would gradually invade all the Islamic sciences, in particular Arabic grammar. This invasion would consist in a progressive - but never total or definitive - substitution of Aristotle's logic (= *al-Manṭiq*) for the logico-mathematical concepts of the first Arab linguists. Contrary to what we have been able to say so far about the *Ḫalilian* school of the 4th century, its adherents, although they did not refrain from using concepts taken from the *Manṭiq*, were always (or mostly) careful not to confuse them with those taken from the mathematical logic of *al-Ḫalīl* and *Sībawayh*. But the same was not true of their colleagues - shallower and much less original - and most of their successors.

After 470 (date of the death of a famous continuator of *Ibn Ǧinnī*, namely *'Abd al-Qāhir al-Ǧurǧāni*). Arab thought in the domain of stagnation which would last several centuries, and it was really only with contact with western civilisation that it would have the chance to become aware of this lethargy and consequently try to overcome it. In the domain of the scientific study of language, most Arab researchers of our time have simply attached themselves to one or another of the western schools of linguistics. It could be - and we heartily hope it to be the case - that some of them manage - (by extending the viewpoint they took as a starting point) to achieve some original work. A second method consists in elaborating the concepts of the Arabic linguistics

of the first centuries of the Hegir (after having grasped all their contours and the principles on which they are based) in themselves on the one hand, and in opposing them to the concepts of contemporary linguistics on the other hand. This presupposes, of course, a deep and extensive knowledge of these concepts, an *objective* and *critical* knowledge (like that which we ought to have of ancient Arabic grammar, of Indian grammar, etc.), that is without prejudice. It is this second method which my colleagues and I have attempted to follow for more than two decades, within the movement which has been given a name we are willing to accept - the neo-*Ḫalilian* School.

The Methods of Research of the 'Ilm al-'Arabiyya and Their Scientific Basis.

The first scholars who took on the task of codifying the *'Arabiyya* were led to carry out surveys in the field of the "*Faṣāḥa*"[1] in order to gather the largest amount of data deriving from this language. The full set of data constituted a kind of corpus (called by them al- *samā'* or *al-masmū'*: the set of data collected ex-auditu), but their view of the corpus is fundamentally different from that of the empiricist structuralism which appeared in the west after Saussure. Empiricist structuralism, as we know, advocates basing a description only on a set of utterances given once and for all "in order to avoid temptation of calling up occurrences which would fit in with the investigator's own theories". The *Arabs' samā'* was instead always open, in the sense that any linguist could at any time note down one or more occurrences. In fact the linguist studying the *'Arabiyya*, while making his analysis, referred not only to the information he had collected himself, but to the whole *masmū'*, all the data collected by others up to that moment. On the other hand this *masmū'* was, for him, completely intangible: if he was allowed to increase its size by incorporating the product of his observations into it, it was only at the price of an *iğmā'*, that is an agreement between his own observations and those of all his colleagues[2].It follows from this that nothing can be stated about the tangible existence of an item, a pattern of occurrence, or the extent of its use, without a precise reference to the full *masmū'*, *thus made objective*. Recourse to a set of data considered as the sum of all observations and made objective by a real *consensus* seems to us infinitely closer to the attitude of the serious scholar in rigour as well as in fruitfulness**.**

This set of linguistic data which served as a base for the grammarians' descriptions was obtained through numerous *surveys* which extended over almost all the Arab peninsula. The surveys which provided the largest part of the

data were those carried out by the originator of these surveys and his followers, namely *Abū 'Amr Ibn al 'Alā'*. The regular or occasional informants has to be *native speakers* who had not stayed too long in areas where dialects other than the *'Arabiyya* were spoken (they were called *fuṣaḥā' al-'Arab*).

The investigators of *'Arabiyya* made use of very elaborate techniques, the main features of which were exclusively receptive observation (*istilġā'*) and active observation; the latter could be carried out through two processes:

Stimulation-suggestion *(called talqīn)* and questioning about aspect of language. In the first case the informant is given a series of stimuli in order to encourage him to talk and suggest to him a topic of conversation (the aim being to obtain instances of variation at all levels). In the second case the aim is to obtain information about the language; factual questions are thus used. Here again there are two methods: the questioning was carried out either by means of an alternative question, where the speaker was asked to choose between two possibilities, or by means of a test-question or stimulus-sentence which constituted a real test, or a "provocation", in the form of a sentence which was given to the informant without further comment.

All this was naturally recorded in writing: the great linguist *al-Ḫalīl Ibn Aḥmad* had moreover developed, for this purpose, a system of transcription which was very convenient for noting the variations.

In the Islamic sciences, the rational systematisation of facts soon took the form - especially among linguists - of a rather remarkable combination of two opposing, and apparently contradictory tendencies: an extremely, meticulous concern for the facts, and a very pronounced obsession with abstract constructions.

For Sībawayh, for example, the combination consisted to the test (which was always the deciding factor) but in also giving great importance to *hypothetico-deductive* constructions and to formalisation in general.

The important concepts on which this step is based are the notions of *bāb*, *naḍir*, asl and *far'*, *istimrār* or *ittirād* and above all that of *qiyās*. The Arabs give the name *bāb* to any set of objects or processes having in common, *not a simple property*, but *a shared structure*. Hence the possibility for a *bāb* to be empty, when there is no linguistic item in the analysis which can correspond to the structure it characterises, or monary when there is only one such item. This is the case with the objects characterised by the lexical structure:

$C_1iC_2uC_3$ (where C_i = the root consonant in position i; what the Arabs call

fi'ul), a structure which is required by the lexical combinations of the *'Arabiyya* but which is nonexistent in actual usage. Likewise, hence the impossibility of applying this term to classes of concrete realisations thus a simple set in the logical sense of the word, but a set whose constitution and/or combinatory rules. A *bāb* is thus made up of homologous elements: these are *naḍa'ir*, plural of *naḍīr*. A more abstract bāb is one which incorporates other *bāb-s* having the same structure but a higher level of abstraction. This is the reason why the name bāb is given to the patterns generating items (lexical or syntactic) at different levels of abstraction, as well as to the ordered sequences of consonants which constitute lexical roots.

The notions of *aṣl* and *far'* depend one an extremely interesting idea, in that they make possible the establishment of an *order* on the paradigmatic axis (which constitutes a real "wastepaper basket" in contemporary functionalism). In fact, the term *aṣl* is applied to any element which also invariably occurs in other forms of elements which are its *furū'* (plural of *far'*) and which contain it and overlap it by virtue of some kind of material and/or formal addition. Everything we have to say about the Arabic linguistic models will serve as ample illustration of these important notions.

As for the *qiyās*, it is in, the fist place an *equivalence* (in the mathematical sense) which can be established between two or more structures (hence between two or more *bāb-s*). This presupposes a constituent analysis of the elements and the establishment of biunivocal correspondences between these constituents. But that is far from sufficient, for it is also necessary for this correspondence *composition* of the elements which are thereby set together. The *qiyās* can be applied to objects-items in their structure - but it is more interesting when applied to *processes*, in particular to transformations (in the broad sense, covering the Harrisian and the Chomskyan transformations); isoschemism thus becomes true isomorphism, and the systems of operations thus placed in correspondence are then necessarily structures which are closed in on themselves, that is group structures.

The concept most similar to *qiyās* is that of analogy, but it involves a much more elaborates kind of analogy (as we have just seen), for it applies to *structures* (in the *mathematical* rather than the phonological sense). It must therefore: (1) allow a *stimulation* of reality by the construction of models (the generative patterns, = *muṭul*, plural of *miṭāl*). (2) reveal much more abstract structures (by a series of equivalences).

The construction of generation patterns (the *bāb-s* transformed into *qiyās* and called miṭal-s) should not lead us to forget that these *bāb-s* and these *qiyās* are first of all perceived in reality (it is only after being stripped of content that they become real simulation instruments). But in reality there exists another kind of uniformity which is perceived, not within a *bāb*, but between two entities which are linked in such a way that the presence of one is always or almost always accompanied by the presence of the other (e.g. subject – mark-u. This is law in the Baconian sense, but by introducing the *qiyās* into this circunstancial or invariable relationship, we produce what the Arabs call *qiyās* mustamirr, that is, a uniformity which is circumstantial and congruential at the same time.

Theories and Models of the *Naḥw*

1. Discourse Acts and Related Concepts

As we have seen the grammarians started out from direct observation of the language behaviour of native speakers of the *'Arabiyya* and were then led to imagine the different components of the communication system, rather like our contemporaries. In any exchange of messages (*ḫitāb)*, there is a *"mutakallim"*, sender of the message, necessarily unique, and a *"muḫāṭab"* target of the message, which can be multiple: between them passes the *ṣawt* or sound which carries the message (or a substitute such as writing or something else). The mutakallim or speaker determines the *"ḥāl al-ḥadīṯ"*: this is the present situation with regard to the message, the situation which serves as a reference for the *muḫāṭab*, in that it constitutes the starting point for the spatio-temporal situating of the communication. But the communication is guaranteed only if the participants share the prerequisite knowledge, given and not deduced, of the *"waḍ' al-luġa* (or *muwāḍa'a*) which is the language code (later *iṣtilāḥ al-taḫāṭub)*. This *waḍ'* is a real "social, convention": considered as such is any articulated sound *(lafḍ)* which is institutionalised and which can be opposed to any *lafḍ* which has not been established as a signifier *(dalīl)* denoting something signified. This is the case with sequence of phonemes such as **saṣ ,*ḍat, *qāǧ*, etc, which have not been adopted by the *"wāḍi'"* or founder of the language (for reasons of phonetic incompatibility). But the code cannot be reduced to more correspondences between signifier and signified, there are also all the arrangements of signifiers which are also conventional but which can be inferred from one another by means of rules which are the *"maqāyīs"* (here the plural of *qiyās*) of the naḥw. Here resides the difference between the *luġa* (or datum of the language) and the *naḥw* (hence the opposition between the *luġawī*, who has the task of collecting

and verifying linguistic data in the field; and the *naḥwī*, who is specialised in the analysis and systematisation of these data). The *waḍiʿ* is contrasted to the *istiʿmāl*, which is the implementation and actual use of the language in utterances. A principle maintained by the Arab linguists in this connection is that not everything allowed by the *qiyās* or congruential system of the language is necessarily found in actual speech.

"Language was invented" claims *Ibn al-Sarrāğ*: (teacher of Abū ʿAlī al-farisī) "mainly for the transmission of information *(fā'ida)* … but if one should take it into one's head to say: "fire is hot", snow is cold", one would have on utterance with zero *fā'ida*" (cf. his *Uṣūl* I, f.43). *Fā'ida* thus appears to be a quantity of information which can be positive or null; the message coveys something to the addressee, or it conveys nothing. It is then called *"mufīd"* or *ġayr "mufīd"*. Hence the importance given on the one hand to ambiguity or labs and on the other hand to the knowledge or data possessed by the addressee *('ilm-al-muḫātab)* and to frequency of usage *(kaṯra)* in explaining the phenomena of omission *(ḥaḏf)*, redundancy *(Ziyāda, tawkīd)* and those relating to context. Thus the ambiguity or labs of certain utterances such as "*kāna insānun ḥalīman"* (a man was noble) (example analysed by *Sībawayh*, *Kitāb* I, 22) can make the message completely predictable and probable (the information, in this example, relates to at least one member of the class of human beings). It will also be noted that only the elements which can be supplied by the context are liable to be omitted; and it is those which cannot be omitted which alone carry the information which the addressee lacks. *Sībawayh* also says: "speakers omit or leave implicit things which are frequent in speech, for they need to reduce what is used very frequently" (*kitāb* I, 294). As for redundancy, the Arab grammarians consider that there are two kinds: redundancy originating from the system, and due to an excess of distinctive functions: *ziyāda li-l-farq* (e.g, case agreements in unambiguous utterances) and accidental redundancy due to variations in the circonstances of communication: there is a surplus of signifiers in order to counteract the deficiencies of communication (*ziyāda li-l-bayān aw al-tawkīd*) in both cases there is predisposition (*iḥṭiyāṭ* V. *Ibn Ğinnī, Ḫaṣā'iṣ*, III, 101-111). This *ziyāda* is said to be mufida, i.e. functional, since it ensures the reliability and efficiency of the communication. This said, the Arab linguists did not base their theory of language entirely on this notion of "*mufīd*" as is currently the case whit the functionalists.

2. Acceptability, Grammaticality and Semantic Interpretation

Nothing is found in the *naḥw* which has not been originally observed in the speech acts (or *kalām*) of Arabic speakers: "The *naḥw*", Ibn al-Sarrāǧ tells us in this connection, is a science which the ancients established on the basis of repeated, systematic observation (*istiqrā'*) of the Arabs' *kalām*"(*Uṣul*, f. IR). Although essentially based on factual observation, the *naḥw* nevertheless cannot be reduced to a simple description of the language, for as *Abū 'Ali, Ibn al-Sarrāǧ*'s pupil, makes clear, it involves "the science of *maqāyīs* (here plural of *qiyās*), abstracted (or induced = *mustanbaṭa*) by *istiqrā'* from the Arabs' speech (al-Takmila, 1). Now the qiyās, as we have seen, is at the same time a constant (a law established through observation) and also generative pattern, or a model which makes it possible to generate an infinite number of grammatical utterances (as conceived by N. Chomsky, but with completely different kind of recursivity). Drawn from or confirmed by experiment, the *maqāyis* make possible the prediction of other discourse acts and the explanation, by means of their structural correspondences, of many phenomena observed in actual discourse.

It is for this reason that the idea of grammaticality - in the exclusive sense of "conforming to the requirements of the *qiyās* is not unknown to the Arab grammarians".

"The *nahw*", claims one the grammarians of the fourth century H., is a scientific discipline which makes it possible to understand the different states (*aḥwāl* = descriptions) of the Arabs' language with regard to the validity of its arrangements, and to distinguish, by this means, the correct utterances from the incorrect ones "(quoted in Iqtirāh, 6) (in reference to the language to the language behaviour of the Arabs since *Al-Rummānī*) (cf. his *Ḥudūd*, 38). In other words, to distinguish the utterances which belong to the *'Arabiyya* (and not to a norm associated with a restricted and privileged social group) from those not belonging to this language, and it is only in this sense that the term *acceptability* was used *(ḥusn, istiḥsān)*. This meant acceptability with regard to the largest number of native speakers (hence the degrees of acceptability expressed by the terms *fasīḥ/afṣaḥ, ḥasan/aḥsan, qabīḥ /aqbaḥ:* these are the reactions of native speakers which the grammarians recorded in their surveys and which they combined by means of statistical lists.

Sībawayh was one of the first of these linguists to point out the relations existing between grammaticality, acceptability of utterances and their semantic

interpretation. "*The kalām*", he says, "can be: *mustaqīm ḥasan* (grammatical and acceptable), *muḥāl* (meaningless)" (al-Sirafi, his annotator, specifies: *mustaqīm muḥāl* = grammatical and meaningless) *mustaqīm kaḏib* (grammatical but unlikely), *mustaqīm qabīḥ* (grammatical and not acceptable, except in poetry for certain forms), *muḥāl kaḏib* (meaningless and unlikely) (cf. his *kitāb* I, 8).

3. Semiologico-grammatical V. Semantic

One of the most important achievements of Arabic linguistics (that of *al-Ḫalīl* and his followers) was the very clear distinction it drew between the strictly semiological and grammatical analysis of the language and that of the meaning deriving from the act of utterance. It was not a question of giving more attention to one at the expense of the other, or even of completely separating the two, but only of avoiding the unfortunate confusion (into which many old and new approaches have fallen) between what derives from the *waḍ'* (Semiologico-grammatical system = signifiers/signified + arrangements) and what belongs to the *'isti'māl* (actual use of the *waḍ'*) with all that implies with regard to communicational framework, logico-semantics, etc.). This distinction is moreover based on an observation (first made by *Sībawayh* in his *kitāb*, I, p 15-16) relating to the *lafḍ* (articulated sounds) and to the *ma'nā* (meaning) when they are combined by the *waḍ'*. The former is then dephoneticised (in this union in the *waḍ'*) to varying degrees: vowels and affixes at the level of lexeme patterns are variables. In the same way the *ma'nā* is desemanticised to varying degrees: the nominal and verbal lexical ends have generic content at the level of the code, and at the level of structures (roots, lexical and syntactic patterns) the content is still more abstract[3]. Finally, the phenomena of synonymy and homonymy - which are essential and not accidental - make the *waḍ'* - code and the *waḍ'* - structure (semiology and grammar) into an entity which is *totally available* and able to fulfil all links of needs. On the other hand, the meaning which the addressee can infer from the situation or from information obtained earlier or provided by intuition, or intellectual knowledge, is *ma'āni* (plural of *ma'nā*) which do not depend on the *waḍ'*, the language but on the infinite domain of semantics (see in this connection *Sībawayh*'s remarks about the information the *lafḍ* can provide, which he contrasts with that which is inferred from everything except the *lafḍ*. (*kitāb* I, 15-16 and also *Ibn Ǧinnī*, *Ḫaṣā'iṣ*, II, 184, 321 and III, 98 ff.). Similarly, language is not defined by its communicative function (or by any other function). This is what is said explicity by a philosopher of language who has properly understood the spirit of the khalilian school: "In saying that speech

is constituted of what is mufīd (what carries information), we do not mean that it must necessarily carry a "*fā'ida*"(information) every time it occurs, but only that it is that by which the *fā'ida* can reasonably be realised, although this property may sometimes be lacking because of a certain state speaker" (*Muġni*,VII, 11).

4. The Syntactic Model

Essential contrasts with the structuralist or generativist approach:
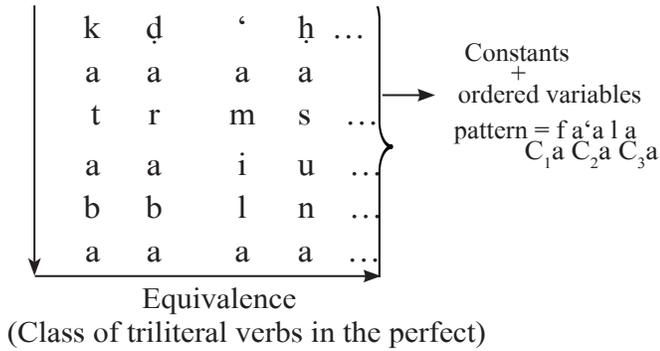
4.1. Synthesis of Class and Order

Form and substance are concepts the Arab grammarians knew well (in Arabic: "*Sūra*" and "*mādda*", which are the translations of the terms used by Aristotle), but they did not content themselves to see in form a simple system of opposition; form for them is, in addition, an arrangement of these oppositions, or, in other words, a structure where the members of one class are directly related to those of one or more other classes. The use of the aristotelian terms for grammarians meant just this, that every structure is opposed, at a lower level of abstraction (not necessarily by embedding), to the set of elements which it "informs", and at the higher level, to a wider structure in which it constitutes an element. Thus the simple opposition of the aristotelian-saussurian type is not the only thing at work in the immanent system of the language, even if it is embellished with a secondary dimension called "contrast". A further requirement is that the elements of the system must be placed in correspondence with other elements belonging to other classes, and it is only then that they acquire, through their *position* (= their "*mawḍiʿ*) in this extensive system, the status of absolute variable (not bound to a class).

4.2. The Four Kinds of Denoters

The Arab linguists identified in the "'*Arabiyya* four kinds of units or, more precisely, four types of denotation: by the root and the pattern the signifying segment and the zero marker.

The *signifying segment* is what the Arabs call *"Kalima"*. It is tempting to see it as the result of a simple segmentation (with commutation) which constitutes a simple analysis of the text and culminates, in western linguistics, in the "morpheme" we shall see below that the *Kalima* does not result from such a simplistic analysis. The Arab grammarians contrast this unit with two other denoters: the root and the pattern, which they obtain by a vertical analysis, in other words by an induction of *ordered variables* on the basis of two operations which are carried out *simultaneouly*: firstly, the establishment of equivalences

between signifying segments, and secondly, the seriation (or ordering) of this equivalence class. The operations for obtaining a pattern can represented as follows:

$$
\begin{array}{ccccc}
k & \d{d} & \text{`} & \d{h} & \ldots \\
a & a & a & a & \\
t & r & m & s & \ldots \\
a & a & i & u & \ldots \\
b & b & l & n & \ldots \\
a & a & a & a & \ldots
\end{array}
\quad\Bigg\}
$$

→ Constants
+
ordered variables

pattern = f a'a l a

$C_1 a\ C_2 a\ C_3 a$

Equivalence
(Class of triliteral verbs in the perfect)

The same procedure is followed to obtain roots: only the equivalence class need be changed.

However, there exist some signifying segments which cannot be analysed in terms of root and pattern. This is the case, first, with all the grammatical morphemes - called *ḥurūf al-maʿānī* or *adawāt* (function-words) - to which we can add all the segments which function as second power markers (substitutable for full nouns), namely the deictics or shifters (*al-'asmā' al mubhama*).

Thus, not everything in the language is segment or accent; each of the entities established by the vertical analysis of the *kalima* possesses its own *dalāla* or denotation in *"kataba"*, *"kātib"*; *"maktūb"*, the sequence / K.T.B. / by itself denotes, in the code of the *'Arabiyya*, the signified "write", and the patterns fa'ala, *fā'il*, *maf'ūl*, specify the signifieds: "completed" "agent", "patient". Thus the kalima when it is analysable into a root and a pattern, is not a minimal signifier (it is such only at the level of the text: none of its components alone denotes its signified).

As can be seen, denoters are not necessarily segments. To insist at all costs on analysing a text merely by segmenting it, as the postsaussurian linguists do, leads to dead ends or unsatisfactory solutions. Such is the concept of "discontinuous morpheme" proposed by those determined to account for pattern variations. The most striking example in this respect is the internal plural in the 'Arabiyya: how, in this out-an-out segmentalist framework, could we explain the shift from *"sāḥib"* (companion) to *"'asḥāb"*: where in this case is the segment which denotes plural?

A last type of denoters identified by the Arabs - and by our contemporaries -

is the zero marker, which Sībawayh called *"ʿalāma ġayr ḍāhira"* (unexpressed marker) or *"tark al-ʿalāma "* (*kitāb*, I, 123, 220). This involves, the meaningful absence of an expression at the level of sound. It characterises, for example, the affix pronoun of the third person of the verb, and the nominal markers for masculine and singular, as opposed to those for feminine and dual/ plural which are represented by the adjunction of a full marker. It can also be applied at the syntactic level: the governing zero, for example, which corresponds to the *ibtidāʾ*, is the marker of the syntactic unit which underlies the simple noun phrase.

Much later than *Sībawayh*, some authors, influenced by philosophy, did not understand that the absence of a marker could have a governing effect. Unable to conceive the phenomena in an operational framework, these grammarians could not make the difference between an absence considered in itself, and an absence considered *within* a structured whole; it is in the latter case that the role of this absence is seen as the fact that it affects a **position** in this whole (cf. the role of the concept of zero in mathematics).

4.3. The Concept of *Mawḍiʿ*

It is just this concept of position within a structured whole (not a system of simple oppositions but a whole in which all the elements contained there are placed in biunivocal correspondence), which is, along with some other concepts, at the base of the analysis of the *naḥw*. Another term, mawqiʿ, is sometimes used in place of the word *mawḍiʿ*; they refer to the "place of occurrence" of an element in language. Al-Rummānī talks in this connection of the "qismat *al-mawqiʿ*", that is the distributional combinations of the elements of language. Could the analysis of the naḥw be called "distributionalist"? We shall see that it involves a much more complex "distribution". In fact, the *mawḍiʿ* is not always identical to the physical position an element can occupy in language (thus, the anteposition of the complement does not change its *mawḍiʿ*). The *mawḍiʿ* is not defined by distribution alone, or by the function of the elements occupying it: it is a position virtually contained within an operational pattern abstracted on the basis of both the syntagmatic and paradigmatic axes. More light will be shed on this concept in the analysis of models of the naḥw (which involve it) to be described below.

4.4. The "*lexie*" as the Basis for Analysis

"The *ism muḍhar* (= overt noun, as opposed to *muḍhar* = personal pronoun), *"Sībawayh* declares, "is never composed of a single *ḥarf* (sound segment), for it must be possible for it to be *followed by silence* and at the same time not *preceded*

or followed another element..." (*kitāb* II, 304). Any verbal string which can be separated from what precedes it (*'ibtidā'* = attack) and from what follows it (*'infiṣāl)*, and considered as minimal with regard to this isolatability (*'infirād)*, constitutes a formal unit, since it is extracted from a single signifier, which the same author calls *"kalima mufrada"* and *"bi-manzilati kalimatin wāḥida"* [4].

The first of these labels is applied to isolatable (and indivisible) strings which contrain no substring which is isolable in other contexts. The same is true of the *ism*, overt and covert (noun and isolatable personal pronoun). The second term is applied to strings deriving from this by one or more expansions which are bound to it in such a way that the string remains indivisible.

Infirād or isolability, thus understood, is of capital importance, for by demarcating potentially autonomous sections of speech, it allows us to reach the *truly* fundamental unit which is located at the intersection of syntax and lexis - and even at the intersection of the syntactic and the communicational, since each section thus defined an does function as a *minimal message* (like a sentence, whatever the number of elements it may contain, the important things being the two pause-breaks as boundaries and the indivisibility). From this level, the central one, the linguist's analyses will thus more either downwards, to look for the signifying segments and their components, or upwards, to see how the minimal sequential units fit into syntactic structures.

This method makes it possible to avoid the arbitrary nature of many current or earlier approaches which take as their starting-point the sentence or the proposition (which is not a given formal unit). Certain contemporary authors have devoted their attention to defining a formal starting-point: this is true of Harris (see Lyons, *Introduction*, 5, l, 2) and of Revzin (see his *Models*, French translation, 15), but none of these authors has tried to exploit fully the concepts he has established.

It is just this possibility of generation *(tafrī')* of derived structures *(furū')* from an elementary string *(aṣl)* by adjoining meaningful expansions *(zawā'id)* to the right or left of the later that the Arabs call *"tamakkun"* (capacity) and *"taṣarruf"*(variability). This makes it possible to establish distinctions (which have the advantage of being purely formal) between isolatable strings. Certain strings have a perfect tamakkun; they can receive all kinds of expansions. This is true of 'ism ǧins (common nouns). Others are less, or much less perfect: the 'alam or proper noun, for example, which cannot take an article or an adnominal complement. In addition, certain adverbs can receive only one or two declension markers *(i'rāb)*, or none at all, and so on.

The *taṣarruf* or variation (additions) from the initial string is subject to rules whose (ordered) application can be represented as follows: See the following figure (Ḥadd of the ism).

THE ḤADD OF THE ISM [5]
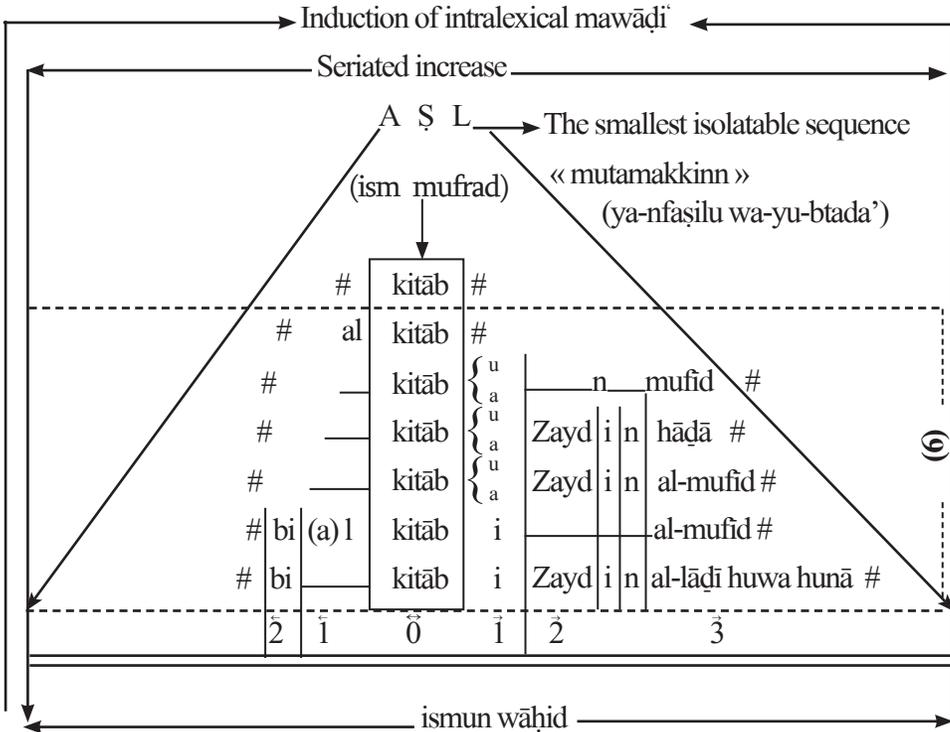
As we can see, the expansions are additions which can appear in one position and also alternate among themselves. Case inflections, *tanwīn* (which alternates horizontally with the article and vertically with the adnominal complement) the qualifier to the right of the kernel on the one hand, and the article and preposition to the left, are all additions with regard to the kernel which they can affect and in the sequence of which they can appear and disappear. We can also note the presence of zones of recursively (*iṭāla* by *'aṭf, takrīr* or *taṯniya*) at this level; these are the *mawḍi'-s* of the adnominal complement (2̈) and that of the qualifier (3̈). we shall see that it is in these positions, among others, that the embedding of strings is possible.

Thus the criteria of *infirād (ibtidā', infiṣāl)* and of tamakkun make possible the recognition, within the lafḍ (the signifying form) and *only within the lafḍ*, a first unit, which is here the ism *'āmm* (or *ǧins* = common noun). The element having the greatest *"tamakkun"* will be that which is *totally unmarked* in its minimal realisation (its *aṣl)*. Thus we have here a formal definition of this kind of unit; any string which allows all possible expansions to left and right without the whole there by losing the status of an indivisible string with regard to its realisation from this starting-point the other units contained within the

expansions can be formally defined. This is possible because they occur in specified positions which are inferred from the examination of the possible occurrence of a given element. These *"mawḍiʿ-s"* in their turn determine-still in formal terms- the grammatical functions of the elements which occupy them.

It is also in this framework that the *"kalima"* can receive a formal definition. In fact, any string of sounds which can occur in one of these *mawḍiʿ-s* is considered a kalima it follows that the "*kalima"* is not necessarily identical with the concept of "morpheme", in fact, the kalima is indeed a signifying segment (which is reached, as we have seen, by the elaboration of the generative pattern of the lexie) but this segment is minimal only with regard to the lexie pattern which contains it. It is a component of the lexie, whereas the morpheme is a unit of meaning (almost always equivalent to a segment) regardless of the level on which it is situated. Hence the confusion made by the descriptivists between the meaningful elements which go into the *making* of the system generating lexemes *(*the *miṯāl-s)*: *faʿala, iftaʿla, istifʿāl*, etc.), which are elements internal to the lexeme, and the meaningful elements external to this pattern, the latter differing formally from the formar in being separable by deletion in the same pattern.

On the other hand, the Arab grammarians consider that the element most able to undergo expansion is also the least costly to realize *(ʼaḫaff)*. Such is the case with the common noun. The verb is the least economical (ʼaṯqal) because it cannot be realised without a subject, inflection, tense realisations, etc. The ṯiqal designates just this expansional load, and consequently is located on both the physiological and the psychological levels.

As can be seen, the paradigmatic axis is necessarily, structured, in that it is the site for transformations, and the latter are necessarilly arranged in a hierarchy, in *aṣl→ furūʿ*. The expansional, transformation or *ziyāda* determine many distinctions, paradigmatic relations, but the latter must not be considered when we are dealing with the syntagmatic axis within a single morphosyntactic class. On the contrary, they must be considered in the structured whole which results from the combination of the two axes. It is thus necessary to consider together without separation, all the columns "where paradigmatic relations are found, firstly in the syntagmatic *order* they exhibit, and secondly in the *movement* which renders the two axes *dynamic* through the progressive and augmentative passage from the minimal string to its derivations and vice versa, in other words in a perspective of *totality* or of a structure closed in on itself and having its own

properties (we shall) see that all the patterns of the *'Arabiyya* constitute such totalities which prove to be group structures). Here again there is a synthesis of class and order. This pattern generating lexies, like all the other patterns, at whatever level, constitutes from one point of view a bāb, since it is a whole structured by an orderly (= generalised) establishment of equivalence of the aṣl and its furūʿ(all the derived strings constitute lexies for the same reason as the starting-point: they are its isotopes; but it also constitutes a *qiyās* and a *ḥadd* (operational definition) since this biunivocal and reversible establishment of correspondence constitutes a model of derivation and thus of characterisation and of ordered generation of linguistic items.

Beside the nominal lexie-generating pattern, there exists another pattern which generates verbal lexies (in fact, three different sub-patterns which correspond to the three verbal moods: perfect, imperfect and imperative).

Before moving to the other levels, situated above and below this central level, it seems approriate to mention two other concepts which refer to the linking cohesion between the components of the same unit. These are the waṣl (or *'ittiṣāl*) and the *binā'*. The Arab grammarians noted that this cohesion becomes looser and looser as we pass from one level of analysis to the one above (see among others the šarh of *al-Rummāni*, II, 86 R and 59V). Thus as far as the level of lexie is concerned, this involves only *"waṣl"* which is a simple concatenation (juxtaposition without effect). The expansion which appear in the pattern are merely concatenated; their disappearance never affects the lexie in which they appear, which keeps its status as a lexies as long as the head remains. This alternation with zero which does not affect the lexie is also found at the level of the syntactic peripherals which are also expansions. This kind of simple concatenation is not realised at the level of lexemes (internal to the kalima which is analysable into root and pattern) or at the level of the syntactic kernel, as we shall see below; in these two cases we have a very strong cohesion which is called *binā'*. Here we have a construction involving the integration or combination of two or more elements within the same structure; this is true of the elements of the kalima which are integrated into the pattern which generates it. This is proved by the fact that the deletion of any one of the components destroys the whole unit: mukrim-*krim (or at least, in some rare cases, it causes the pattern to change: *tafaʿala - faʿala*). There also exist intermediate levels where the waṣl is not as loose but where at the same time the *binā'* is not as strict. This happens where we have agglutination of the kalim (in this case we talk of *ḍamm*): the

markers of the feminine, the dual and the plural are linked to the nouns as in a *binā'*, but they can disappear without breaking up the unit.

As far as the level above that of lexie is concerned, it will be noted first of all that the units found there are not the result of simply combining lexies. Neither the *lexie* nor the *kalima* constitutes the minimal unit at this level. Moreover, the relations between the elements at this level are quite different.

With regard to the sequences of lexies attested in speech:

(1) # *'Abdullāhi qā'imun* # and (2) # *qā'imun 'Abdullāhi* # 1), *Sībawayh* claims that there is *binā'* between these two lexies and not a simple *waṣl*, which is correct since the deletion of one of the two lexies (without reference to context) destroys the unit. It will be noted, besides, that the same strings can be found in larger strings: (3) # Inna *'Abdullāhi qā'imun* # (4) # *kāna 'Abdullāhi qā'iman* # (2). It is clear that (3) and (4) derive from (1) by *ziyāda* or adjunction of "Inna" and of "*kāna*".

Given that there is the same relation of *binā'* between these lexies, and since (3) and (4) derive from (1), it is possible to draw the correspondance between them term by term:

(1) # ∅	'Abdullāhi	qā'imun	#	
(2) # inna	'Abdullāhi	qā'imun	#	
(4) # kāna	'Abdullāhi	qā'iman	#	(7)

It will be noted that the elements occupying the left most column (at the beginning of the strings) seem to be related to the inflections contained within the lexies. This relationship is rightly considered by the Arab grammarians as *government* (ʿamal) [8]. The governing elements in fact determine the inflectional markers of the governed elements. This allows as to compose this set to another sequence of lexies containing a verbal lexie: (5) # *ḍaraba 'Abdullāhi 'Amran* # [9] where "*ḍaraba*" similarly functions as a governing element (and where *'Abdullāh* is considered as the subject (*fāʿil*) of the verb (*fiʿl*) "*ḍaraba*"). It will be noted on the other hand that, as opposed to the other strings, (1) involves the zero expression of the governing element, and it is this zero expression which the grammarians call *ibtidā'*.

On the other hand, the same grammarians raised an important point, namely that there exists one element among these governed which *can never be anteposed to its governing element*: this is the element governed with regard to naṣb (mark a) by those of the class "inna", and for rafʿ (mark u) by those of the

class "*kāna*" and "*ḍaraba*", in other words all the items which appear in second position in this set. With regard to the *binā'* required by the zero expression of the governing element this is the *mubtada'*, that is the item *governed by zero* (= *al-'ārī 'anī l-'awāmil al lafḏiyya*) 3) which corresponds to the term which cannot be anteposed. The item obligatorily governed in postposition is what *Sībawayh* calls "awwal *mā tašġalu bihi al-'āmil*" (*kitāb* I, 245 and also 41) = the term which first "absorbs" the governing element. This subordination (order + dependence on what precedes) is simulated by *Sībawayh* by means of the strings we have just seen by:

R (Syntactic governing element) → T₁(term first governed, T₂ (term governed second).

Where only T, has to be placed after R in the sequence. In speech we can thus have the following variations: (R, T_1, T_2), (R, T_2, T_1), (T_2, R, T_1). We must however understand that the *binā'* or structural integration of one element with another is not between R and T_1, but between the ordered pair $(R - T_1)$ and T_2. It will be noted that this ordered pair can be found alone, without T_2, in speech (as in # *qāma 'Abdullāhi* # or # *qumtu* # (I go up). Finally, be content of these entities must be interpreted at the level of case as well; in R we must have either zero or what is called an exponent verb, such as *"kāna"*, which is a true temporal exponent, or a non-verbal exponent of the class of "inna" (the corroborative particle): layta (wish), *la'alla* (expectation), *Ka'anna* (comparison), etc., or a non-exponent verb, Such as "*ḍaraba*". The content of R in fact determines the case content of the governed terms. Thus, if R = ∅ ,T_1 necessarily contains a *mubtada'*[10] which is at the formal level the name given to the content of T_1, but which at the level of case can be interpreted as the subject of a ḫabar, the latter, which is the content of T_2 in this kind of structure, being interpretable as the item carrying information *about* the term assumed which is the content of T_1. If R = exponent (verbal or otherwise), the kernel of the string does not change since these exponents are assigned to it as such. We talk only of "ism and ḫabar of *kāna* or '*inna*" = noun and ḫabar governed by these exponents. Finally, if R= non-exponent verb, we obtain a string which, although homologous to the preceding one, nevertheless has its own properties. T_1 must then have a subject (fā'il) and T_2 an object complement (*maf'ūl*), T_2 then being liable to delete.

It has been possible to object that since # *qāma* #'*Abdullāhi* # (= A got up) and # '*Abdullāhi qāma* # have the same meaning, it should therefore be permitted to antepose T_1 to R (their content here being subject and verb *fā'il, fi'il*). Several

arguments have been put forward in response to this objection, the main point of which can be summarised as follow. (Mubar., Muqtaḍab, IV, 128):

"There exists a position or *mawḍi'* after the verb which can be occupied only by its subject. Now a *mawḍi'* cannot be eliminated; it can simply be left empty. This is clearly shown in the following pairing:

#		'Abdullāhi	qāma	∅	# =	'A.	he	got up
#		'Abdullāhi	qāma	'aḫūhū	# =	'A.	his brother	got up

This same alleged subject at the grammatical level is capable of being governed by another element, so there exists a *mawḍi'* preceding this type of string which is here empty.

#		∅		'Abdullāhi qāma		# = 'A. , he got up
#		ra'aytu		'Abdallāhi qāma		# = I saw 'A. get up

'Abdullāhi is thus *mubtada'* and not *fā'il* in the first string.

The formula $(R \rightarrow T_1)\ T_2$ (where the arrow linking T_2 to the ordered pair indicates the *binā'* dependency) constitutes, in fact, a true generative pattern capable of characterising all the kinds of syntactic kernels (we shall see below that there exist other syntactic elements which are external to this kernel). Thus there exists at the level above the lexie a pattern generating items where all the constants of the lower levels are transformed into variables abstraction of the content of the elements and abstraction of inter lexical ordering, with the exception of the ordering governing element of *binā'* $\rightarrow$ subordinate element (without which we would have complete indeterminacy at the formal level). Thus, the formula we have just examined makes it possible to limit considerably the possible syntactic combinations. Let us take the string: # *ḍarab 'Abdullāhi 'Amran* #. It seems to us aberrant to examine all the possible combinations of the three segments (given that we are at a level above that of the combinations of kalim). We are thus forced to take into consideration this important fact noted by the Arabs, namely that a syntactic governing element always subordinates a term, and to and of first and second governed terms in these constituents.

As will be seen below, these syntactic entities may also receive as content from the lower level not only lexies but also signifying segments and even syntactic units from their own level, namely the formula $(R \rightarrow T_1)\ T_2$ itself. All these units form which incorporates them E.g.: See the following table:

Binā'

R_1 = signifying segment	T_1		T_2		Outside binā'
# Kāna	____	Zayd un	qā'im	an	#
# Kun-	__ tu	____	qā'im	an	#
# ḍaraba	____	Zayd un	'Amr	an	#
# ḍarab-	____tu	____	'Amr	an	#
# a'ṭā	____	Zayd un	'Amr	an	ṭawban

R_2 = R, T_1	T_2		T_3
# ḥasiba Ḫālidun	Zayd	an	qā'iman
# ḥasibtu	Zayd	an	qā'iman

R_3 = R, T_1, T_2	T_3	T_4	
# a'lama Ḫālidun Bakran	Zaydan	qā'iman	(11)
# a'lamtu Bakran	Zaydan	qā'iman	

We ought also to point out that such a formula which is dependent on a level of abstraction higher than that of the lexie and the signifying segment is not necessarily bound to a stratum materially higher than that of the other units. In fac, there is syntax even within lexies and even in the kernel of the lexie; thus # *ḍarabtuhū* # (I hit him) is certainly a (verbal) lexie [12] but it is analysable as $(R \rightarrow T_1) = ḍarabtu$ and $T_2 = h\bar{u}$ and constitutes at this level of abstraction a purely syntactic structure [13].

Another important observation: the positional variations in the content of T_1 and T_2 are very wide – and this is a fundamental difference between this level and the lower ones - but they are subject to certain constraints. Thus, when the case inflections appear in the verbal sequence (for phonetic reasons) the sequence is fixed $(R — T_1, T_2)$. On the other hand, the nature of R's content may also fix the positions of T_1 and T_2 this happens with items belonging to the class of "inna". These are invariable exponents with regard to their internal structure consider only the status of governing element (*ġayr* mutamakkin or *mutaṣarrif*); according to the Arab grammarians this implies the fixing of the position of T_1, which must not be separated from R, except if T_2 contains a complement of place or time (*ḍarf*).

On the same level with regard to the role played by the contents of these syntactic entities, but one concerned with case, it is worth noting this observation by *Sībawayh*: "Bear in mind that the *mubtada'* absolutely requires that the / content of the / its own / or constitute a spatial location". (*kitāb*, I, 278). The author means that when in the earlier formula R= ∅, or exponent (formula of *ibtidā'),* the content of T_2 must *necessarily have the same referent* as T_1 or refer to a place where the referent of this content is to be found E.g.: # *'Abdullāh qā'imun* ≠ and ≠ *'Abdullāhi fī dārihī* ≠ ('A. is at home). But this does not apply to the string where R= non-exponent verb.

In the same connection, it will similarly be noted that *Sībawayh* and those who have understood him well have carefully refrained form drawing a parallel between this structure and its interpretation at the communicative level. As will be seen; *Sībawayh* deals with what he calls *isnād*, which is the minimal relation which must hold between a musnad (topic or subject) and a musnad ilayh (comment or predicate) for a *"Kalām mustaġnī* (complete sentence) to be realised. Now the only possibility for drawing a parallel between these two dimensions is at the level of the string R ∅, e — T_1, T_2 (where e = exponent). In fact, it happens that T_1 and T_2 in this structure can be interpreted as "musnad" and "musnad 'ilayh", but this is not always and necessarily a term-by-term correspondence: the mubtada' and mabnī alayh (T_1 and T_2 in the previous formula), formal syntactic elements, are not always interpreted as subject and predicate (a very common case: #*'alā Zaydin daynun* #: Z. has a debt). On the other hand, many lexies can function as sentence (# *darabtu* # *Zaydan* #, # *ṣah* #(hush !), etc.) and on the contrary many binā'-s extend beyond the minimal utterance through the presence of *mafˁūl* which is necessary for there to be *binā'*, but not necessary for there to be *isnād* (relation of subject to predicate).

Nuclear and Peripheral Elements

Until now we have dealt only with the fundamental syntactic structures and variations in their content. But as we have said, there also exists an extranuclear component at this level. It is worth noting that here again there are units which function as determinants in consideration of the content of the kernel (R, T_1, T_2), indeed in more or less the same way as the interlexical determinants, that is the *zawā'id* or expansions which appear- in fixed positions nonetheless - to the right and the left of the kernel of the lexie.

A syntactic determinant is added in the verbal sequence to the kernel, which can be reduced to R→ T_1; what then prevents it from being confused - formally

speaking- with T_2? *Sībawayh* replied to this by observing that only the *ḫabar* (content of T_2 in the *ibtidā'* formula) and the *maf'ūl* (content of T_2 in the formula $(R \rightarrow T_1)$ T_2 can sbstitute themselves, as such, to the mubtada*')* for the first and to the *fā'il in the second.* Thus the subject of the verb can give way to the complement which acquires in this *mawḍi'* a status equivalent to the subject of the verb:

	R	T_1	T_2	
#	ḍaraba	Zaydun	'Amran	# (Z. hit 'Amr)
#	ḍuriba	'Amrun	#	('Amr was hit) [14]

The first of these determinants is the complement of manner (al-ḥāl). This addition as such, can in no way be substituted for the content of T_1, in the way that the maf'ūl can (see *kitāb*, I , 20):

	R		T_1	T_2	Determinent
#	ḍarab -	tu	—	'Abdallāhi	qā'iman # [15]
#	ḍuriba	-	'Abdullāhi	——	qā'iman #

The same is true of the item *"munṭaliqan"* in a string like: # *hāḏā 'Abdullāhi munṭaliqan* # [16]. This element is no longer the ḫabar (content of T_2 in the *ibtidā'* formula) of the string from which this expression derives, namely: # *'Abdullāhi munṭaliqun* # [17] but a lexie with respect to *ḥāl* (the ḫabar is transformed into a determinant and thus transfered outside the kernel). *Sībawayh* claims in this connection that the *maf'ūl* in the previous example (*'Abdallāhi*) and the ḫabar in this example (*'Abdullāhi*) constitute an obstacle between the *ḥāl* (= *qā'iman*, munṭaliqan) and the verb or *"hāḏā"* (see *kitāb* I, 20, 57, 260).

Another determinant, called *"tamyīz"* (specifying) following *Sībawayh*, functions like the *ḥāl*; in both cases there is a *mawḍi'* which separates the position of the governing element and that of the determinant, and prevents the latter from being confused with the element which occupies it. But there is a difference of size between these two determinants: the *tamyīz* can be governed, like the *ḥāl* by a syntactic governing element[18] but it can also be governed by the lexie: noun + item functioning in the *mawḍi'* of the *tanwīn*, or by inflexible terms equivalent to this lexie (e.g, kam = how *many* ?).

NOMINAL LEXIE ⟶		Extralexical tamyīz determinent
Kernel of lexie $\overset{\sim}{0}_{+}\overset{\sim}{1}$	Intralexical determinent $\overset{\sim}{2}$	
ḫātamu	—— n ——	fiḍḍatan
ḫātamu	fiḍḍatin	————
ˁišrū	—— na ——	dirhaman
akṯaru	——hum—	malān
miṯlu	—— hū——	quwwatan
——kam——		kitāban

(19)

As can be seen, there is an intermediate level between that where the determinants of the syntactic kernel are found and that where the expansions or specific determinants of the lexical kernel are found[20].

The other syntactic determinants are the complements called "mafˁūl maˁah" (of accompaniment), the *"mafˁūl* lah (of cause), the *"mafˁūl muṭlaq* (corroborative or specifying process), the *mafˁūl* fīh or ḍarf (of time or place) and the *mustaṯnā* governed in the *naṣb* (marks exception).

we will notice, as the Arab grammarians finaly did, that the determinants which are peripheral to the syntactic kernel are all governed in the naṣb (here mark a). This mark seems to differentiate them formally from the elements which function within the kernel. The *naṣb*, however, is not enough to establish this distinction in certain cases, such as the content of T_2, which can receive the *naṣb* (mafˁūl and ḫabar of kāna). It is in fact, as *Sībawayh* has noted, the potential presence of the content of T_1 which makes this distinction possible. The *naṣb* seems rather to be a distinctive element for any element which is not indeed in a minimal lexie (as a component) and does not have the same reference as the kernel of this lexie (*huwa ġayruhū* walaysa min *ismihī, Kitāb* I, 276).

The Phenomena of Recursiveness

The ism[21] has as its primary positions the *mawḍiˁ-s* where the items are governed. Conversely, an item governed by another can only be an ism in its *aṣl*. This said, there exist some structures where the governed terms are constituted

by verbal lexies. These series then have the status of an ism, and this happens both at the interlexical level and within the lexie, as we shall see. *Sībawayh* claims in this connection that the verb in the imperfect receives the mark u when it occupies one of the ism positions which he lists as follows: 1. position of mubtada' (T_1 in /∅ → T_1 /); 2. that of *mabnī 'alā al-mubtada'* (T_2 in / ∅ → T_1, T_2 /); 3. that of the last *maf'ūl* in strings resulting from the *ibtidā'*. (Ti 2 in / R_{ve} → T_1, T_2 , T_{i2} / 4. that of ḥām (Dh in/ R; T_1,T_2, Dh /); 5. that of the adnominal complement (intralexical position 2⃗); 6. that of the qualifier (intralexical position → 3⃗). Some examples will serve to illustrate these embeddings:

Structures embedded in T_2 and Dh:

R ∅	T_1	T_2			Dh		
		R	T_1	T_2	R	T_1	T_2
# ∅	zaydun	ya-qūlu	∅	ḍāk#			
# sami'-	tu	-	zaydan		ya-qūlu	∅	ḍāka #

Structures embedded in 2⃗ and 3⃗:

0⃗	1⃗	2⃗				3⃗			
		R	T_1	T_2		R			
# yawm	a	ya-qūlu	∅	ḍāka		#			
# ra-ǧul	u	——— n ———				ya-qūlu	∅	ḍāka	#

These embedded syntactic units, however, to be *recognised as such*, that is, functioning as *asmā'* (plural of ism), need to be linked to the central element of the embedding string[22] by means of a duplication of this element in the form of a pronoun (*ḍamīr* called *rābiṭ* = connector). This is the case in: # Zaydun, *ya-qūlu abūhu ḍāka #* [23] where "*abūhu* occupies the *mawḍi'* of ∅ in the example cited above (∅ is in fact the marker of the 3rd person singular).

There also exists another kind of embedding which is done by means of an integration element; this element forms, with the syntactic unit it incorporates, a string which can occur in mawdi'-s reserved for governed terms, namely the *asmā'*. These integrators are, in the *'Arabiyya*, "'*an*" and "*mā*", specific to units with a verbal kernel, and "'*anna*" which integrates only structures whit a nominal kernel.

E.g.

R	T_1 or T_2	T_2 or T_1		
		Int	R	T_1
ḫif —	tu	an	ya-ḫruǧa	Zaydun
'urīdu	$\varnothing$	an	'a ḫruǧa	$\varnothing$
aˁǧaba-	-nī	mā	ṣanaˁ -	-ta

(24)

Given that these embedded strings have the status of an ism , they will therefore have, at the level of the code, the status of a verbal substantive (*maṣdar*: *'an aḫruǧa* ↔ *ḫurūǧī*). Another integrator: "Kay" (which marks the object) can substitute for "'an". The whole being thus nominalised (→ a nominal lexie) can thus be preceded by prepositions. Two of these prepositions by their frequency even lead to the dropping of "'an"(which remains present to all intents and purposes: *muḍmara*).

We can note here that *"'anna"* transforms the syntactic unit into a string having the status of a *maṣdar*.

Another kind of integrator is the relative *al-ladī* and its derivatives, as well as *"man"*, *"mā"*, and *"ayy"*. It has been noticed that it can be found in all the positions of ism (because the string it forms with the series it integrates does not have the status of a *maṣdar*).

These embedded syntactic units constitute an indivisible whole with the element which incorporates them, and behave like nominal lexies. The same is true, moreover, of the units embedded in the six positions of the verb. They cannot be anteposed or have any effects on the items which precede them. Nor is it possible to insert into them an element external to them, whether preceding or following.

A recursion by duplication of items rather than by embedded of syntactic units (called *iṭalā*) consists either in repeating the item contained in the same *mawḍiˁ* (multiplication of its content[25]) or in diving the *mawḍiˁ* itself into two. In the first case we have what is called *"išrāk"* (later *'aṭf nasaq*) which corresponds to coordination in the European languages, or "taˁaddud", which is a multiplication by simple juxtaposition, and which can occur only in the six verbal positions listed above. In the second case we have a redundancy which has to mitigate the unfavorable conditions of commuication and which is realised

in the form of *"tawkīd"* or reinforcement through repetition or through the use of certain specific items or of *"bayān"* (*'atf bayān* = clarification). The "badal" which results is also a restatement or supplementary information (~ appositive).

Super-Government

There exist some function-words which can occur at the beginning of sequences having the structure R, T_i, D, which suggests the existence of *mawḍiʿ-s* outside this structure. On the other hand, it has been noticed that the space occupied by an item may correspond to series of n *mawḍiʿ-s*. This is the case with the *interrogative "hal"* which indeed seems to cover more than one *mawḍiʿ*, since it cannot be substituted for its homologue *"'a"* in expressions like *"'a lam taḥruǧ"*, *'a sa-yaḥruǧu"*, *"a in ḥaraǧtu 'āqabtanī"*,[26]. Moreover, "in" is itself not substituable for " sa -", which suggests that "hal " covers three *mawḍiʿ-s*, its own and those of these two particles. The same is true of the negative exponent *"mā"*. It cannot be substituted for any of the elements which alternate alone with "a", "lam" or "sa" and "in". Now "lam" and "sa-" are *elements internal to the verbal lexie* (hence always to R). The same is true of "inna", this element cannot substitute for the interrogative *"'a"* alone, but for the series which includes. R itself {'a+lam/ in…+R} and which is equivalent to {hal +R}. The Arab grammarians also observed that the order of the two *mawḍiʿ-s* (which have to contain a on the hand the other particles on the other) -which we shall label in our metalanguage α and β - is always fixed, that no item can be inserted into them with respect to R, T, or D, and that the α position constitutes, besides, *the limit* beyond which no item located in the preceding or following sequence can be moved by ante- or postposition. This means that no item located in front can be governed by α governing syntactic element introduced by α. Hence the name "*ḥurūf al- ibtidā'* or *ḥurūf mubtada'a* which is given to the exponents appearing in α, where *ibtidā'* has the meaning "position of complete non-dependence on *what precedes*, and which can be replaced by the term ṣadr (*al kalām*)= absolutely initial, to avoid the ambiguity of the first term (which also designates the position of R $\varnothing$)[27].

As for the β position, it is the *mawḍiʿ* where there is alternation of exponents which introduce two syntactic units, between which is established a dependency called *"taʿlīq"*, which *al- Ḫalīl* compared to that established between (R→T_1) and T_2. The *"taʿlīq* is thus a kind of *binā'*, but tocated at a higter level. In both cases we have: 1, an *ibtidā'* or position of grammatical non-dependence of (R, T_1) and position of absolute non-dependence (= to the second power (of β, R); Z, an obligatory postposition of the first governed term

in relation to the exponent[(28)]. Symbolising the two series by $\mathfrak{C}_1$, $\mathfrak{C}_2$ and $(\alpha, \beta,)$ by R, we get the formula:

R	$\mathfrak{C}_1$	$\mathfrak{C}_2$
α, β	$([R \rightarrow T_1]\, Ti_{>1} \pm D)_1$	$([R \rightarrow T_1]\, T_{i>1} \pm D)_2$

The Arab grammarians' conception of more abstract initial structuring position makes it possible to transcend the fundamental but partial formulas of the level where government of R occurs by showing this more abstract structuring effect, which is the super-government of R and the indirect *super-government* of the elements governed by R.

The *mawḍiʿ* $\mathfrak{R}$ is interpreted, at the level of the code, as an *illocutionary* position. But the Arab grammarians here esthablished differences based on the primitive (*aṣl*) or secondery (*farʿ*) nature of the *denotations*. Thus certain sememes denoted by these exponents are more primitive then others; the most primitive are precisely those which are marked excluvisely in α, namely the assertive (*ḫabar*) and the simple interrogative which derives from it. The corresponding marks are $\varnothing$ and 'a. These two have, in turn, two degrees of redundancy: insistent assertion and interrogation marked respectively by " *lā* -"or " *inna*" (or the two) and by " *hal*". Opposed to the semantic feature of assertion as derived features, are the simple and complex interrogative and all the other semantic features belonging to the category called *ṭalab* (postulative): order, wish, expectation, etc. However, this is not simple; "hal", in fact extends beyond α and covers β and the intralexical *mawḍiʿ* Ī specific to the verbal lexie, in other words a part of R, since it not compatible with the elements that appear in this position (lam, lan, etc.). The same is true of "mā" which however does not cover α since it can be preceded by "'a ". There are also items which extend far beyond their exponent *mawḍiʿ* to include *mawḍiʿ*-s which are far from the term governed in T_1 or T_2, that is an ism. Such are the markers of the complex interrogative: "man" = who, "mā" = what "'ayna" = where, "matā"= when, etc., for they include, as well as the interrogative feature, that of the object ism or ism-ḍarf, and must for this reason have a function at the level of case: subject, object complement, complement of place, etc. In the part of R which is bounded by the initial part of the verbal lexie (which can occur there) function exponents which are similarly ranked into *aṣl* and *furū'*. Thus the most primitive semantic features marked (necessarily by $\varnothing$) in this position are

affirmation (V. negation) and temporal indeterminacy (V. present, past, futrure) . Finally in the β position we have the semantic features ″conditional″ and ″hypothetical″ (V. the primitive features: non-conditional or non-hypothetical), which imply, as we have seen, a second string, which is at the level of code, its logical consequence. The primitive feature found immediately after ∅ is marked by ″in″ = if and ″law″ (if). The derived features are created from it, as for the postulative, by the incorporation of additional features (the same exponents that serve as interrogatives can function as exponents here ″man″ = whoever, ayna (*mā*) = wherever, etc.). See p. 43 a sample of the matrix which incorporates these denoters.

The notion of the overlapping or blocking of several mawḍi'-s has considerable explanative value since it makes it possible to account first of all for the correspondance betweeen sequences (=Harris's transformations) which are not clear or dot not appear at all the verbal string, and secondly for the numerous fixed positions. On the other hand, it is absolutely necessary - as a component at the phrase level - for the notion of a ranking of the other axis into primitive and secondary items. The Arab grammarians used these two notions (which cannot be separated) to explain the fact that certain items can modify the inflectional marker of the term they govern, while others cannot. In the framework we have just described, to have such an effect an element must, according to the rules, *(taqdīran)* occupy a *mawḍi'* which is syntagmatically anterior (R in relation to T1, for example) and paradigmatically posterior (second rather than first. Cf. the primitive and secondary features above). Thus ″lam″ and ″ lan ″ as exponents of the verb have an effect on its inflection because they both contain *two secondary* features: negation + past, negation + future. *Al-Rumānī explains: "Given that ″safwa″ modifie*s the verb from only one point of view, its presence alone suffices; as for the other elements, given that they modify the verb from two points of view, their introduction is not enough - a *supplementary marker is required for* this feature which is likewise supplementary ″ (*Šarḥ*, III , 91 V.).

5. The lexical level

We shall say a few words about this level since we have already talked about the pattern and root of the kalima.

For *al-Ḫalīl* evreything begins, at this level, by the possible combinations of *ḥurūf* -phonemes: "The language of the Arabs″, he tells us, ″ is entirely based on four kinds of combinations: binary, triliteral, quadriliteral and quinquiliteral… The binary kalima has *two* possibilities for variation Q/C and D/Q. The triliteral has six and is called ″*masdūsa*″ (= hexatropic group) DRB/BRD/BDR/RDB/

RBD. The quadriliteral has twenty-four possible variations. In fact, the product of the number of *ḥurūf* contained in this kalima and the number of possibilities for variation. In fact, the product of the number of its *ḥurūf* and that of the variations of the quadriliteral, 5x24, is indeed 120″*(kitāb* al-'Ayn, I, 66). *al-Ḫalīl* not only establishes the formula for calculating all the possibilities for permutation between the kalim constituents, or *factorial* of n *ḥurūf*, but also had the good idea of representing- whit the aim of operating on sings - all these permutations in a diagram, namely a double-sided circle (cycle). The *kitāb al-'Ayn* conceived by *al-Ḫalīl* was thus to present an exhaustive list of the roots involved in these combinations and to give an interpretation according to the language's code, specifying the existential status of each combination (muhmal = non- existent/ musta'mal = existing in actual use) and the list of all the kalim deriving from it.

The combinatory rules for roots are completed and considerably limited- by those for the patterns of kalim, which are by this fact truly fixed models. More than 1200 have been counted, but they do not exceed 300 if the hapax are exluded. The qisma, or combinations for the triliteral ism[29] involves 12 patterns; the problem amounts to linking each state of C1 (f) with a state of C2 ('), which is equivalent to obtaining the cartesian product of $\{f\} = \{a, i, u\}$ by $\{'\} = \{a, i, u, \varnothing\}$.

It will be noted that the *'Arabiyya* has retained only 10 patterns specific to the ism for the primitive triliteral, and 5 and 4 patterns for the two other kinds of ism.

6. Distortions of Discourse and Causal Explanation.

As we saw earlier, the *waḍ'*, code and structure in actual usage or *istiʿmāl*, suffer distortions which, when they become the predominant usage, need to be explained. The first grammarians talked of "*ʿilla*"; this is the *cause* of a deviation in the behaviour of an item when compared to the behaviour of the structured whole to which it belongs (or the pattern which characterises it). In other words it is a factor of disorder or imbalance for a *qiyās-bāb*. This *ʿilla* takes as its starting point the spontaneous manipulations or lapses of speakers in realising these *qiyās*, lapses themselves due to the fact that they are subject to other constraints, other rules, physiological, psychological or social. The situations where *ʿilla*-distortions are frequent are those *Sībawayh* calls "*saʿat al-kalām*" or free use of poetic code or established usage. In the first case the reason invoked is mainly "*istiḫfāf or ḫiffa*" (as opposed to *istiṯqāl, ṯiqal*); this according to *Ibn Ǧinnī*, is

" the tendency to seek out what is not felt to be costly and to avoid what is not" (*Ḥaṣā'iṣ*, I, 162-163. This is for him "the principle of principles"). This tendency towards economy is of course at work at the level of expression or the type of oral presentation of the Koran which is called "*ḥadr*"(as opposed to *tartīl*), characterised by a speed of delivery which leads to a considerable reduction in articulatory activity. In syntax, ḥiffa like *'illa* appears in ellipses. But, it may happen that an economical realisation which existed only at this level becomes generalised and even comes to constitute the only accepted usage. It is on these grounds that attempts are often made to explain *(ta'līl)* forms which form *qiyās* but which are deviant for more primitive *qiyās* such as "*bāˈa*", "*qāma*", whose *bāb* is almost homogeneous, but which ought to have been ₊ *bayaˈa* and ₊ qawama which are predicted by the original system (synchronically speaking). As for the idiomatic expressions called *"'amṯāl"*[30], their very high frequency of occurrence is invoked. They are ranked with poetry, which allows considerable distortions (cf. poetic licence) because it is intended to be circulated. The principle of economy is opposed to a contrary principle which is that of "amn al-labs" or avoidance of ambiguity.

ℜ		ℭ₁				T₁	T₂	ℭ₂
		R						R, T₁, T₂ ± D
α	β	V1	V2	V3	N. régis	T₁	T₂	
∅	——	∅	——	——	∅	Zaydun	munṭaliqun	..#
∅	——	————————laysa————————				Zaydun	munṭaliqān	fa-’anṭaliqa..#
∅	——	————————mā————————				Zaydun	munṭaliqān	..# "(30)
’a	——	∅	——	——	∅	Zaydun	munṭaliqun	..#
’a	——	mā	——	——	∅	Zaydun	munṭaliqun	..#
’a	——	————————laysa————————				Zaydun	munṭaliqān	..≠
		hal			∅	Zaydun	munṭaliqun	fa-’anṭaliqa..#
’a	——	∅	——	∅	ḫaraga	Zaydun	..#	
		hal			ḍaraba	Zaydun	‘Amran	..#
		’inna				Zaydun	munṭaliqun	..#
		layta				Zaydun	ya-nṭaliqu	..#
∅		lā				kitāba		fa-aqra’a..#
∅	——	∅	——	∅	kāna	Zaydun	munṭaliqan	..#
∅	’in	——		——	kāna	Zaydun	munṭaliqan	’intalaqtu..#
’a	’in	lam		——	ya-qum	Zaydun		‘āqabtuhū..#
∅	law				ǧa’a	Zaydan		la-’akramtuhū..#
		————————halla————————			ǧa’a	Zaydun		fa-’ukrimahū..#
la-	in	——	——	——	ta-ḍrib	∅	‘Amran	’aḍribka..#
——	——	——	’iḍrib	——		∅	‘Amran	..#
——	——	lā	ta-ḍrib	———————		Æ	‘Amran	..#
——	la-	’aḫruǧanna		——		∅	..#	
		————————’iḍribanna————————				∅	‘Amran	..#
∅	law	lā	——	——	∅	Zaydun	(’iḍmār)	la-halaktu..#
…	...	...	...	...	...	...	…	…

(31)

This is used to explain why certain omissions are not made in certain cases: such is the case with the verbal lexie in *"li-yuḍrab Zayd!"* (may Zayd be hit), for here there is no way of distinguishing the second person from the third.

Some other principles *('uṣūl)* are also invoked, first, the structural levelling of a *bāb* which has undergone distortion at the phonetic (or other) level, to prevent the *bāb* from being made of disparate elements. This is in fact *analogical rebuilding*. Thus in verbs of the class of *"ya-zid"*, the first radical which is a -w- is dropped in the third person, because it is between *a y* and *ī* (phonetic *'illa*),

but this 'illa has spread to all persons.

Two other principles concern errors of manipulation (*ġalaṭ* or tawahhum) and the interference of regional variants. The first case concerns the "mistakes" made by native speakers in handling, the *qiyās*. This "error" can be the act of one individual (accidental or not), and is then called *ḫaṭā'* (inaccuracy); but it may be indespread (it is then called *"luġa"*) or even the only form accepted. It will thus be noted that the *"ġalaṭ"* is not in itself an error, but a usage which originates from a mistake. Usually it involves an alignment which is not necessarily implied by any isoschemism. For example, *"muṣība"* was felt to exhibit the pattern *"faʿīla"*, so its plural has been formed as *"maṣā'ib"* on the basis of this singular.

7. Formalisation of Causality and Free Variation of the *Waḍ'*

The Concept of Causal *Taqdīr*

The Arab grammarians also tried to formalise their explanations of distortions by trying to incorporate all these intuitive causes into a system of formal relations which could in turn be incorporated into the system of the *qiyās*. In fact, the taʿlīl or causal explanation in its formal aspect consists in opposing the form of an item which appears deviant in respect of its *bāb* to the form required by the *qiyās* of this *bāb*; that is, between an actual, observable form *(ḏāhir al-lafḏ)* and a potential form which is that which ought to appear. The actual item having another form in reality, we shall then talk of simulation and potentiality *(tamṯīl and taqdīr)*. The formal explanation resides precisely in the reduction of these two entities to a single one. We consider, in fact, that there must exist between them relations of the same kind as those which combine, in the same pattern, items dependent on the same structured whole, namely the *reversible transformations* which are everywhere at work in the constitution and functioning of generative patterns. It is *on the basis of these transformations* that a pattern is established which is also a *qiyās*, but a pattern is established which is also a *qiyās*, but a qiyās which overlaps above both the form required by the intragrammatical qiyās and the form observed in use. The extragrammatical, or more precisely the 'illa and its effect, are thus incorporated into the system of generative patterns of the language.

The *taqdīr* thus presupposes two operations which are not simple: first, the reconstitution of the *aṣl* muqaddar (the potential form), and second, the discovery and at the same time the elaboration of the transformations or implicational relations which link the potential *aṣl* to the observed form. This amounts to

setting up an *aṣl* and then comparing it to the reality in order to reduce one to the other.

The potential form is obtained thus: the actual items A, B, C necessarily appear in the *mawḍiʿ-s* specific to them. Now, these *mawḍiʿ-s* form a pattern which generates the forms that A.B.C must have. Thus the items kataba, fahima, karuma (write, understand, be generous) have the pattern faˑxla (where x - a, i, u), which should in principle also characterise *qāma*, *ḫāfa*, *ṭāla* (get up, be afraid, be long). Now it is not possible to pair them directly. We therefore apply the combinatory rules for this kind of them, to obtain the primitive potential forms *qawama, * ḫawifa, and, *ṭawula. We then induce the transformation - relation (which can moreover be a series often very complex transformation) which they have in the potential form, namely: second radical w/y → long segment with the quality a. An example of a long and arduous series of transformations is that concerned with the plural form *"ḫaṭāyā"* which has as *aṣl*ₓ *"ḫaṭāyī *"* (*Ibn Ǧinnī* identifies six martaba-s or transformational stages here. See his *Ḫaṣā'iṣ*, III, 5) [32]

In fact, it is a matter of real sign calculation since it is possible that one or more of these forms resulting from transformations have never had any kind of existence and may even be unpronounceable. This is the case with *maqūwl*, which originates from the *aṣl* *maqwūl, and which, through a series of transformations, must result in the observed form *"maqūl"*.

There are however, good reasons for clearly differentiating two kinds of *taqdīr* and consequently two kinds of ''transformations examined above: the derived forms on *furūʿ* derive from an aṣl by a structural change and/or a phonetic or semantic addition to this *aṣl*. The transformations of this type (comparable to those of Harris) thus necessarily of *taqdīr* (here causal), the transformations simply have to coherently result in the observed form (the chomskyan type). In these conditions, the problem of deciding where the semantic and/or phonetic information necessary for interpretation is located *does not arise* in this framework of the *naḥw*.

When there is a deviation of usage at the level of the *waḍʿ* -code, we do not talk of *'illa* but of *'ittisāʿ* or free extension and variation. It is then no longer the naḥw which is concerned, but the *balāġa*, which is the study of individual use of the language. This is concerned with these derivations, but generally with the exploitation of the expressive potential of grammatical material (*maʿānī al-naḥw*). This involves explaining, as *'Abd al-Qāhir al-Ǧurǧānī* observes, the

choice of a particular expression among all those allowed by the *naḥw* (*Dalā'il*, 67-68) The Arab linguistics also took account of primitive semantic features. Considered as such are the meanings which are marked *by a signifier* (full or zero), in contrast to those which are obtained by extragrammatical indices, namely the *dalālat al-maʿnā* (presupposition or implication). It is just this set of primitive meanings which constitutes the *waḍ'* -code. The study of the phenomena of *ittisāʿ*, which are a regulated modulation of this code, will thus consist, for the linguist, in thoroughly examining this modulation by: 1. varying the respective contents of the positions: α, β and D and/or the expansions at the lexie level together with the order of the items, and 2. holding the content of the syntactic kernel constant, in order to obtain sets of utterances or utte-elements. He must also draw up a list of the relations established in actual utterances between these variations and the situations where they occur. The formal integration of these primitive relations will consist, as for the causal *taqdīr*, in linking the situations and the observed forms by means of a potential element, namely the *mawḍiʿ-s* where these actual forms occur and the generative pattern which follows from them. We then determine the aṣl in this pattern by reducing to zero all the extra-nuclear content as well as the observed order. The *aṣl* is then linked to its *furūʿ* (which are here the expressive variants) by transformations; the latter in turn determine a wider code which is the expressive code of the language at a particular moment in its history.

8. The Axiomatisation of the *Naḥw*

The formal system established by the naḥw is defined by four sets: the *awḍāʿ*, the *maqāyīs* and the *uṣūl* and *furūʿ* implied by the maqayīs. The *'awḍāʿ* (plural of *waḍ'*) are simply the terms or alfāḍ: 1. the *awḍāʿ al-naḥw* which constitute the technical vocabulary or metalanguage of the *'ilm al-'Arabiyya* (subdivided into *amṭila* or symbols conventionally chosen, e.g. fiʿl), and *alqāb*: ism, fiʿl, ḥarf, rafʿ, naṣb, etc.) and 2. The *awḍāʿ al-luġa* which are the specific items of the language. The *maqāyīs* constitute sets of combinatory rules. From the logical point of view, these are true deduction schemas. The series of *awḍāʿ* which are implied by these *maqāyīs* can be either uṣul (plural of *aṣl*), that is, series from which other series can be deduced, and among which there are uṣūl which cannot be deduced, from any other aṣl, or *furūʿ* which are the series which are deduced from the *uṣūl*. The set of operations needed

to transform an *aṣl* into its different *furūʿ*, through a number of stages called *marātib*, constitutes an 'amal or calculation.

Strictly speaking, the formal system of the *naḥw* is only represented (at each level) by the finite series of intragrammatical transformations or series of marātib (the latter are illustrated by the sign calculations called *al-masā'il* or *masā 'il al-tamrīn*). But it is felt that the system as such could not be represented or even constructed without calling upon something external to it, namely a language allowing the linguist to discuss it and elaborate on it.

In the calculations contained in the attempts at explanation, and in the *masā'il*, everything is *conventional*. The best proof of this lies in the fact that the aṣl and the *far'* can be completely imaginary and have no counterpart in reality. The freedom thus gained by the linguist leads him to try all kinds of pattern and at this point to pay no attention to whether they exist or not. The formalised qiyās is thus completely cut off from intuition. It must even be cut off from the laws of reason, the only law of reason maintained there is coherence and the principle of non-contradiction (without which no rational knowledge is possible). This is what *Ibn Ǧinnī* tries to show in a chapter entitled "on the *mustaḥīl*" (= the impossible or the absurd), in which he shows that the qiyās can relate even to premises which are materially false, since what is retained is the hypothesis and the set of rules which are set up.

Arabic Phonetics

Phonetics occupies a prime position for the ancient Arab linguists, particularly *al- Ḫalīl* and his pupil *Sībawayh*. In their discussions, the question almost always arises of the influence of accidents in the stream of speech on the form of lexical units. In fact, phonetics intervenes at all the levels of description and explanation. In *Sībawayh's Kitāb* the description of the sound system of the *'Arabiyy*a comes right at the end. Moreover, it forms a kind of introduction to his discussion concerning one of the important phenomena of combinatory phonetics which affects the system of the 'Arabiyya, namely the *idġām* or geminative contraction [33]

      - The level of ḥurūf
      - The concept of ḥarf

The *ḥarf*, the kalima - and the *kalām* as a unit of communication - are the segmental units of the 'Arabiyya (the last two of course comprising an underlying structure which is not identifiable with the segment as such).The kalima is the unit which appears in one of the positions contained in the lexie pattern. It is a meaningful segment whose minimal nature follows from the pattern in question and not from its content (which is minimal only in relation to it). The *kalām* is formed not only of kalim but of syntactic units which may contain kalim.

The kalima is analysable into ḥurūf (plural of *ḥarf*). These are then segments of sound, but without meaning as such. On the other hand, the kalima can be formed of only one ḥarf (the final-t in the 3ʳᵈ person feminine singular of the verb, for example). But how can we get to the *ḥarf* ? In other words, by what objective criteria can it be delimited in space. On what basis did the Arabs finally manage to distinguish the ḥurūf and establish their system? Before answering these questions, it is first necessary to examine their very original conception of an articulatory dynamics based on the concepts of ḥaraka and *sukūn*.

The Concepts of *Ḥaraka* and *Sukūn*

"We cannot", declares *al-Rummānī,* "pronounce a *ḥarf* in isolation, but only concatenated *(yū-ṣalu)* with another *ḥarf* for concatenatory linking is in fact the aṣl of *kalām*" (*Šarḥ*. V, f.23). Thus, the *ḥarf* has material existence only in a chain of *ḥurūf*; it can be realised only within a sequence. We can conclude from this that it constitutes a *sequential element*, a unit transitional in nature. This observation is very important since it constitutes the axiom on which all segmentation of the *kalām* is based.

The verbal non-autonomy of the *ḥarf* is moreover based on the still more primitive concept of the articulatory *"idrāǧ"* or *"waṣl"*, which is a sort of dynamic insertion. As we have said elsewhere, "the *ḥarf* is *embedded* among *ḥurūf* and develops in time through an aero-organic movement of sound, with which it is mingled and in which it is born and dies as a *phase* or sequential transition, leaving its place to another *ḥarf* without a break".

On the *"idrāǧ"* (cf. the *"darǧ al-kalām"* = the chain of words) are based the concepts of ḥaraka and *sukūn*. The ḥaraka can be defined as the aerial, organic and usually acoustic movement or impulse which is required for a *ḥarf* to be produced in a continuum of sound. "The *ḥaraka* makes it possible for the *ḥarf* to be realised" claims *al-Rummānī*(*šarḥ* , V, p 15R); "the *ḥarf* followed by a *ḥaraka* implies the passage of this ḥarf *towards* another *ḥarf* (Ibid., 22 V). The *ḥaraka* is thus, in this conception, an *aero-organic impulse which makes possible the sequential articulation and consequently a transition to another ḥarf.*

The passage to another ḥarf implies *a change of maḫraǧ* or place of articulation. There is then, as Saussure describes it when talking of what he calls "explosion", a relaxation of the organs (or an opening movement).

The *sukūn* is the state which contrasts with that of the *ḥarf mutaḥarrik* (*ḥarf* + *ḥaraka*): the ḥarf is produced with a "closing" movement of the organs. This constitutes a halt in the movement of the *ḥaraka*: the latter thus necessarily

precedes the *ḥarf sākin* in order that this halt can be produced by means of an "implosive" obstruction of the vocal tract. The ḥarf sākin can thus be described as implosive.

However, the *ḥaraka* is not the same as the vowel sound (which is always produced by an opening action of the organs) which can accompany it. This is proved by the fact that the *ḥaraka* can be the subject of an *"Iḫtilās"*, which is a rapid glide from one consonantal place of articulation to another in which this vowel sound is thereby made quite indistinct or even non-existent. *Ibn Ǧinnī* gives an example: *"Šah-rramaḍān"* where the case vowel which should follow the first r is completely non-existent. What remains is only the aero-organic impulse which underlies it (the two r's do not constitute a geminate since what precedes is implosive)" (see *Sirr al-Ṣināʿa*, I, 64-65). In this framework of *ḥaraka/sukūn*, al-Mubarrad (*Ibn al-Sarraǧ's* teacher) states: "we can only begin (a sequence) with a *ḥarf mutaḥarrik*, just as we can only make a pause with a *ḥarf* sākin. If we asked someone to pronounce an /isolated/ *ḥarf*, we would have asked for something impossible… for/ that/ would be the same as asking him to produce a *ḥarf* mutaḥarrik and *sākin* at the same time!" (*Muqtaḍab*, I, 36). Similary, two *ḥarf-s sākin* cannot come together except at a pause (which brings in a supporting sound *"ṣuwayt"*) and when the first *ḥarf* is a long segment (ḥarf madd) (see *Ḫaṣāʾiṣ*, II, 328). We have published a study in which we propose to label these concepts kineme and zero kineme (state of *taḥarruk* = kinesic and its opposite akinesic). See *al-Lisāniyyāt*, Algiers, 1971, Vol. 1, pp. 63-84).

The Delimitation and Inventory of *Ḥurūf*

From the purely articulatory point of view, "the ḥarf constitutes the ultimate point where the cutting of the sound / in the organs of speech/ takes place" (*Ibn Ǧinnī*, Sirr *al-Ṣināʿa*, I, 16). This "cutting" consists in an obstruction of the vibrating air coming from the larynx, an obstruction which "trims" or cuts out the amorphous sound to give it the acoustic form characteristic of the *ḥurūf*. The place or position of the organs producing the *ḥarf* is called "maḫraǧ" and also *"maqṭaʿ"* (and "madraǧ" by *al- Ḫalīl*). This term also functions as a verbal substantive and applies to *the concrete realisation of the ḥurūf*, hence the meaning of a variant realisation it also possesses (cf. Ǧāḥiḍ, *Bayān*, I, 34).

It is obviously on the basis of the kalima, which we already know how to demarcate (thanks to the lexie pattern) that we reach the *ḥurūf*. The definition and demarcation are here, too, purely formal, since the starting point is again the lafḍ. In fact, linguists consider as *ḥarf any sound contained in a kalima which cannot*

be realised in isolation *without a prothetic* element, if it is *sākin*, and *without a supporting sound* if it is mutaḥarrik. Such is the case with the realisation in isolation of the "k" of "laka" and the "b" of "iḍrib": # kan # and # ib # (Kitāb II, 62). The prothetic element is that which occurs when a kalima begins with an akinesised *ḥarf (sākin)* such as "*(i)-mru'un*" and "*(i) ṯnānī*" (glottal stop + i); the supporting sound is generally the akinesised glottal -h which occurs in speech at a pause after an obligatorily kinesic ḥarf. These elements are thus the marks of the minimal articulation. All the sounds possible within the *kalim* which are substitutable for these minimal units within the *kalim*, without causing them to lose the status of *kalima*, will also be considered as *ḥuruf (ta-qaˁu mawqiˁa …).* On this basis the Arab linguists were able to draw up the inventory of all the sounds belonging to the *'Arabiyya* , in other words the set of kinds of variant heard from the lips of native speakers.

The Formal Characterisation of *Ḥurūf*

al- Ḥalīl and *Sībawayh* were able to demonstrate the full set of distinctive features of the *ḥurūf* - in the form of realisation schemas- by studying *very closely co-articulation* (and thereby observing the behaviour of the organs of speech), namely the phenomena of ibdāl and idġām (mutation and geminative contraction), of *ḥaḏf* (elision), of qalb /*makānī*/ (metathesis), etc., in the framework of *taqrīb* or *idnā'* (assimilation) or its converse *(ibdāl li-'iḥtilāf l-ḥarfayn)*[34]. On the other hand, the axis along which the places of articulation of the organs of speech are laid out is compared to a *series* of *mawḍiˁ-s*; each *mawḍiˁ*, thus becomes a place of alternation for the features which serve to distinguish the *ḥurūf* depending on this *mawḍiˁ*. The set of these *mawḍiˁ-s*, thus seriated, forms a true matrix, based this time on the articulatory substratum. The features have here similarly been thought of in terms of *ziyāda* or expansion, the distinctions established in each column are not achieved by a simple opposition but by successive additions, thus by transformations. This *ziyāda* is here called *faḍīla*.

In the matrix, we have first of all the conventionally least marked sounds, namely vowels, those which accompany the *ḥaraka* and those which prolong these sounds, which are called *ḥurūf al-madd*. These "chronemes" are considered as *ḥurūf* because, in the morphophonology of the *'Arabiyya*, they can be substituted for other *ḥurūf*. The Arab linguists in effect consider length as segmental rather than suprasegmental. The sounds of the *ḥarakāt* are, from the acoustic point of view, *ḥurūf* (since they occupy a space) but deficient (*nāqiṣa, ḥurūf ṣaġīra*)

since they do not substitute for other segments. These *ḥurūf* have as *fadīla* the fact that they have an articulation which *opens* to the passage of air (*muttassiʿa li-hawā'i al-ṣawt. kitāb,* II, 265). This feature, called *līn*, contrasts with *ġalaḍ* (soft, aerial/ hard, solid). At the opposite end of the axis of transformations are the *ḥurūf al-šadīda* which have a maximal *ǧalaḍ*; there is total but momentary obstruction of the vocal tract, hence a greater expenditure of energy. These are the stop consonants. Between these two poles (*līn*→ maximal *ǧalaḍ*) are found three classes of *ḥurūf*: 1. the *layyina* consonants[35] w and y. The air passes as in the *ḥurūf al-madd*, but they have a closing action. Immediately afterwards: 2. the ḥurūf al-riḫwa; here the obstruction is partial, which allows the passage of vibrating air (*yaǧrī fīhā el-ṣawt* → fricatives). 3. the *ḥurūf bayna bayn*, intermediate between the previous category and the stops; their realisation involves both a closure *(luzūm al mawḍiʿ)* and an opening elsewhere. This is the case with the nasals, l and r and the pharyngealʿ. Within these classes other distinctions can be drawn. Thus the *ḥurūf ǧawāmid* can be *mutbaqa* or otherwise (pharyngealised, emphatic, dark), and *maġhūra* or *mahmūsa*, which corresponds more or less to the distinction between voiced and voiceless (see the generative matrix of *ḥurūf)*. See the next page.

The realisation schemas of the *'Arabiyya* called *uṣūl* were those used by the majority of speakers of *faṣīḥ-s*, but other schemas were also noted in certain regions or certain tribes; these constituted *luġāt* which is a particularly unusual and isolated variation. It is dependent rather on purely individual deviations (see *Kitāb*, II, 404-406)

GENERATIVE MATRIX OF ḤURŪF

Ṣifāt / manners (rows) × Maḥāriǧ places (columns)

Place groups (columns): Pharynged Ḥalqī faucal zone = C1–C3 · Uvular lahawī = C4 · Velar 1 min aqṣā al-ḥanak = C5 · palatal = Šaǧrī = C6–C7 · Alveo-dental = ale naṭʿī = C8–C11 · Inter-dental laṭawi = C12 · Labial = šafawī = C13–C14

Manners / Ṣifāt	Aqṣā al-halq (post-pharynx)	Awsaṭuhū (medio-pharynx)	Adnāhū (oro-pharynx)	al-Lahāt (uvula)	Sifāq al-Šaǧr (velum)	Min al-adrās al-ʾulyā ḥattāl-ṭanāyā	Wasaṭ al-ḥanak (medio-palate)	al-Adrās ḥatta l-ṭanāyā (premolars to incisors)	Fuwayqa al-ṭanāyā	Uṣūl al-ṭanāyā (alveolar)	Fawqa al-ṭanāyā	Aṭrāf al-ṭanāyā (extremity of incisors)	+ šafa suflā (inferior lip)	Bayna al-šafatayn bilabilal (the two lips)
Novels Ṣawt — al-ḥaraka — Deficient	´ = a						ِ = i							ُ = u
Novels Ṣawt — al-madd — Full	∟ = ā						ي = ī							ـو = ū
semi-vowels = layyin							ي = y							و = w
Fricatives rihw — Non Pharyngalised = munfarid — Voiceless	هـ = h	ح = ḥ	خ = ḫ				شّ = š			س = s		ث = ṯ	ف = f	
Fricatives rihw — munfarid — Voiced			غ = ġ							ز = z		ذ = ḏ		
Fricatives rihw — Pharyngalised = muṭbaq — Voiceless										ص = ṣ		ظ = ẓ		
Fricatives rihw — muṭbaq — Voiced						ض = ḍ								
Semi-fricatives = baynahumā — lateral = munḥarif								ل = l						
Trill = mukarrar									ر = r					
Nasal = ʾaġann										ن = n				م = m
Special property of ʿayn		ع = ʿ												
Stops = šadīd — Non Pharyngalised — Voiceless					ك = k						ت = t			
Stops = šadīd — Voiced	ء = ʾ						ج = ǧ				د = d			ب = b
Stops = šadīd — Pharyngalised lised = muṭbaq — Voiceless											ط = ḍ			
Stops = šadīd — Voiced				ق = q										

Third-level articulatory descriptions: Mā ya-lī l-ṣadr (C1) · Ḥalq = pharynx (C2–C3) · + aqṣā al-lisān (root of tngue) (C4) · + ḍahruhū (dorsum) (C5) · + ḥāfatuhū (latéral p.) (C6) · + ḍahruhū (dorsum) (C7) · + hāfatuhū (lateral .p) (C8) · adḥal fi-l-lisān (prc. dorsal p.) (C9) · + ṭarafuhū (ḍawlaq = apex) (C10–C11) · + šafa suflā (inferior lip) (C13) · al-šafat‾ān (the two lips) (C14)

Conclusion

Two basic differences distinguish this kind of approach from that of post-Saussurian structuralism. In the first place, the latter generally operates by simple abstraction (the characteristic of all radical empiricism). In fact, everything in it is based on *intensive* and *exclusive* abstractive processes: the individuals in classes which themselves are related only by inclusion, *intersection* or exclusion. To move from one element to another, one always feels obliged to pass through the class containing them, unlike in the case of *qiyās*, which depend on a *constructive* and extensive abstraction: the elements belonging to two *bāb-s* are directly related. This then reveals a more abstract structure which incorporates them and extends beyond them.

The other difference lies in the fact that we do not reduce all science to the science of phenomena; we recognise there also a "science of action" which is not necessarily identical to the prescriptive disciplines. In fact, we can, see in the qiyās mustamirr (in the *ḥadd*) not only the simple description of a relation holding between two classes of phenomena, but also the *constant* which makes it possible to *predict* the development of new acts of discourse and consequently to regulate language behaviour according to the requirements of the language system.

On the other hand, no linguistic theory based on pure structuralism or generative grammar has the syntheticist viewpoint on the grammatical qiyās, which regards the syntagmatic and pradigmatic axes as the *integrated and dynamic* components of the same unit, namely the matrix of a set of items and never attempts to separate them or to give more importance to one or the other of these components. This conception takes into account the fact that language is composed of both *objects* and actions *bearing on these objects*.

Now, once we are concerned with characterising types of action, simple commutation, even if supplemented with an accessory look at the contrasts that appear on the syntagmatic axis, or simple distributional analysis, even if followed by a similary, secondary look at pragmatic variations, is not sufficient to define and characterise these actions. Such a conception (of an integration of class and order, of the categorial and the serial) implies that of a structuralisation of the pradigmatic axis: the substitutions made there are first of all true transformations, but these are affected through reversible expansion from an irreducible kernel. Finally, such a viewpoint implies that we cannot be satisfied with an analysis into immediate constituents, even if formalised into tree diagrams, and even if

this analysis is supplemented by a battery of transformations, for it is the *transformations themselves* which (in the conception of the *naḥw*) define the constituents on the utterance.

In fact, no linguistic theory seems to us to fit in completely with the *radical* an *integral operationism* of Arab linguists. Chomsky's reintroduction in scientific form of the concepts of transformation and grammatical rule, and the application of the concept of formal generation to language seem to us to offer confirmation of certain aspects of the naḥw, but generative grammar does not (yet) seem to be moving in the direction of a more systematic integration of the transformational into the syntagmatic.

Another kind of transformation consists in incorporating into what is grammatical forms which have been distorted. This brings the *naḥw* closer to generativism, but it is nevertheless different: by the fact that only the distorted forms, together with the free variations in pattern content, receive formal explanation by recourse to this kind of transformation and secondly by the fact that the recursivity of its rules is extensive.

The linguistics of the *'Arabiyya* has not been limited to the study of the meaningful form (the lafḏ), as we have seen: it is concerned above all with the *kalām*, that is the utterance and the act of uttering. In fact, the study of what the *waḍ'* (code and structure) becomes in the act of utterance, and consequently of the variations and distortions actually affecting it, is for Arab linguists only another dimension to be added to the naḥw in the strict sense, since in this study they continue to make use of the same biunivocal and seriated correspondences. The signifying forms at the syntactic level certainly have a semantic content in acts of utterance, but they are not isomorphic with this content (the Arab grammarians give much importance to this). The solution consisted in establishing a mediation or potentialisation (which presupposes these reversible and seriated correspondences: semantic feature implied by the *lafḏ* alone R semantic feature implied by the components of the communication where R is often an implication based on natural logic).

We can say at the end of this study that the Arabic linguistics which we have just outlined, and which contrasts with the speculative and/or exclusively normative grammar inherited from later centuries, depends on an operationist conception which can on these grounds be composed, but not reduced to, certain concepts of contemporary science with regard to its fundamentally operationist approach.

Notes

1- « Faṣāḥa » for *Sībawayh*, is the property of being a native speaker who has not spent much time in linguistically cosmopolitan areas.

2- What seems to justify this viewpoint is the fact that there is no difference, after all, between the data of the physician (or biologist) and those of the linguist: in both cases any researcher *must be able to verify*, at any moment, an observation obtained by another researcher.

3- Let us remember that for us the notion of structure cannot be reduced to intensive relations of opposition and co-membership (the importance of which has been These very much exaggerated by the post-saussurian structuralist linguists). These relations or the classes deriving from then must still be *commounded among themselves before we can talk of structure.*

4- Later authors than Sībawayh have given this entity the very-significant name of *"lafḏa"* (= unit of *lafḏ*) (see Ibn Yaʿīš, Šarḥ al-Mufaṣṣal, I, 19 and Raḍī, Šarḥ al-Kāfiya, 1, 5).We have elsewhere proposed calling it *"lexie"*. See for more details (on the whole of this study) my work " Arabic linguistics and general linguistics".

5- Ḥadd as a production and reception model, or in the formal framework of the naḥw: a structure capable of *characterising* (in the mathematical sense) *the items of the language.*

6- *Furūʿ* derived from the aṣl and having the same *isotopes* or *mawādiʿ* in speech.

7- ʿA. is standing, ʿA. is indeed standing, ʿA. was standing.

8- This concept of *government* is unknown among the Greco-Latin and Indian predecessors of Arabic grammar. In the West it is only met for the first time among the author of the late Middle Ages.

9- ʿA. hit *ʿAmr.*

10- The *mubtada'*, contrary to what some believe, did not receive this name because it is at the beginning of a string, since it may happen to be proposed (ibtidā' here = independence from what precedes).

11- Translation: Z. was standing; I was standing; *Z. hit ʿAmr; I hit ʿAmr; Z. gave ʿAmr* a garnent; Ḫālid thought that Z. was standing; I thought that Ḫ was standing; Ḫ. informed Bakr that Z. was standing; I informed Bakr that Z. was standing.

As can be seen, the content of R is capable of dividing T_1 and T_2 into T_3 and T_4.

12- Minimal string none of whose members is realised separately in the same string.

13- Which is not the case with the lexie *"kitābu Zaydin hāḏā"* (Zayd's book, that is). In

fact, only the verbal lexie is analysable into syntactic elements

14- This substitution accompanies a structural transformation of the verb as can be seen.

15- I hit 'A. standing; 'A, was hit standing.

16- There is 'A. who is going away.

17- 'A. goes away.

18- As in: / # 'imtala'a al-'inā'u/mā'an # the vase has filled with water.

19- = a ring of silver; twenty dirhams; the most wealthy; equal in strength; how many books?; how many books !

20- At this intermediate level are found the terms governed by lexies whose kernel is constituted by an item related to a verb: a verbal substantive, participes, etc.

21- Which covers everything which is not a verb or a grammatical morpheme.

22- The generativits, « matrix ».

23- = Zayd, his father says that.

24- I was afraid that Zayd would go out; I want to go out; I appreciated what you did.

25- This follows from the expandable nature of some mawḍi' -s

26- =didn't you go out? -wil he go out? - If I go out, will you punish me? "lam" marks the negation of a process in the past; "sa –" is a verbal prefix of the future, and "in" = if.

27- Sībawayh also calls it "isti'nāf ".

28- One difference is to be noted: the second series, unlike T_2, cannot be anteposed to its governing element or to the first series.

29- Without adventions elements (muǧarrad)

30- mā" is here treated as "laysa" (this is the Ḥiǧāzian variant).

31- Translation : Zayd goes away-Zayd doesn't go away in order that I go away-Zayd does no go away - Does Zayd go away? - Doesn't Zayd go away? (same meaning with laysa) - Does Zayd go away in order that I go away? Zayd goes away! (strong affirmation) - Would to God that Zayd would go away - I have no book to be able to read Zayd was going away – If Zayd goes away, I do the same - If Zayd does not get up, I will punish him - If Zayd came, I would welcome him - (By God), if you hit 'Amr, I will hit you - Hit 'Amr! - Do not hit 'Amr! - (by God), I will go out! Hit 'Amr! (insistent order) - If Zayd had not been there, I would have died.

32- It is important to emphasise the fact that these transformations are never arbitrary, they are implied by the already established system of patterns of the language and thus contribute to expanding and enriching this by incorporating into it new, more abstract

patterns.

33- This corresponds to a contraction of two consonants caused by the dropping of a vowel and resulting in a geminate. The translation of this word as "assimilation" is an error (Ibn Ğinnī is the only one to have taken it in this sense. His Ḫaṣā'iṣ, II, 139), as is proved by the existence of idġām without assimilation as in " t̲awbu Bakr" > *"t̲awbbakr"* (see Kitāb, II, 408-9).

34- As we have said, combinatory phonetics occupies an important place in Arabic linguistics.

35- The distinction between consonants and vowels is found, but in a framework based on continuity and movement, in the opposition (not discontinuous): ṣawt or ḥarf ḍawā'ib / ğawāmid. The terms muṣawwitāt/ṣawāmit are the translations of the Ereek terms phoneenta/aphona.

Bibliography

- Abd al-Ǧabbār (al-Qādī), al-Muǧni, vol. VII, le Caire, 1961.

- Abū ʿAlī al-Fārisī, al-Takmila, Bibliothèque Nationale, Le Caire, manuscrit n° 100.

- Al-Ǧāḥiḏ, al-Bayān wa-l-tabyīn, 4 vol. en 2 tomes, éd. A. Harūn, Le Caire, 1960.

- Al-Ǧurgāni (ʿAbd al-Qāhir al-), Dalāʿil al- ʾigāz, éd. R. Rida, Le Caire, 1335 H.

- Al-Ḫalīl (Ibn Aḥmad al-Farāḥīdī), Kitāb al- ʿAyn, t. I, éd. A. Darwīš, Baghdad, 1967.

- Al-Mubarrad (Abu l- ʿAbbās), al-Muqtaḍab, 4 vol., Le Caire, 1385 -1388 H.

- Al-Raḍī (al-Astrābādī), Šarḥ al-Kāfiya, Istamboul, 1275 H.

- Al-Rummānī (ʿAli b. ʿīsā), Šarḥ Kitāb Sībawayh, 4 vol., microfilm n° 88, de l'Inst des Manuscrits arabes, Ligue Arabe.

- _____ , Kitāb al-Ḥudūd, éd. M. Ǧawād, Baghdad, 1969.

- Al-Suyūṭī, al-Iqtirāḥ, Hayderabad, 2e éd., 1359 H.

- Hadj-Salah (A.), Linguistique arabe et linguistique générale. Essai de méthodologie et d'épistémologie du ʿilm al-ʿArrabiyya, Thèse dactylographiée, Paris - Sorbonne, 1979.

- Ibn Ǧinnī, al-Ḫaṣāʾiṣ, éd. M. al-Naǧār, 3 vol., Le Caire, 1952-1956.

- _____ , Sirr Ṣināʿat al-ʾ iʿ rāb, t. I, Le Caire, 1954.

- Ibn al-Sarrāǧ (Abū Bakr),Uṣūl al-naḥw, Bibliothèque Générale, Rabat, man. n° 34.

- Ibn Yaʿīš, Šarḥ al-Mufaṣṣal, 10 vol., Le Caire, s.d.

- Lyons (J.), Introduction to Theoretical Linguistics, Cambrige, 1968.

- Revzin (I, -I), Les modèles linguistiques, trad. franç., Paris, 1968.

- Sībawayh, al-Kitāb, Būlaq, 2 vol., 1317 H.