

# PARSING ARABIC TEXT WITH ATN : PROBLEMS AND IMPLICATIONS

HASSAN A. TAMAN Ph. D.  
ALEXANDRIA UNIVERSITY  
EGYPT

## INTRODUCTION :

A natural language that fails to accommodate the whole range of computer applications, or accommodates them to an unsatisfactory degree will cease to exist. The reason being that processing language data by means of machine is no longer a linguistic luxury, but rather an indispensable implication of modern technology. By "whole range of computer applications" is meant here all possible forms man-machine communication in natural language be it simple processing of language data, or highly sophisticated intelligent systems.

Arabic has been the subject of numerous studies in a computational framework, ranging from getting the computer to recognize Arabic characters (See Amin et al. 1980, 1982, Gazal 1987, Mounajed 1987), and some primary analysis of Arabic morphology and syntax (e.g. Moussa 1987, Beesley 1989a, Beesley et al. 1989b, Fehri 1987), to speculations about man machine communication in Arabic (Mrayati and Makhoul 1987) Arabic speech synthesis (Mouradi et al 1987) and computer-aided translation (Vauquois 1987).

In fact, the above and similar attempts are far from being completely successful, particularly when it comes to parsing and content processing; in addition to the various problems of processing natural language in general (see Garvin 1986), Arabic has some peculiar characteristics that make some aspects of this endeavour relatively more complicated. The aim of this paper is to address some of those peculiarities as they impinge on parsing Arabic input with ATN. Implications and suggestions for natural language processing in general will also be presented.

## COMPUTABILITY AND TYPOLOGICAL PROPERTIES OF ARABIC ORTHOGRAPHIC AMBIGUITY :

Arabic orthography has the peculiar characteristic of representing only consonants and long vowels; short vowels go unrepresented. In spite of the fact that a system of diacritics is prescribed to eliminate this problem, it is employed very inconsistently if ever. The result is a tremendous amount of ambiguity that cripples the processing of Arabic by machine. This problem is often overlooked in the writings on Arabic language data processing, or is sometimes treated unsatisfactorily (see Taman, 1986).

It should be noted that orthographic ambiguity cannot be handled in a backtracking component of RTN or ATN due to its magnitude. The system, MIDS, (Taman, 1986) eliminates this ambiguity to a satisfactory degree through a "Graduated Context Search" technique. To illustrate the difficulty of parsing Arabic input strings that are orthographically ambiguous, let us consider the following ambiguous string:

1 -	nšr	katb	mçrwf	ktb	alshr	hdh
	1	2	3	4	5	6

In the above sentence, elements 1 and 2 can be interpreted in any of the following ways:

a -	* V	N	ungrammatical
b -	* V	N	ungrammatical
c -	V	N	grammatical
d -	N	V	grammatical

## ملخص البحث

إن الهدف من هذه الدراسة تبيان بعض النظريات العربية والغربية المطبقة على الجملة العربية ثم البرهنة على أن هذه النظريات ليست كافية لسر البنية النحوية والدلالية للجملة العربية. لذلك تقترح هذه الدراسة نموذجاً لسانياً حديثاً لتحليل الجملة العربية نحويًا ودلاليًا. يستمد هذا النموذج الجديد مقوماته من ثلاثة مصادر.

الأول: النظرية الجمالية العربية التي وضعها العرب القدامى في القرن الثامن الميلادي.

الثاني: النظرية التوليدية والتحويلية التي وضعها عالم اللسانيات نوم تشومسكي وبالتحديد منهج العامل والربط الإحالي (G. B. Theory).

الثالث: النظرية الدلالية التي وضعها عالم اللسانيات ولتر كوك والمسماة (Case Grammar).

إن هذا النموذج اللساني الحديث سيطبق على الجملة العربية الأساسية ليبين كيف يمكن لهذه الجملة أن تتولد نحويًا ودلاليًا. أضف إلى ذلك أن هذا النموذج سوف يرصد عمليتي التقديم والتأخير للعناصر اللغوية في الجملة العربية ويرصد بالتالي القواعد التحويلية التي تقوم بهاتين العمليتين ثم الضوابط المفروضة على هذه القواعد وذلك من أجل توليد جمل صحيحة.

تأمل هذه الدراسة - من خلال هذا التحليل النحوي والدلالي الدقيق لتوليد الجملة العربية - أن تسهم في تطوير اللسانيات الحاسوبية - المعلوماتية العربية لتكون أكثر فهما لبنية اللغة العربية علمياً وتقنياً.

- \_\_\_\_\_. 1980. "Introduction to Generative Semantics." Mimeograph. Washington, D.C.: Georgetown University.
- Dik, S. 1978. Functional Grammar. Ed. North Holland. Amsterdam.
- Dinneen, Francis P. 1967. An Introduction to General Linguistics. Holt, Rinehart and Winston, Inc.
- Dreshner, E. and Hornstein, N. 1979. "Trace Theory and NP Movement Rules." Linguistic Inquiry 10.
- Emonds, Joseph E. 1976. A Transformational Approach to English Syntax : Root, Structure-Preserving, and Local Transformation. Academic Press.
- Fehri, Fassi, Abdelkader. 1981. "Complementation et Anaphore en Arabe Moderne : Une Approche Lexicale Fonctionnelle." Thèse de Doctorat d'Etat. Université de Paris III.
- Ibn yaʿqūb (d. 1250). Šarḥ ʿal-mufašṣal. Ed. ʿālam ʿal-kutub, Beyrut, 1970.
- Ibn Hišām (d.1368). ʿal-muğnī. Ed. Al-mubārak and M. A. Ḥamadallah. Dār al-fikr, 1969.
- Ibn jinni (d. 1002). ʿal-xaṣāʾiṣ. Ed. ʿan-najjār M.A. Dār ʿal-hudā liṭṭibāʿah wan-našr, Beirut, 1952.
- Killeen, C. 1966. "The Deep Structure of the Noun Phrase in Modern Written Arabic." Unpublished Ph. D. dissertation, University of Michigan, Ann Arbor, Michigan.
- Snow, J. 1965. "A Grammar of Modern Written Arabic Clauses." Unpublished Ph. D. dissertation, University of Michigan, Ann Arbor, Michigan.

- Bresnan, Joan. 1976. "On the Form and Functioning of Transformations." In Linguistic Inquiry 7 (1).
- \_\_\_\_\_. 1980. "The Passive in Lexical Theory." Occasional Paper No. 7. M. I. T.
- \_\_\_\_\_. 1982. The Mental Representation of Grammatical Relations. In Press. M. I. T.
- Chafe, Wallace L. 1970. Meaning and the Structure of Language. University of Chicago Press.
- Chomsky, Noam. 1957. Syntactic Structures. The Hague, Paris : Mouton.
- \_\_\_\_\_. 1965. Aspects of the Theory of Syntax. M. I. T. Press.
- \_\_\_\_\_. 1970. "Remarks on Nominalization." In Readings in English Transformational Grammar. Ed. R. A. Jacobs and P. S. Rosenbaum. Washington, D.C. : Georgetown University Press.
- \_\_\_\_\_. 1971. "Deep Structure, Surface Structure, and Semantic Interpretation." In Semantics: An Interdisciplinary Reader in Philosophy, Linguistics, and Psychology. Edited by Steinberg and Jakobovits, pp. 183-216.
- \_\_\_\_\_. 1975. Reflections on language. Pantheon Books.
- \_\_\_\_\_. 1977. Essays on Form and Interpretation. North-Holland Publishing Company.
- \_\_\_\_\_. and Lasnik, Howard. 1977. "Filters and Control." Linguistic Inquiry 8 (3).
- \_\_\_\_\_. 1977. "On WH-Movement." In Wasow Culicover and Akmajian, eds., Formal Syntax. Academic Press.
- \_\_\_\_\_. 1978. "On Binding." Linguistic Inquiry 11 (1).
- Chomsky, Noam. 1981. Lectures on Government and Binding. Foris publications.
- \_\_\_\_\_. 1981. "Lectures on Government and Binding." Lecture, M. I. T. class. Fall 1981.
- Cook, Walter A. 1976. "Five Models of Case Grammar." Mimeograph. Georgetown University, Washington, D. C.
- \_\_\_\_\_. 1979. Case Grammar: Development of the Matrix Model. (1970-1978) . Washington, D. C. : Georgetown University Press.
- \_\_\_\_\_. 1979. "Case Grammar and Generative Semantics." Georgetown University Press on Languages and Linguistics, Number 8. Washington, D.C.: Georgetown University Press.

## BIBLIOGRAPHY

- Aj-jurjāni (d. 1078). Dalāʾil ʾal-ʾiʿjāz. ed. Maktabatu l-qāhirah, 1961.  
\_\_\_\_\_. ʾaj-jumal. Ed. Ḥaydar, A. Damascus, 1972.  
\_\_\_\_\_. ʾal-ʿawāmil ʾal-miʾah. Ed. in ʾstanbul, 1924.
- Al-ʾistrābāʾi (d. 1289). Ṣarḥ ʾal-Ṣāfiyah. Dār ʾal-kutub ʾal-ʿilmiyyah. Beirut, 1975.  
\_\_\_\_\_. Ṣarḥ ʾal-kāfiyah. Dār ʾal-kutub ʾal-ʿilmiyyah. Beirut, 1957.
- Al-waer, Mazen. 1979. "The Semantic and Syntactic Frame Structure of the Verb 'SEE' in Arabic and English." Mimeographed. Georgetown University, Washington, D.C.  
\_\_\_\_\_. 1979. "Teaching English as a Foreign Language from a Case Grammar Point of View." Mimeographed. Georgetown University, Washington, D.C.  
\_\_\_\_\_. 1980. "Beyond the Symbolic System." Mimeographed. Georgetown University, Washington, D.C.  
\_\_\_\_\_. 1982. "On Some Controversial Issues of Transformational Generative Grammar." Allisāniyyāt, Vol. 6. Algerian Linguistics Institute, Algiers.  
\_\_\_\_\_. 1980. "Linguistics: From Structuralism to Transformationalism." ʾal-maʿrifa, Nos. 220 and 221. Ministry of Culture and National Guidance in Syria, Damascus.
- Anshen, F. and P. Schreiber. 1968. "A Focus Transformation of Modern Standard Arabic." Language 44 : 792-97.
- Aoun, Joseph. 1979. "Ambiguity and Metric : The Symmetry Constraint." Alfikr Al-Arabi, Nos. 8 and 9.  
\_\_\_\_\_. 1981. "ECP, Move- $\alpha$ , and Subjacency." In Linguistic Inquiry 12 (4).  
\_\_\_\_\_. 1981. "Parts of Speech." Ms. MIT.
- Awwad, M.A. 1973 "Relativization and Related Matters in Classical Modern Standard and Palestinian Colloquial Arabic." Unpublished Ph. D. dissertation, Brown University, Providence, Rhode Island.
- Boas, Franz. 1964. Introduction to the Handbook of American Indian Languages. Georgetown University Press, Washington, D.C.
- Brame, M. 1970. "Arabic Phonology: Implications for Phonological Theory and Historical Semitic." Unpublished Ph.D. dissertation. M.I.T.

## 5. Conclusion

This work drew its theoretical framework from three sources: the transformational generative grammar proposed by Chomsky, the case grammar matrix model proposed by Cook, and the Arabic grammar proposed by the early Arab grammarians in the eight century A.D.

The investigation showed that Arabic sentence consists of three constituents, namely, MI 'topic' and 'subject', M 'predicate' and F 'adjunct'. The relation holding among these constituents is IS predication. When such constituents are organized, the outcome is sentence. The structure of sentence is produced either by base - generated rules or by transformational rules which generate various general and specific meanings.

In the structure (15b), the MI-subject is postposed to the right of the verb. In structure (16b), the MI-Topic is preposed to the left of the verb, but semantically, they have one general meaning, i.e., "Zayd stood" (Ibn jinni; ?al-xaṣāʔiṣ 1, p. 343). Ibn jinni's idea is similar to the semantic framework of Cook (1979). Cook viewed different syntactic categories to have one semantic structure which is determined by the case roles. The case roles can assign one semantic structure of the above two examples (15b) and (16b).

(17). gāma, [ + ————— A, o ] / A = o

Ibn jinni fully knew that syntactically the MI-subject never precedes its governor, i.e., (M) or the verb. But for a focus function one can prepose the MI-subject in the base and substitute it by an empty category which is (e) or damīr mustatir which must be adjacent to the right of the M. Ibn jinni stated these facts when he said :

?alā tarāka ?iḏā suʔilta ʿan Zaydin min qawlinā (qāma Zaydun )  
 sammayta-hu fāʿilan wa ?in suʔilta ʿan Zaydin min qawlinā  
 (Zaydun qāma ) sammayta-hu muḩtadaʔan lā fāʿilan wa ?in kāna  
 fāʿilan fi l-maʿnā (Ibn jinni, ?al-xaṣāʔiṣ 1, p. 343).

This paragraph translates as follows: "You see, when you are asked about Zayd in qāma Zaydun [ Zayd stood ] you will call it 'subject,' but when you are asked about Zayd in another structure such as Zaydun qāma [As for Zayd, he stood ] you will call it 'topic' and not 'subject' even though they are semantically agents." This means that Ibn jinni was distinguishing syntactically between the preposed MI-Topic and the postposed MI-subject even though they mean the same thing semantically.

This suggests that Arab Grammarians and linguists distinguished between two types of semantic structures. The first type was a general semantic notion: i.e., we have different syntactic structures with one general semantic structure (Ibn jinni). The second type was a specific semantic notion: i.e., we have different syntactic structures which bear different specific semantic functions (?aj-jurjāni).

As a matter of fact, the first semantic type can collapse under the second semantic type because different syntactic orders of the constituents have one general meaning that is a particular event: actor of the event and the object which is acted upon. The specific semantic function, however, results from the intentions of the speaker-hearer's knowledge of his language which concentrate on one and only one physical reality of the above three roles, i.e., event, actor, and object. That was the cornerstone in the logical structure of the semantic theory as perceived by ?aj-jurjāni).



(12). a.  $\frac{?a}{Q}$   $\frac{\text{munṭaliqun}}{\text{the departed}}$   $\frac{\text{Zaydun } t}{\text{Zayd}}$  ?

Is Zayd departed ?

b.  $\frac{\text{Na}^{\text{c}}\text{am}}{\text{Decl}}$   $\frac{\text{munṭaliqun}_i}{\text{the departed}}$   $\frac{\text{Zaydun } t}{\text{Zayd}}$

Yes, Zayd is departed.

(13). a.  $\frac{?ā}{Q}$   $\frac{l - \text{munṭaliqu}}{\text{the departee}}$   $\frac{\text{Zaydun}}{\text{Zayd}}$  ?

Is the departee Zayd ?

Base-Generated

b.  $\frac{\text{Na}^{\text{c}}\text{am}}{\text{Decl}}$   $\frac{?al - \text{munṭaliqu}}{\text{the departee}}$   $\frac{\text{Zaydun}}{\text{Zayd}}$  ?

Yes, the departee is Zayd.

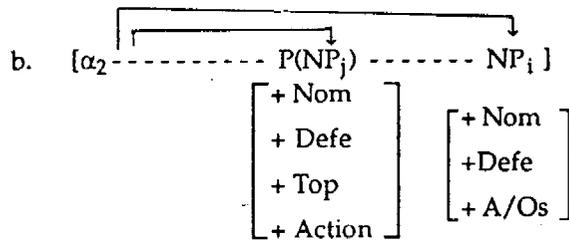
In the above examples, the fronted constituents are preposed by either Move- $\alpha$  or by the base. They are preposed as a result of presuppositional questions.

The most important semantic function of the preposed constituents is to convey concentrated semantic information which is very important to the speaker-hearer's communicative knowledge. And as we have seen, even though the structural process is different in preposing these constituents, the semantic process is identical, because both Move- $\alpha$  constituents and base-generation constituents convey the same semantic information which is considered to be "focus" or "theme" in Dik's (1978) terminology, "interest" or "importance" in ?aj-jurjāni's terminology, and "topicalization" in Chomsky's (1977) terminology. The virtue of ?aj-jurjāni, however was that he tried to seek a more comprehensive explanation of the preposed and postposed phenomenon. It was not enough for him to say that the preposed or postposed constituents serve as focus or topicalization only. He went beyond such an explanation to establish a presuppositional semantic system by which we can figure out the nature of the preposed and postposed constituents. He proposed a semantic system with a question-declarative process which can explain the computational relations in the human mind (?aj-jurjāni; Dalā'il ?al-?i'jāz, pp. 74-76).

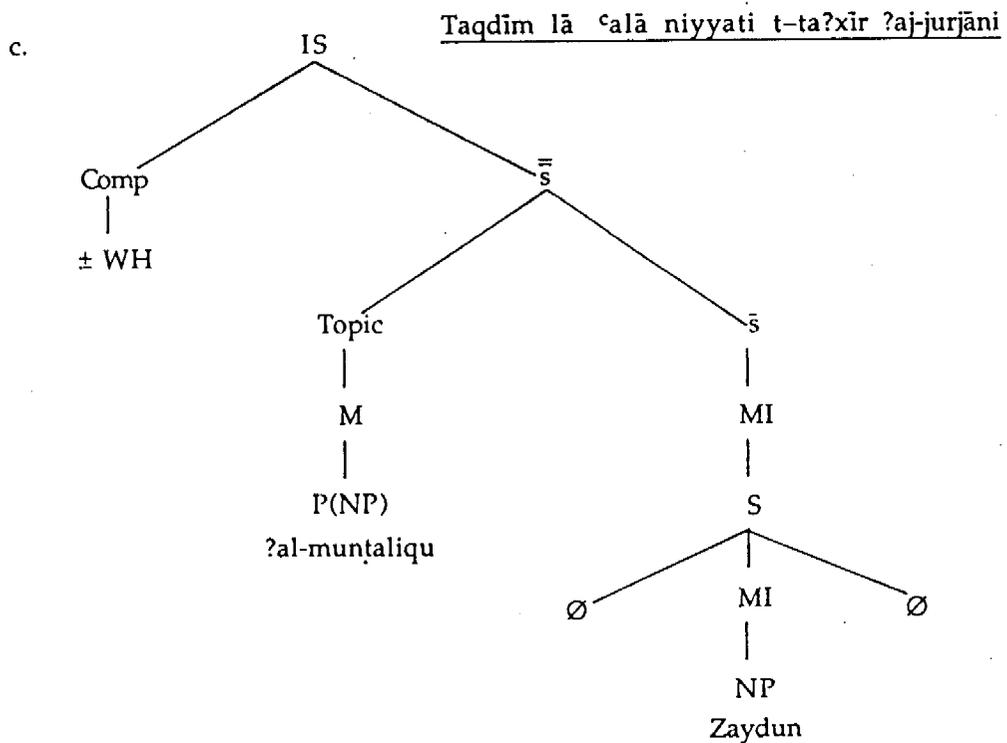
The general system of preposing and postposing constituents which bear different semantic functions is exhibited in the following Figure (14).

(14).		[ Comp -----	M -----	MI -----	F]
		+ WH -----	V -----	S -----	O
a.	If	[ 1 -----	2 -----	3 -----	4]
	Then	[ Decl -----	2 -----	3 -----	4]
b.	If	[ 1 -----	3 -----	2 -----	4]
	Then	[ Decl -----	3 -----	2 -----	4]
c.	If	[ 1 -----	4 -----	2 -----	3]

- (7). a.  $\frac{?a}{Q}$   $\frac{^c\text{amrun}}{^c\text{amr}}$   $\frac{\text{ḍaraba}}{\text{hit}}$   $\frac{[e]}{\text{ec}}$   $\frac{\text{Zaydan}}{\text{Zayd}}$   
Is it  $^c\text{amr}$  who hit Zayd?
- b.  $\frac{\text{Na}^c\text{am}}{\text{Decl}}$   $\frac{^c\text{amruni}}{^c\text{amr}}$   $\frac{\text{ḍaraba}}{\text{hit}}$   $\frac{[e]_i}{\text{ec}}$   $\frac{\text{Zaydan}}{\text{Zayd}}$   
Yes, it is  $^c\text{amr}$  who hit Zayd.
- (8). a.  $\frac{?a}{Q}$   $\frac{\text{Zaydan}}{\text{Zayd}}$   $\frac{\text{ḍaraba}}{\text{hit}}$   $\frac{^c\text{amrun}}{^c\text{amr}}$   $\frac{t}{?}$   
Is it Zayd, that  $^c\text{amr}$  hit ?
- b.  $\frac{\text{Na}^c\text{am}}{\text{Decl}}$   $\frac{[\text{Zaydan}_i]}{\text{Zayd}}$   $\frac{\text{ḍaraba}}{\text{hit}}$   $\frac{^c\text{amrun}}{^c\text{amr}}$   $\frac{t_i}{?}$   
Yes, it is Zayd, that  $^c\text{amr}$  hit.
- (9). a.  $\frac{?a}{Q}$   $\frac{\text{rākiban}}{\text{riding}}$   $\frac{\text{jā?a}}{\text{came}}$   $\frac{\text{Zaydun}}{\text{Zayd}}$   $\frac{t}{?}$   
Is it by riding, that Zayd came ?
- b.  $\frac{\text{Na}^c\text{am}}{\text{Decl}}$   $\frac{[\text{rākiban}_i]}{\text{riding}}$   $\frac{\text{jā?a}}{\text{came}}$   $\frac{\text{Zaydun}}{\text{Zayd}}$   $\frac{t_i}{?}$   
Yes, it is by riding that Zayd came .
- (10). a.  $\frac{?a}{Q}$   $\frac{\text{jā?a}}{\text{came}}$   $\frac{\text{rākiban}}{\text{riding}}$   $\frac{\text{Zaydun}}{\text{Zayd}}$   $\frac{t}{?}$   
Did Zayd come by riding ?
- b.  $\frac{\text{Na}^c\text{am}}{\text{Decl}}$   $\frac{\text{jā?a}}{\text{came}}$   $\frac{[\text{rākiban}_i]}{\text{riding}}$   $\frac{\text{Zaydun}}{\text{Zayd}}$   $\frac{t_i}{?}$   
Yes, Zayd came by riding.
- (11). a.  $\frac{?a}{Q}$   $\frac{\text{fī}}{\text{in}}$   $\frac{\text{d - dāri}}{\text{the house}}$   $\frac{\text{Ta?abbaṭašarran}}{\text{Ta?abbaṭašarran}}$   $\frac{t}{?}$   
Is it the house that Zayd is in ?
- b.  $\frac{\text{Na}^c\text{am}}{\text{Decl}}$   $\frac{\text{fī}}{\text{in}}$   $\frac{[\text{d - dāri}_i]}{\text{the house}}$   $\frac{\text{Ta?abbaṭašarran}}{\text{Ta?abbaṭašarran}}$   $\frac{t_i}{?}$   
Yes, it is the house that Zayd is in.

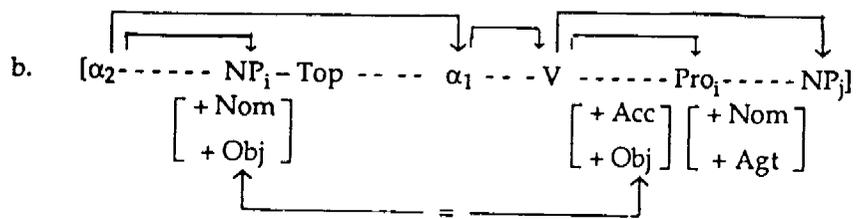


The preposed constituent in the base can be seen through the following process in the configuration (6c) below.

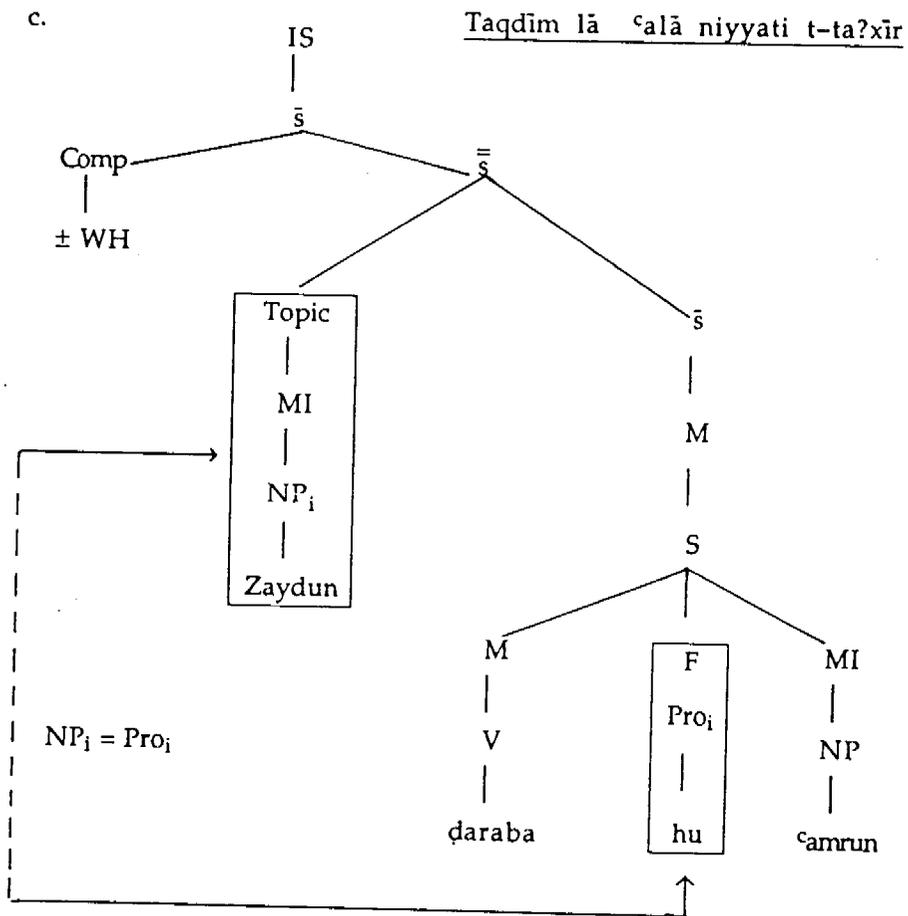


As seen in the above examples, the base-generated a preposed constituent as the topic of the sentence either in the complex nominal structure as in (5) or in the equational structure as in (6). Preposing either by Move- $\alpha$  or by base-generation in the preposed position reflects a presupposition on the part of the speaker. This semantic presupposition forces us to consider the declarative sentence to be dependent on the question formation, because whenever the position of the constituent can be in the question formation it will appear in the declarative sentence. The preposed constituent in either case serves as a focus on which the speaker-hearer's knowledge of their language concentrates (?aj-jurjāni; Dalāʿil ʿal-ʿiʿjāz, pp. 70-80).

These systematic principles of the presuppositional semantics which ʿal-jurjāni tried to establish can be clarified through the following examples.

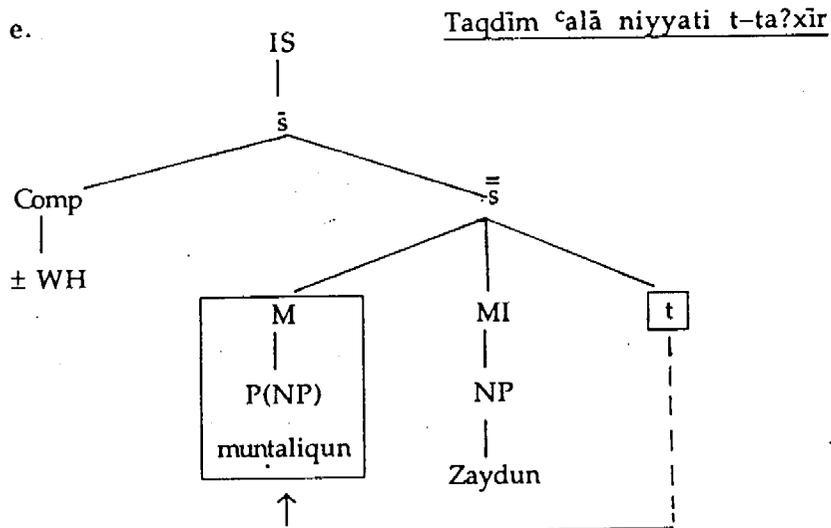


The preposed base-generated constituent can be seen in the configuration (5c) below.

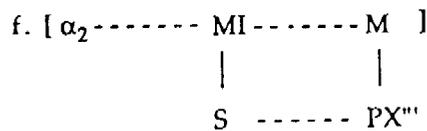


There is the same preposed base-generated constituent in the equational sentence, this being shown in the following examples.

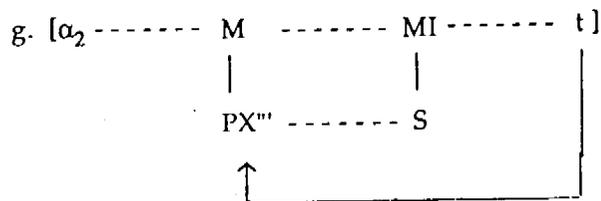
- (6). a. ʿal-manṭaliqu Zaydun  
 the departee Zayd  
 The departee is Zayd.



The semantic structure in (4e) is a result of the basic nominal equational structure whose base-generated word order is in (4f).



The syntactic structure of (4f) can bear one transformational semantic structure whose preposed constituent serves as a focus as in (4g).



Thus, Move- $\alpha$ , operating on both verbal sentences as in (3d), and equational sentences as in (4d) bear semantic relations of focus and interest. The only syntactic condition on Move- $\alpha$  in the equational sentence is that the constituent PX''' must be indefinite as we saw earlier.

The other theoretical issue in ʿaj-jurjāni's theory is that Move- $\alpha$  cannot be involved in some other structure even though their constituents are preposed. Such preposed constituents would be generated in the base. The criterion of preposing constituents in the base can occur in both types of nominal structures, i.e., complex nominal structures and equational structures. Let us consider the following examples.

Taqdīm lā ʿalā niyyati t-taʿxīr

- (5). a.  $\frac{\text{Zaydun}}{\text{Zayd}}$   $\frac{\text{ḍaraba - hu}}{\text{hit him}}$   $\frac{\text{ʿamrun}}{\text{ʿamr}}$   
 As for Zayd, ʿamr hit him.

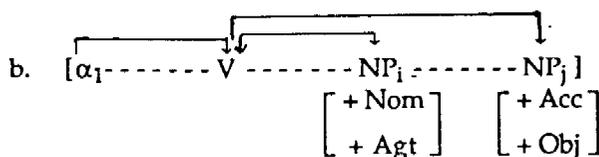


hit ʿamr]. The second type of preposing is 'preposing which is not intentionally postposing' and here the characteristics of this structure are changed syntactically and semantically such as, for example, Zaydun darabtu-hu [As for Zayd, I hit him]. Here, Zayd is not preposed from the object position, and the verb is not the governor of Zayd assigning it a case role and case marker, but rather, it is assigned a nominative case from topicalization and the verb will be busy with its resumptive pronoun and the verb and its resumptive pronoun are the predicate of Zayd ."

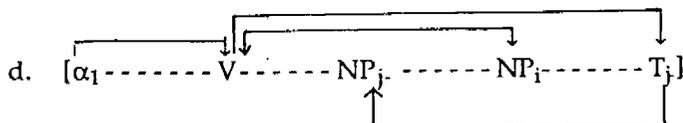
This paragraph summarized the syntactic and semantic notions of Move- $\alpha$  and base-generation. It means that constituents can be preposed transformationally via Move- $\alpha$  or via base-generation principles. The result is different semantic structures as shown in the following examples which were introduced by ʿaj-jurnāni.

Taqdīm ʿalā niyyati t-taʿxīr

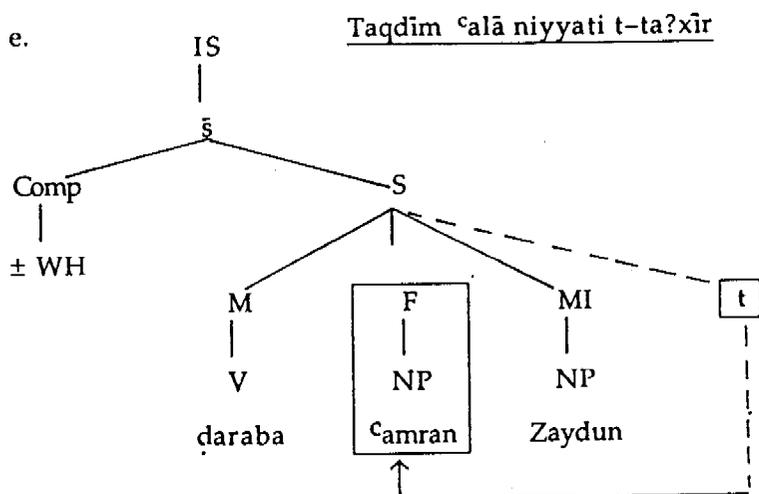
- (3). a.  $\frac{\text{daraba}}{\text{hit}} \frac{\text{Zaydun}}{\text{Zayd}} \frac{\text{ʿamran}}{\text{ʿamr}}$   
 Zayd hit ʿamr.



- c.  $\frac{\text{daraba}}{\text{hit}} \frac{\text{ʿamran}}{\text{ʿamr}} \frac{\text{Zaydun}}{\text{Zayd}} t$   
 Zayd hit ʿamr.



The transformational process of such a structure as (3d) can be seen in the following configuration.



As a matter of fact, ?aj-jurjāni tried to interpret the functional aspects of the word order from a clear systematic framework which goes beyond the shallow explanation of preposing and postposing constituents (?aj-jurjāni; Dalā?il ?al-?j?jāz, p. 206). ?aj-jurjāni proposed two types of preposing and postposing of constituents in the basic sentence. The first type he called Taqdīm ʿalā niyyati t-ta?xīr, i.e., "preposing which is

intentionally postposing. ?aj-jurjāni meant by this criterion the transformational constituents which are a result of Move- $\alpha$ , i.e., even though the constituent is moved to another position it still has a trace of its syntactic and semantic properties that trace is governed by its original operator in the verbal or nominal structures". The second type he called Taqdīm lāʿalā niyyati t-ta?xīr, i.e., "preposing which is not intentionally

postposing. ?aj-jurjāni meant by this criterion the preposed constituents which are a result of base-generated principles in the nominal and verbal sentences in Chomsky's (1981) sense. Here the preposed constituent has different syntactic and semantic properties which are different from the transformational ones." Summing up, the difference between these two types of preposing is the difference between the constituent preposed by Move- $\alpha$  leaving a trace behind, and the constituent preposed by base-generation which leaves no trace.

?aj-jurjāni explained in a sophisticated way the whole syntactic notion of Move- $\alpha$  and base-generation and the whole semantic notion of preposing and postposing constituents when he said:

wa ʿlam ?anna taqdīma š-šay?i ʿalā wajhayn ; Taqdīmun yaqālu ?inna-hu  
ʿalā niyyati t-ta?xīr , wa ḏālika fī kulli šay?in ?aqrarta-hu maʿa  
t-taqdīm ʿalā ḥukmi-hi li-laḏī kāna ʿalay-hi wa fī jinsi-hi l-laḏī kāna  
fī-hi kaxabari l-mubtada? ?iḏā qaddamta-hu ʿalā l-mubtada? , wa l-mafʿūli  
?iḏā qaddamta-hu ʿalā l-fāʿil , kaqawlika ( muntaliqun Zaydun ) wa  
( ḏaraba ʿamran Zaydun ) . wa taqdīm lā ʿalā niyyati t-ta?xīr , walākin  
ʿalā ?an tanqula š-šay?a ʿan ḥukmin ?ilā ḥukm . wa tajʿala-hu bāban ḡayra  
bābi-hi wa ?iʿrāban ḡayra ?iʿrābi-hi kaqawlika ( Zaydun ḏarab-tu-hu ) .  
lam tuqaddim ( Zaydan ) ʿalā ?an yakūna mafʿūlan manšūban bi-l-fiʿli  
kamā kān . wa lākin ʿalā ?an tarfaʿa-hu bi-l-ʔbtidā? wa tušgila  
l-fiʿla bi-ḏamīri-hi wa tajʿala-hu fī mawḏiʿi l-xabari la-hu"  
(?aj-jurjāni; Dalā?il ?al-?j?jāz, p. 73).

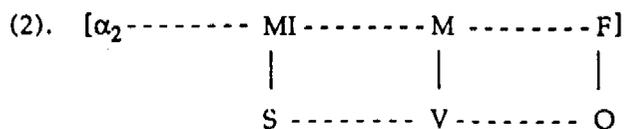
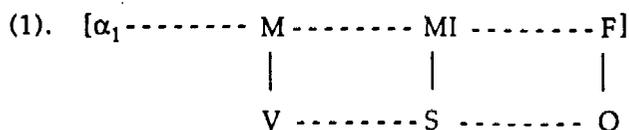
This paragraph can be translated as follows: "You should know that preposing is of two kinds: preposing which is called 'preposing with the intention of postposing,' and this may occur in all structures where the preposing will not change the structural characteristics. This can be the predicate of topicalization, or the NP-object, as for example, muntaliqun Zaydun [a departee is Zayd] or ḏaraba ʿamran Zaydun [Zayd

#### 4 The Internal Semantic Structures of the Basic Sentence

In the previous sections, I have tried to show that the Arabic sentence can be represented adequately if we analyzed it within the theoretical and terminological framework of the Arabic language itself. Such a framework can account for the whole body of syntactic and semantic data of the Arabic sentence. Trying to adjust the principles of this traditional framework within the new linguistic framework, we found that the logical structure of the Arabic theory can meet the logical structure of Chomsky's syntactic theory and Cook's semantic theory with some modifications which we were forced to make when converting and fusing both logical structures.

In this section, however, I shall explain the semantic and functional structures which the basic sentence and its transformations can reveal, and I shall explain the semantic nature of the logical structure in the Arabic theory.

As seen before, the most important categories of the Arabic structure are M and MI. When the structural relation is established between these two essential categories by the domination of the IS category, the structure will be flexible to accept extra syntactic and semantic categories, i.e., F. The syntactic structure of such essential and extra categories have a systematic word order which represents two underlying semantic representations for the sentential verbal structure in (1) and for the sentential nominal structure as in (2).



These main word orders which result in two semantic structures were emphasized by the majority of the Arab Grammarians. They also allowed these two systematic word orders to be flexible and exhibit varieties of transformational structures of semantic and functional roles. As a matter of fact, the semantic and functional roles which these transformational structures reveal were not clear enough in the work of the Arab Grammarians because they were interested in a purely syntactic analysis of the Arabic language. They were exactly similar in this respect to the early syntactic work of Chomsky (1957). It was the semanticists or those who were called in the Arabic tradition "Rhetoricians" who explained exhaustively and elaborately the functional aspects of transformational structures of the basic sentence. ?aj-jurjāni, for example, devoted most of his life to showing the semantic and functional genius and excellence of the Arabic structures revealed in Qur?ān. He explained, among many things, the structural and functional flexibility of the Arabic language which are a result of Move- $\alpha$ . ?aj-jurjāni explained the syntactic phenomenon of preposing and postposing constituents to the right or to the left of the verb. This phenomenon reveals the systematic aspects of the semantic and functional roles which are very similar to those functional roles in Dik's (1978) theory. According to Dik (1978), the most important constituent, which one is concerned with, is the preposed constituent in certain structures. This criterion was used by ?aj-jurjāni who proposed that the constituents preposed in certain structures are those the person bears upon or wants to know about more as we shall see.

The structure of the nominal sentence can be given as in (4).

- (4). [ $\alpha_2$  - NP<sub>i</sub> - (V) - NP<sub>j</sub> - X'']

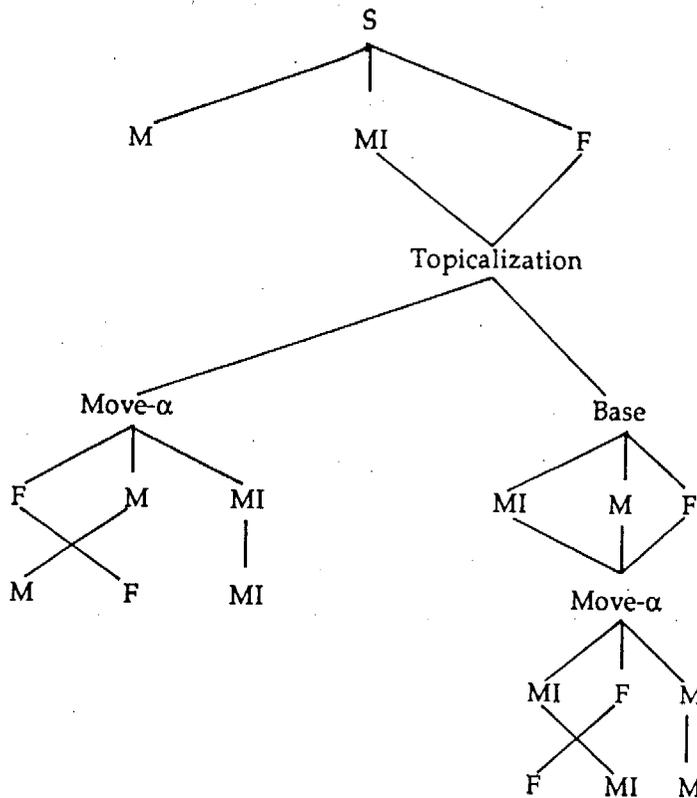
This structure in 4 which accounts for the nominal complex and equational sentence is subject to Move- $\alpha$ . The second type of Arabic sentence is "verbal," i.e., every structure starting with M followed by the MI subject and one of the categories under X''. The structure of the verbal sentence can be given as in (5).

- (5). [ $\alpha_1$  - V - NP<sub>i</sub> - (NP<sub>j</sub>) - (X'')]

The NP<sub>j</sub> - object and X'' are optional constituents but the verb and NP<sub>i</sub> -subject are obligatory. This structure is also subject of Move- $\alpha$ .

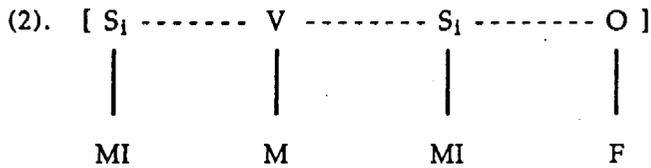
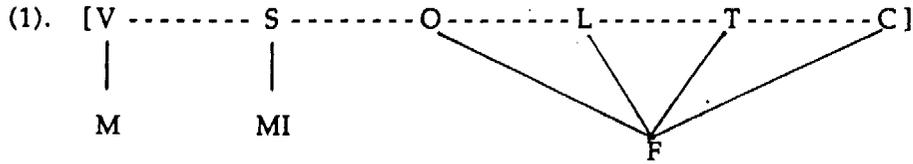
Assuming these facts about the Arabic sentence, we notice that the NP-subject is generated by the base either under the Topic-node in the case of a nominal sentence, or under the S-node in the case of a verbal sentence. The other constituents, however, either in nominal or verbal structures can move transformationally according to Move- $\alpha$ . Functionally, however, as Arab Grammarians had stated, there was no difference between the sentences with these constituents preposed to the left of the verb or to the right of it either transformationally or by the base, because they serve one pragmatic and semantic role, i.e., "Topicalization." The process of preposing and postposing constituents in the basic structure of the Arabic sentence can be shown in Figure 6: Preposing and Postposing in the Arabic sentence.

FIGURE 6



### 3. The Internal Syntactic Structures of the Basic Sentence

This section will focus on the basic order of the Arabic sentence which is generated by the base, and it will show the possible derived structures which are the result of Move -  $\alpha$  principles. After that this section will explain the structural and functional aspects of such moving elements. Arab Grammarians considered the internal structure of the basic sentence to be in the following order:



All structures which differ from (1) and (2) are considered structures derived by Move -  $\alpha$  transformationally. These basic structures of order were stated by the majority of the Arab Grammarians, particularly Ibn yaʿīš (d. 1250) who stated the order of the basic sentence as follows:

Faʿiḍān rutbatu l - fiʿli yajibu ʿan yakūna ʿawwalan wa rutbatu  
l - fāʿil ʿan yakūna baʿda - hu wa rutbatu l - mafʿūli ʿan yakūna  
ʿāxiran wa qad taqaddama l - mafʿūl liḅarbin mina t - tawassuʿi  
wa l - ʿihtimāmi bi - hi wa n - niyyatu bi - hi t - taʿxīr .

(Ibn yaʿīš; ʿarḥ ʿal-mufaṣṣal 1, p. 76)

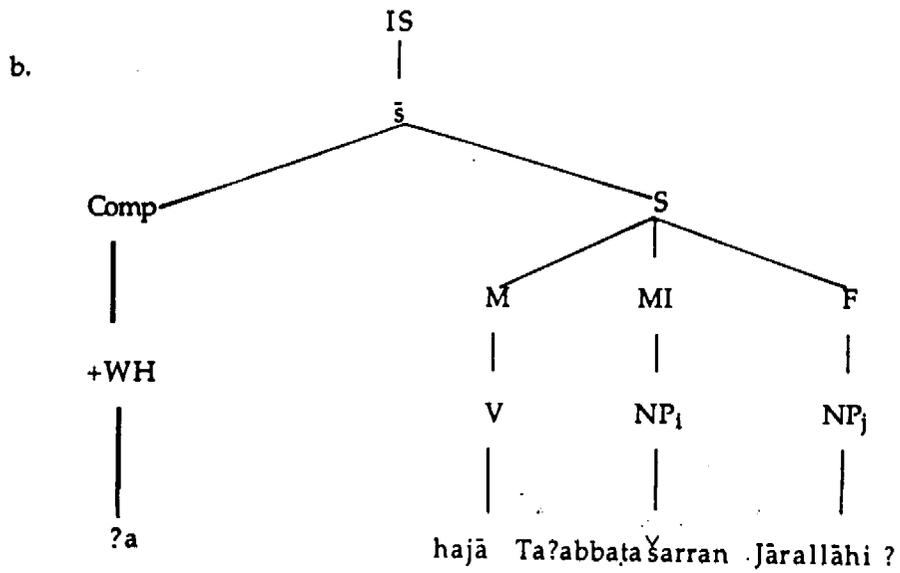
This paragraph can be translated as follows: "So, the verb must appear in the first position, the subject after the verb, and the object must come finally. The object is preposed when it is focused. The focused object is, however, understood as being in its postverbal position."

As a matter of fact, even the basic structure and its derived structures consist of the same constituents; they differ functionally or pragmatically when these constituents are preposed to the left or postposed to the right of the verb.

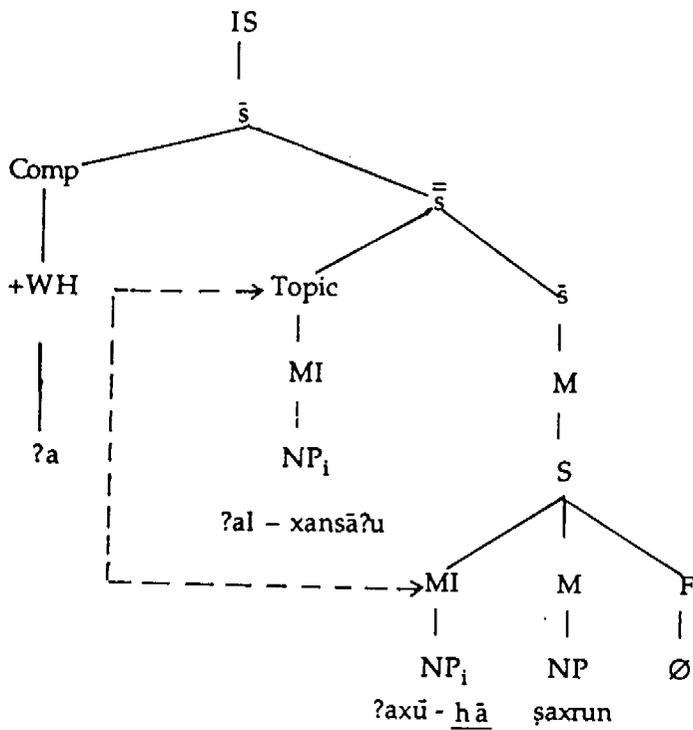
As seen before, the Arabic basic sentence has two types of structures both of which are generated in the base. The first type is the "nominal structure," i.e., every structure starting with MI as a theme or topic, followed by either sentential rheme or one of the following categories which collapse under the category X<sup>'''</sup>.



- (12). a.  $\frac{?a}{Q}$  hajā Ta?abbataŸarran Jārallāhi  
 satirized Ta?abbataŸarran Jārallāhi  
 Did Ta?abbataŸarran satirize Jārallāhi ?



- (13). a.  $\frac{?a}{Q}$  l - xansā?u ?axū - hā şaxrun  
 ?al - xansā?u brother - her şaxrun  
 As for ?al-xansā?, is Şaxr her brother?



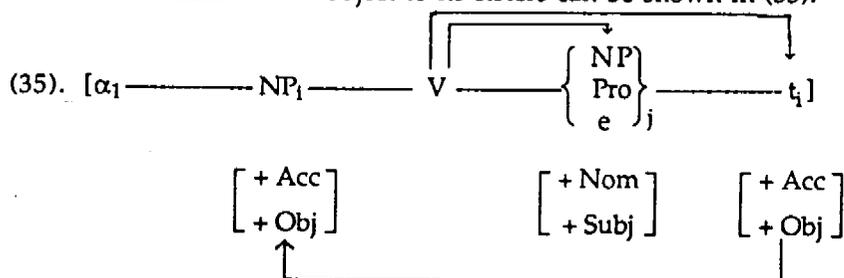
of the language which the Arab Grammarians give us has been almost totally neglected by Western linguists. I would like to think that I am approaching this work in the spirit of the Arab Grammarians. At least this is true in one regard--the problem that so enthused the Arab Grammarians, that of determining the *ʔaṣl* or deep representation of the language" (Brame 1970, MIT Ph. D., p.v). Chomsky (1982) expressed the same fact, but in a different way when he stated that: "... I was quite interested in the tradition of Arabic and Hebrew grammar of the Medieval period.... My father was a specialist in Medieval Hebrew and Arabic grammar, and I studied it with him... Much of my own thinking about language actually was influenced by some of that work.... Some of my early studies in Medieval grammar led to some ideas about rule system which then entered into my work on generative honology and language systems" (Al-wear 1982): *ʔal-lisāniyyāt* , Vol. 6, ʔ6). But if "Arabic Grammar has reached its lowest ebb under the thumbs of Western scholars" it has been attacked by its native speakers who would like to imitate and impose whatever is new linguistically on a certain portion of the Arabic syntax and neglecting the other portions just to meet and justify the principles of the approach or model they adopted. Strictly speaking the imposing of a certain model, which is a result of a certain data on another data which is different, on a philosophical ground, is not objective scientifically (Boas 1964 and Dinneen 1967).

It seems, then, that it would be more sufficient to adopt the logical structure of the Arabic language itself. But I shall also try to incorporate such a logical structure with what can be universal in the logical structure of Chomsky's syntactic theory and Cook's semantic theory. In doing so, I believe we not only capture the adequate nature of the Arabic basic sentence, but we enrich the general linguistic theory as well. Holding these assumptions, the syntactic rules of the basic sentence would be the following.

- |                      |        |   |
|----------------------|--------|---|
| (1). IS              | —————> | $\bar{s}$   |
| (2). $\bar{s}$       | —————> | Comp $\left\{ \begin{array}{l} s \\ \bar{s} \end{array} \right\}$   |
| (3). $\bar{\bar{s}}$ | —————> | Topic - $\bar{s}$   |
| (4). Comp            | —————> | $\pm$ WH  |
| (5). + WH            | —————> | <i>ʔa</i> , Hal; lima . . . .   |
| (6). - WH            | —————> | <i>ʔ inna</i> , <i>ʔan</i> , Kay . . . .  |
| (7). S               | —————> | M - MI - (F)  |
| (8). M               | —————> | V - (NP) (AP) (PP) (AdvP) (VN) (S)  |
| (9). MI              | —————> | NP - (VN) (S)   |
| (10). F              | —————> | NP - (O) (L) (T) (M) (I) (P) (C) (R)  |
| (11). MI             | —————> | NP $\left\{ \begin{array}{l} (A) \text{ gent} \\ (Os) \text{ object of state} \\ (E) \text{ xperiencer} \\ (B) \text{enefactive} \end{array} \right\} \#$ |

These rules are capable of generating both types of Arabic sentences, i.e., nominal and verbal. In addition, they are capable of differentiating between different lexical categories which are in binary set features. Applying such rules to the basic sentence in Arabic, we can have the following configurational structures as in (12) and (13).

Those syntactic relations, however, are different from that holding in the NP-object as we have seen. Here the NP-object moves freely and is not bound to a particular pronoun because it is not generated under the Top-node, but under the S-node and it is transformationally moved, according to move- $\alpha$ . The movement here is to a sister-adjoined position. The relation of NP-object to its sisters can be shown in (35).

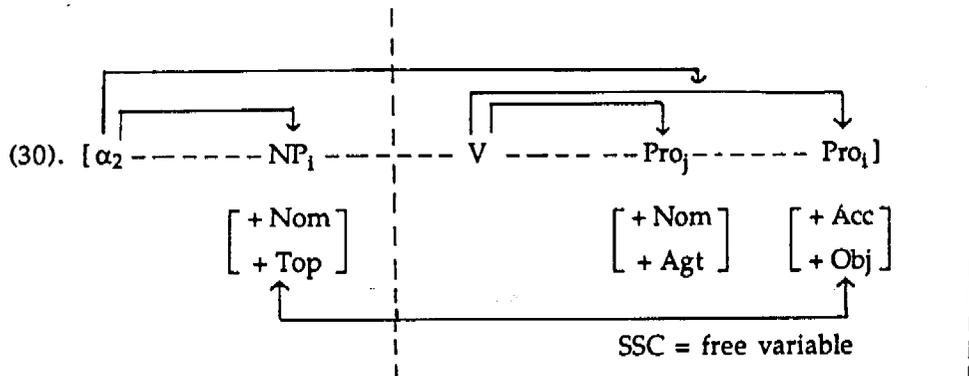


## 2.2. Toward a Realistic Theoretical Framework

In the last section we saw that the basic Arabic sentence data fit some contemporary theoretical frameworks on the one hand and deviate from these frameworks on the other. However, we can try somewhat to fuse and incorporate the rules of the Arabic framework with those of Chomsky's in order to describe and explain the internal structure of the Arabic sentence in a very adequate and objective way. This is possible provided that we may incorporate the logical structure of the Arabic rules within the logical structure of Chomsky's rules, i.e., to convert the rule system proposed by the early Arab Grammarians in certain rule systems proposed by Chomsky (1977, 1978, and 1981).

This proposal, however, might be objected on the grounds that the logical structure of linguistic theory of the early Arab Grammarians is different in its assumptions and its philosophical and dimensional nature, from the nature of the logical structure of Chomsky (Joseph Aoun 1981). The argument against this objection is that the logical structure of the Arabic theory was built and proposed on the grounds of different and rich linguistic data which were available to Arab Grammarians in their era. They investigated a large body of data from foreign languages through the process of translation and its scholastic principles especially the process of translation which was established in what is known as dār ?al-?ikmah , i.e., "the house of wisdom," a code for a governmental organization for translation from different languages. This was in the eighth century A.D. In addition, Arab Grammarians did not use strange linguistic data which we cannot understand nowadays. They studied the Arabic language and its dialects which we are using today, even though some morphological changes have taken place. But from a purely scientific point of view they did exactly what we are presently doing, i.e., they used an exact, objective, and verified methodology in establishing the logical structure of the Arabic theory. For example, they travelled from place to place listening to adults as well as to children to know the exact linguistic knowledge of the speaker-hearer of his language in order to formalize a very adequate logical structure of the Arabic language. The purpose of their investigation was to capture the representations of the computational and abstract relations of the linguistic system in the human mind (?aj-jurjāni. Dalā?il ?al-?i?jāz). The motivation for such investigation was in Brame's (1970) word, to discover the "deep representation" of the language. I can add that by discovering the "deep representation" of the language they could prove the high esteem and excellence of the Qur?ānic Structure to which they devoted most of their lives.

These facts, however, were recognized by two Western objective linguists: Michael Brame (1970) and Noam Chomsky (1982). Brame, in his MIT dissertation (1970), stated the following: "It is my belief that Arabic Grammar in particular has reached its lowest ebb under the thumbs of Western scholars. Much of the subtlety and insight into the nature



In some nominal structures, however, we find that the NP-subject or Topic has no overt resumptive element and yet the sentence is still grammatical. In the logical structure of the Arabic theory it was assumed that the resumptive element is bound and coindexed to its antecedent (NP-Topic) either overtly as we have seen or covertly, i.e., the resumptive element is not showing in the surface structure but it is present in the underlying structure. The covert resumptive element was called by Arab Grammarians ḍamīr mustatir "a hidden pronoun."

The overt resumptive element was called ḍamīr Qāhir, i.e., "appearing pronoun." The covert or hidden resumptive pronoun can be shown in the following examples.

- (31). a. Hindun waṣalat ?ilā Dimašqa  
 Hind arrived to Damascus  
 As for Hind, she arrived in Damascus.

- b. Hindun<sub>i</sub> waṣalat  $\left\{ \begin{array}{c} e \\ hiya \end{array} \right\}_i$  ?ilā Dimašqa  
 ec

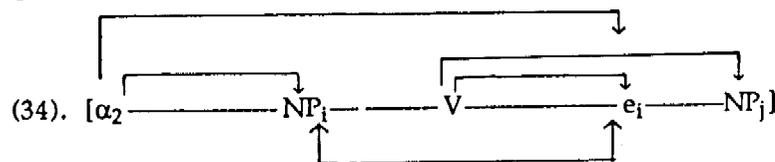
- (32). a. ?al-mūsayāni yūḡanniyāni šīʿran  
 The two moseses sing poetry  
 As for the two moseses, they are singing poetry.

- b. ?al-mūsayāni<sub>i</sub> yūḡanniyāni  $\left\{ \begin{array}{c} e \\ humā \end{array} \right\}_i$  šīʿran  
 ec

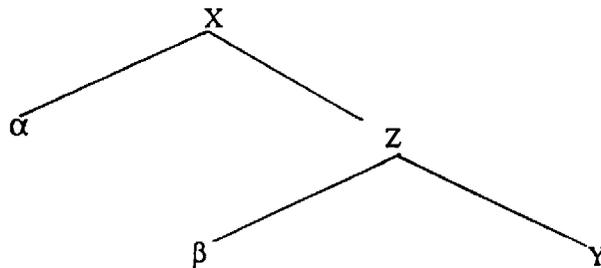
- (33). a. ?al-Mayyātu l - ḥasnāwātu yarmīna - hū bi - tṭarfi  
 The Mayys the beautiful blink - him with eyelashes  
 As for the beautiful Mayys they blink at him with their eyelashes.

- b. ?al-Mayyātu<sub>i</sub> l - ḥasnāwātu yarmīna - hu  $\left\{ \begin{array}{c} e \\ hunna \end{array} \right\}_i$  bi - tṭarfi  
 ec

In (31b), (32b), and (33b) we can see that the resumptive pronoun is covertly hidden, and has become an empty category (ec) which is coindexed with its NP-Topic antecedent. The anaphoric relation in such cases could be represented by this rule:

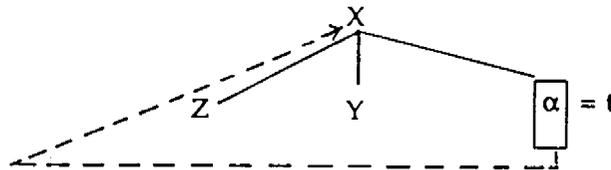


(26).

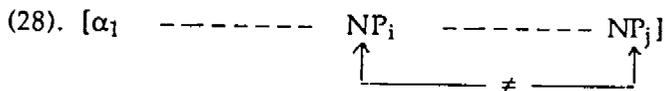


$\alpha$  in (26) C-commands  $\beta$ , but not vice versa. This rule of C-command is very applicable to the NP-subject in the nominal sentence, but not to the NP-object in the verbal sentence. The rule which can be applied to NP-object is move- $\alpha$  to its sister-adjoined element as in the process diagrammed in (27).

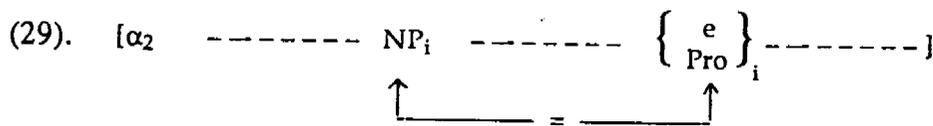
(27).



The movement of  $\alpha$  as in (27) is restricted to the constituent which is dominated by X. This process however, is different from the Wh movement of Chomsky (1977). The movement of the NP-object in the Arabic sentence carries with it the case marker, i.e., the accusative marker. In addition it carries the case role or  $\theta$ -role which the verb assigned it. As a matter of fact, the movement of the NP-object in the logical structure of the Arabic theory can meet the principle of Lasnik (1976). Lasnik's principles allow the two NPs in the verbal sentence to move freely because they are not coindexed, i.e., they have disjoint reference. This relation can be shown in this rule.

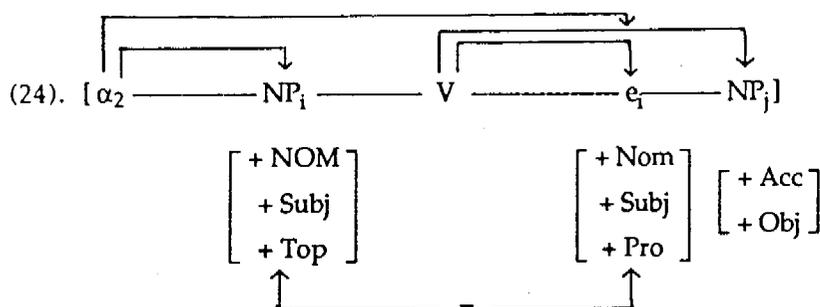


Whereas this rule applies adequately to the verbal structure, the NP-subject of the nominal structure does not meet Lasnik's principle because the NP-subject cannot be free here. It must be bound to its antecedent and be coreferential with it. The relation holding between them can be seen in (29).

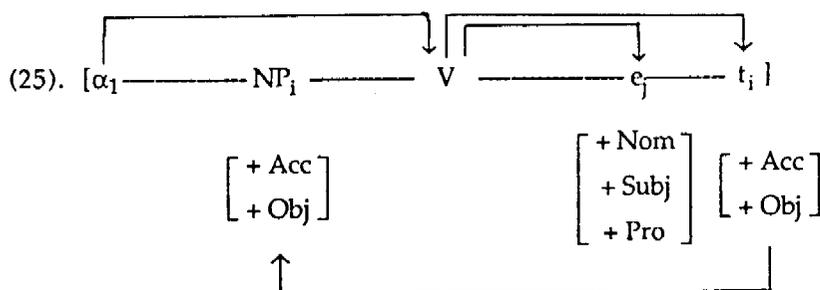


This kind of syntactic process in (29) applies within the framework of Chomsky (1977) but not in an adequate way. In the case of the Arabic NP-subject, the antecedent Topic or Theme must be bound to the full or empty element which coindexes with the NP-Topic, but the sentential rheme or comment does not have to obey the Tensed Island Constraint (TIC) or the Specified Subject Constraint (SSC), which were proposed by Chomsky (1978) because the sentential rheme or comment in Arabic has the SSC which is tensed. This kind of relation enables the sentential comment to be a free variable in Arabic in spite of the fact that the resumptive element which is in the sentential comment coindexes with the NP-subject antecedent, or Topic. This relation can be shown in the following configuration.

These two rules are universal as proposed by Chomsky. They apply to English as well as to other languages. Applying these rules to the basic sentence in English we can see that the Topic applies to the NP-subject and to the NP-object. The two NPs are base-generated according to these rules. In Arabic, however, there is a distinction between the NP-subject which meets exactly Chomsky's topicalization, and the NP-object which disagrees with Chomsky's topicalization. According to the logical structure of the Arabic theory, the topicalization of Chomsky can only apply to the complex nominal sentence whose structure is as in (24).



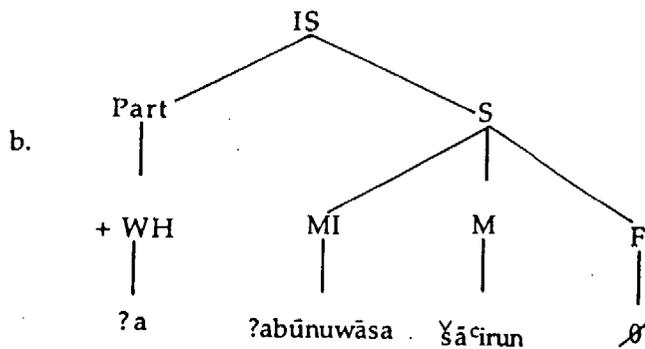
The NP-Topic in (24) is generated in the base under the node Topic. The other constituents in (24) are considered to be sentential comments which are dominated by the S node. The NP-object in the verbal structure, however, is different from the NP-subject because it occurs in a different position and develops a different process. The structure of the NP-object can be given in (25).



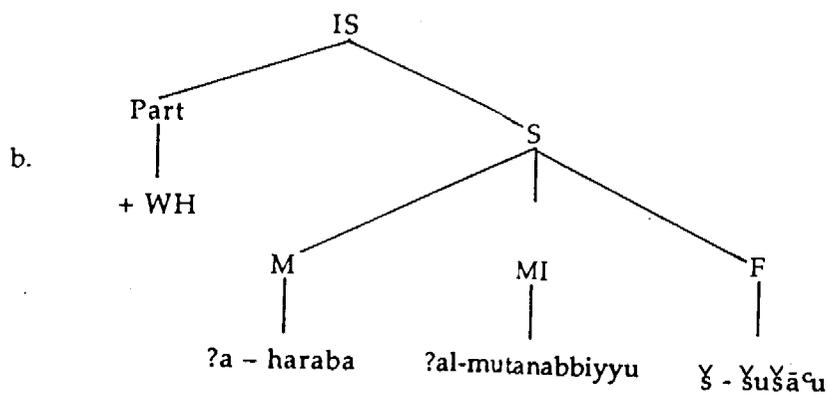
As seen in (25) the NP-object is not generated in place (as in (24)) in the base but is moved transformationally to the front of the structure leaving a trace behind. This means that the NP-object cannot meet Chomsky's rule of topicalization. Arab Grammarians however, agreed with Chomsky on move- $\alpha$  principles by which the NP-object can move transformationally within the node ( $\bar{s}$ ) where the NP-object is considered a sister-adjoined to the verb. This means then that the NP-object cannot move transformationally to a higher node to be a sister-adjoined to ( $\bar{s}$ ) which is dominated by ( $\bar{s}$ ).

Thus the principle of C-command proposed by Chomsky (1977-1978) can only apply to NP-Topic which is generated in the base in Arabic. The general principle of C-command is that  $\alpha$  C-commands  $\beta$  if  $\alpha$  does not contain  $\beta$  and  $\beta$  is dominated by the first branching category dominating  $\alpha$ . The C-command principle can be shown in the following configuration.

- (20). a.  $\frac{?a}{Q} \frac{?abūnuwāsa}{?abūnuwāsa} \frac{Yā'irun}{poet}$   
 Is ?abūnuwās a poet?



- (21). a.  $\frac{?a}{Q} \frac{haraba}{escaped} \frac{l-mutanabbiyyu}{?al-mutanabbiyy} \frac{Y - YūYā'ū}{the courageous}$   
 Did the courageous Mutanabbiyy escaped ?



As a matter of fact, the logical form of the sentential Arabic theory as indicated in the above configurations is similar to the logical form of Chomsky's (1977) REST. Chomsky (1977) proposed new base-generated rules which can account for different syntactic modifications that had been made. The new rules were the following :

(22).  $\bar{s} \longrightarrow \text{Comp} - \left\{ \begin{array}{c} s \\ \bar{s} \end{array} \right\}$

(23).  $\bar{\bar{s}} \longrightarrow \text{Topic} - \bar{s}$

This configurational fact was expressed by Sibawayhi and the majority of the Arab Grammarians who came after him, and it was elaborately explained especially in ?al-?istrābāōi (d.1293) and Ibn Hišām (d. 1368). Let us consider some examples which clarify this theoretical issue.

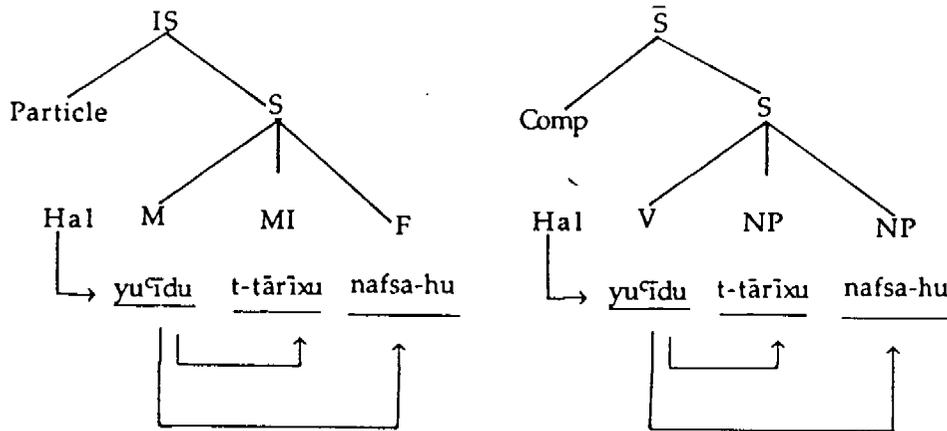
- (18). a. 
$$\begin{array}{c} \text{Part} \\ \hline \text{?a} \\ \hline \text{Q} \end{array} \quad \begin{array}{c} \text{S} \\ \hline \text{Zaydun} \quad \text{?axū - ka} \\ \hline \text{Zayd} \quad \text{brother - your} \end{array}$$
  
Is Zayd your brother?
- b. 
$$\begin{array}{c} \text{Part} \\ \hline \text{Mā} \\ \hline \text{Neg} \end{array} \quad \begin{array}{c} \text{S} \\ \hline \text{Zaydun} \quad \text{šā'iran} \\ \hline \text{Zayd} \quad \text{poet} \end{array}$$
  
Zayd is not a poet.
- c. 
$$\begin{array}{c} \text{Part} \\ \hline \text{?a} \\ \hline \text{Q} \end{array} \quad \begin{array}{c} \text{S} \\ \hline \text{Musāfirāni} \quad \text{?al-ʿisayāni} \\ \hline \text{traveller} \quad \text{?al-ʿisayāni} \end{array}$$
  
Are ?al-ʿisayāni travellers?
- (19). a. 
$$\begin{array}{c} \text{Part} \\ \hline \text{?a} \\ \hline \text{Q} \end{array} \quad \begin{array}{c} \text{S} \\ \hline \text{tuḡannī} \quad \text{Mayyun} \\ \hline \text{sing} \quad \text{Mayy} \end{array}$$
  
Is Mayy singing?
- b. 
$$\begin{array}{c} \text{Part} \\ \hline \text{Mā} \\ \hline \text{Neg} \end{array} \quad \begin{array}{c} \text{S} \\ \hline \text{qāla} \quad \text{š-šī'ra} \quad \text{Zaydun} \\ \hline \text{say} \quad \text{poetry} \quad \text{Zayd} \end{array}$$
  
Zayd never versified poetry.
- c. 
$$\begin{array}{c} \text{Part} \\ \hline \text{Hallā} \\ \hline \text{Q-like} \end{array} \quad \begin{array}{c} \text{S} \\ \hline \text{?anšadtani} \quad \text{šī'ran} \quad \text{yā} \quad \text{ḡulām} \\ \hline \text{recite-me} \quad \text{poetry} \quad \text{oh} \quad \text{boy} \end{array}$$
  
Oh boy, would you recite me poetry?

We noticed from the above examples that the nominal structures in (18) and the verbal structures in (19) consist of particles in Comp which are sister-adjoined to the node S but never to the constituents that are dominated by S, because the essential constituents hold only among M, MI and F which are sister-adjoined dominated by the node S. The node (Part) however is an extra syntactic category which is a sister-adjoined to the node S. The node Part and S are dominated by a higher node which organizes the whole configurational process of the complete structure, i.e, the IS node. The configurational structures of the sentences (20a) and (21a) can be shown in (20b) and (21b).

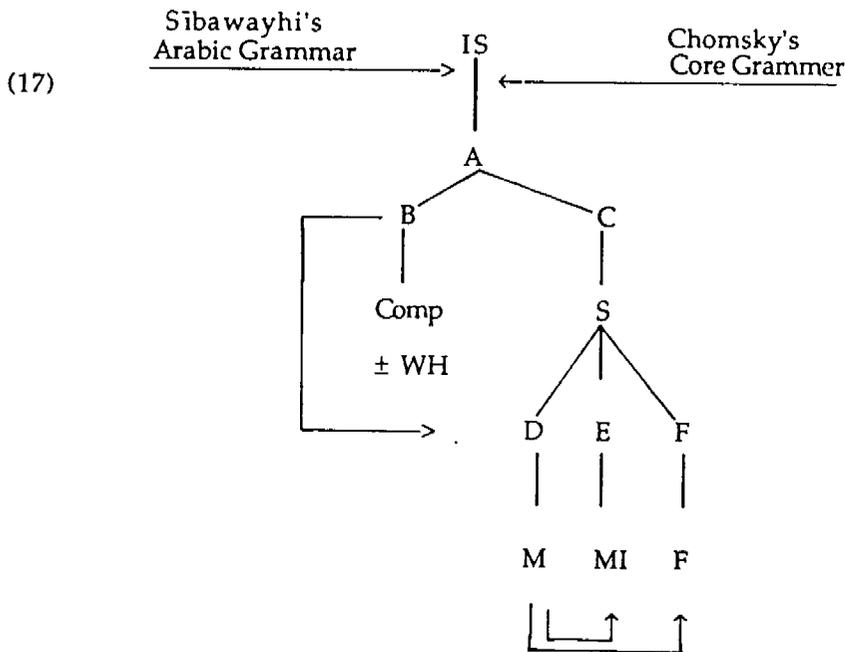
All particles which might modify the sentence do not effect the essential structure of the node S. As a matter of fact, Arab Grammarians considered what is called in REST as "Comp," to be beyond the domination of the essential structure of the sentence, i.e., the essential constituents are not dominated by the same node which dominates Comp or particle. This means that the logical structure of the Arab Grammarians is very similar to the logical structure of Chomsky (1977). This similarity can be seen by the comparison between the two logical structures:

(15) Sībawayhi (d.793)

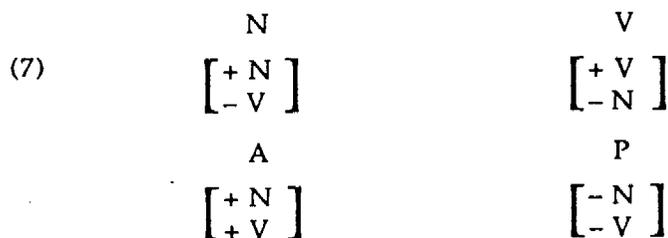
(16) Chomsky (1977)



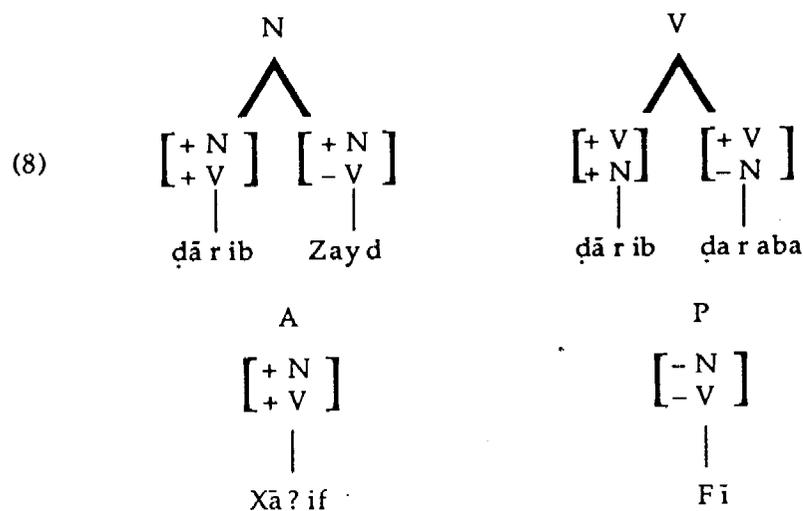
As seen above, the two configurations of Sībawayhi (15) and Chomsky (16) represent one abstract structure which can account for Arabic as well as for other languages. The more adequate configuration which represents the underlying structure as understood by Sībawayhi and Chomsky would be as in (17).



Applying a strict transformational approach to the Arabic data, particularly the lexicalist theory of Chomsky (1970), will indicate some facts about Arabic lexical nature which cannot meet exactly Chomsky's assumption about the lexical hypothesis. Chomsky (1970) classified the major lexical categories by what is known as "Binary sets of features". This means that the general domain of these features can capture the nature of more than one syntactic category of the lexical rules. The major binary set can be shown in the following way.

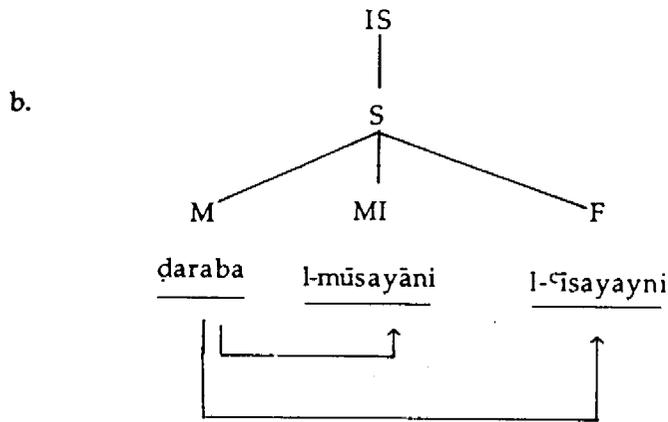


The binary set of the Arabic data is different slightly from that of Chomsky's because of the derivational nature of the Arabic language which enables the lexical category to share another lexical category with some syntactic and semantic properties. The binary set of the lexical category in Arabic can be shown in (8).

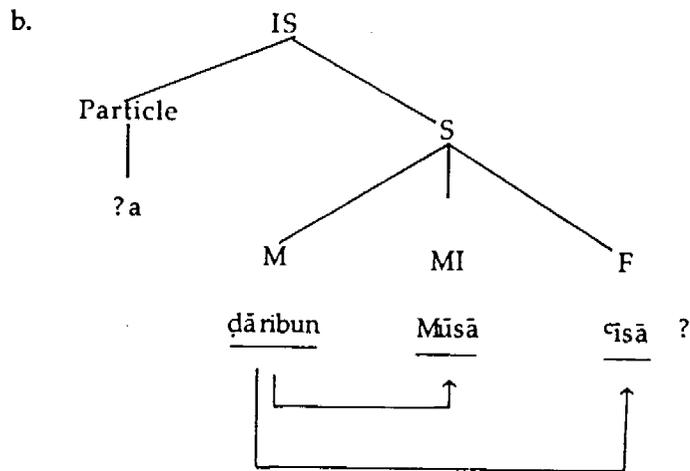


Formalizing rules based on these categories one can see that the rules which can account for the Arabic data would be of the following nature.

- (9) IS  $\longrightarrow$  S
- (10) S  $\longrightarrow$  M - MI - (F)
- (11) M  $\longrightarrow$  V - (NP) (AP) (PP) (AdvP) (VN) (S)
- (12) MI  $\longrightarrow$  NP - (VN) (S)
- (13) F  $\longrightarrow$  NP
- (14) NP  $\longrightarrow$  (A)gent-(O)bject-(E)xperiencer-(B)enefactive  
(L)ocative-(T)ime-(M)anner-(P)urpose  
(I)nstrument-(C)ause-(R)esult.



- (6). a. ʾa    ḍāribun    Mūsā    ʿisā ?  
 is        hitter        Moses    ʿisā  
 Is Moses the hitter of ʿisā?



As seen in the configurations above, the predicate M in (5b) is a verb, but it is a verbal noun which is derived from the verb in (6b). This means that the VN can function exactly as if it were a verb, i.e., it can occur in the position of the verb and inherit its syntactic and semantic governing roles. It requires, as in (6b), an NP-agent and an NP-object and assigns them case roles and case markers. As we have seen before, the (VN) as a governing element can be either an active or passive which governs the active and passive sentence. As a matter of fact, there is a large body of Arabic morphology which can express different functions depending on the constituents they inherited from the verb. Arab Grammarians talked about three categories which function as if they were verbs. The first is called *ʾismu l-fāʿil*, i.e., active verbal noun (AVN) which occurs in the active sentence and governs it. The second is called *ʾismu l-mafʿūl*, i.e., the passive verbal noun (PVN) which occurs in the passive sentence and governs it. The third is called *ʾaṣ-ṣifatu l-muṣabbahatu bi-ʾismi l-fāʿil*, i.e., the verbal adjective which is similar to the verbal noun (AdjVN) which occurs in an adjective position to modify a preposed noun and to govern at the same time a postposed noun. I shall consider all these categories to collapse under the VN category.

$$(3). V' \longrightarrow V - N''' - (N''') - (P''')$$

The most recent theoretical framework of the Arabic sentence was proposed by Fihri (1981). Fihri borrowed a more developed theoretical framework in TGG which is called "A Lexical Functional Grammar" proposed by Bresnan (1976-1982). Fihri tried to apply the constituent structure (C-structure) of this theoretical framework to the basic sentence in Arabic. His understanding of the basic structure is similar to that of Bakir's (1980), except that Bakir's framework was Jackendoff's  $\bar{x}$ -theory. The "Lexical Functional Grammar" framework resulted in the following rule.

$$(4). S \longrightarrow \begin{array}{c} V \\ \uparrow=\downarrow \end{array} - \begin{array}{c} NP \\ (\uparrow\text{SUBJ})=\downarrow \end{array} - \begin{array}{c} NP \\ (\uparrow\text{OBJ})=\downarrow \end{array}$$

As a matter of fact these theoretical frameworks of the Arabic basic sentence can account for some Arabic data, but they cannot capture other data since they cannot describe the counterexamples of the sentential structures in Arabic. An adequate framework is necessary to describe and explain the Arabic data from a purely Arabic perspective point of view. After achieving such description and explanation, one can incorporate the Arabic theoretical framework with that of the modern framework. From this point of view one can have an adequate justification not only for Arabic data but for the general linguistic theory as well. I shall present here the basic syntactic and semantic assumptions of the underlying structure of the Arabic sentence from a purely Arabic terminology. After that I shall raise the possibility of fusing and converting such syntactic and semantic assumptions in the modern theoretical framework in order to capture the essential and adequate structures of the Arabic sentence.

The structure of the Arabic sentence consists of three constituents, two are essential, the third is peripheral. The first essential constituent is called *musnad* (M), i.e., the predicate of the sentence regardless of the syntactic nature of that predicate which might be verb (V), a verbal noun (VN) or a sentence (S). The second essential constituent is called *musnad ?ilay-hi* (MI), i.e., the argument or the subject which the predicate describes and depends on in conveying the message. The constituent MI can be different syntactic categories such as noun phrase (NP), prepositional phrase (PP), adjective phrase (AP), adverbial phrase (AdvP), verbal noun (VN), and sentence (S), but never a verb. The third peripheral or extra constituent is called *Faḍlah* (F), i.e., all constituents which are neither M nor MI. F enters the structure as an extra constituent which contributes to the meaning of the sentence and deepens it. The relation which holds among these structural constituents is called *?isnād* (IS), i.e., configurational predication which governs the sentential constituents and dominates them (?al-?istrābāḍi; ?arḥ ?al-Kāfiyah, pp. 8-22).

Applying this theoretical framework to the basic sentence in Arabic we can understand the logical form of the basic sentence as in (5) and (6).

$$(5). \quad a. \quad \begin{array}{c} \text{ḍaraba} \\ \text{hit} \end{array} \quad \begin{array}{c} \text{l-mūsayāni} \\ \text{the two moseses} \end{array} \quad \begin{array}{c} \text{l-ʿisayayni} \\ \text{the two ʿisas} \end{array}$$

The two moseses hit the two ʿisas.

framework. The new and realistic framework would be more comprehensive and adequate in capturing the syntactic and semantic domains of the Arabic sentence.

The new theoretical alternative will be tested through the internal syntactic investigation of the word order in the verbal and nominal sentences. It will be tested also through the investigation of the possible transformational constituents within a particular structure. The freer and more restricted movement of the constituents will be studied through the move- $\alpha$  principles proposed by Chomsky (1981). In addition, the new theoretical alternative will be tested through the study of the semantic structures in the verbal and nominal sentences. It will be shown that such an alternative will capture the exact nature of move- $\alpha$  which result in different semantic structures.

## 2. The Theoretical Frameworks of the Basic Sentence

What I mean by the theoretical frameworks of the basic sentence are the methods and the analyses which are used by contemporary Arab and Western linguists to investigate the basic structures of the Arabic sentence. What I shall do here is to explain the perceptions of some Arab and Western linguists of these structures and show afterwards that there is need for a realistic understanding of the Arabic sentence within both a traditional and a contemporary scope.

### 2.1. The Contemporary Arabic and Western Framework

The contemporary theoretical framework of the basic sentence in Arabic varies from one linguist to another depending on the theory which each linguist adapts. Thus one who tries to investigate the theoretical framework of the Arabic sentence will face different proposals concerning the configurational structure and its rules that capture the syntactic and semantic domains of the Arabic sentence. The configurational structure and the rules that account for the Arabic sentence come from two sources: the first was proposed by Western linguists who conceived of the structural framework of the Arabic sentence from their modern linguistic background. The second was proposed by Arab linguists who understood the structure of the Arabic sentence through a particular modern linguistic approach. The problem in both sources of analysis is that they approached the immediate and more applicable data which meet the theoretical principles of the theory they adapted, and neglected a large data which cannot meet the principles of their theory.

Snow (1965), Killean (1966), Lewkowicz (1967), and Awwad (1973) for example, believed the structure of the Arabic basic sentence to consist of the following rule.

$$(1). \quad S \longrightarrow \left\{ \begin{array}{l} NP + VP \\ NP + Pred P \end{array} \right\}$$

Other linguists, however, conceived of the Arabic basic sentence differently. Anshen and Schreiber (1968) understood the structure of the basic sentence to consist of the following rule:

$$(2). \quad \longrightarrow VP + NP$$

More recently, some Arab linguists deviated from the above structures trying to benefit from more developed theoretical issues in linguistic theory. Bakir (1980) for example, tried to apply Jackendoff's  $\bar{x}$ -theory to the basic sentence in Arabic. In addition, he followed the theoretical principles of Chomsky's (1977) Revised Extended Standard Theory (REST) and applied it strictly to Arabic sentences. The application of  $\bar{x}$ -theory and REST to the basic structure of Arabic sentences resulted in the following rule.