

ON THE SYNTACTIC CONFIGURATION
OF ARABIC TOPIC AND FOCUS
DISCOURSE FUNCTIONS

by

Kamel Abdelbadie Elsaadany

This paper investigates the TOPIC and FOCUS functional discourse constructions in Arabic. It answers these questions: 1) how should Arabic topicalized/focalized constructions be syntactically/functionally defined? 2) How should Arabic TOPIC and FOCUS constructions be configured via the LFG Functional Uncertainty (FU) apparatus? 3) To what extent does the LFG framework suffice to account for the TOPIC and FOCUS functional constructions in Arabic? And 4) how could the incorporation of i-structure configuration help defining the FOCUS function and overcome the scoping problems that existed in the topicalization of the PRED of the matrix f-structures? The study provides a non-derivational analysis of the syntax-discourse interface of Arabic TOPIC and FOCUS constructions through integrating both the LFG-FU and the i-structure. The suggested incorporation of an i-structure for the FOCUS function proves that the [+New] feature is what licenses FOCUS and considers FOCUS a topicalized construction that contains some [+New] information. Because of the occurring mismatches of TOPIC and FOCUS in the f-structure, the adoption of information framework provides an independent i-structure that is responsible for representing information that defines the elements in the FOCUS that are [+New] and the ones that are [-New]. The study concludes that the incorporation of both the f-structure and i-structure in the suggested functional/informational framework has solved many of the resulted scoping semantic/discoursal problems that existed in the topicalization of the PRED of the matrix f-structures. The informational framework and its incorporated i-structure complement the f-structure, which includes an additional i-structure that independently captures the [+New] and [-New] features of the Arabic FOCUS functions.

1. *Introduction*

This study investigates Arabic TOPIC and FOCUS Discourse functions within the framework of Lexical Functional Grammar (LFG). The syntactic constructions of TOPIC have been widely studied in many different languages and from derivational approaches. Few studies have researched the phenomenon from a non-derivational framework such as LFG. Among the scholars who have tackled this viewpoint are Saddock (1991), Jackendoff (1987; 1997) and Bresnan and Kanerva (1989). Some other studies dealt with topic constructions in languages other than Arabic, such as Workman (2009), Speyer (2010), Mimura (2009), Fung (2007), and Webelhuth & Ackerman (1999).

The FOCUS structures have not been given much attention in the LFG framework. King (1997) is a major LFG study that has proposed the use of an i(nformation)-structure, which captures discourse function information in both the c(onstituent)-structure and s(ematic)-structure. Without such i-structure, the FOCUS function will cause scoping problems in the f(unction)-structure configuration. There are no studies that researched Arabic topicalization in general and FOCUS functions in particular. This current study will fill this gap and propose a non-derivational functional analysis for accounting for this phenomenon.

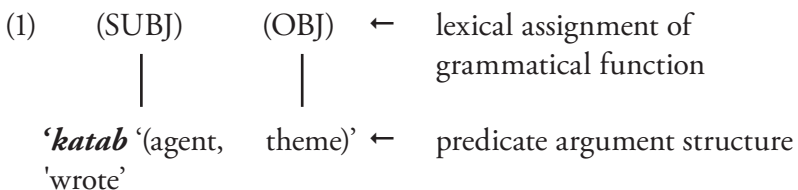
This paper is organized as follows: Section 2 provides a brief overview of the LFG syntactic framework. Section 3 discusses the notion of topicalization and Arabic TOPIC and FOCUS functional constructions. The LFG theoretical framework and its functional apparatus, Functional Uncertainty, for long distance dependency are discussed in Section 4. Section 5 provides a solution to the scoping mismatches in the f-structure of TOPIC and FOCUS in Arabic. Section 6 concludes and discusses the suggested solutions to the problem.

2. LFG framework formalisms

The LFG framework is a non-derivational and constraint-based grammar that uses parallel correspondence architecture (Bresnan 2001; Falk 2001). Contrary to Transformation Grammar, LFG does not assume any constituent movement; grammaticality, as Falk (2001) argues, is determined by satisfying certain constraints. The main parallel structures assumed in the LFG are the l(exical)-structure, c-structure, f-structure, and a(rgument)-structure. Scholars working in the LFG framework such as King (1997) also suggest other structures such as i-structure and s(ematic)-structure. The established connection among these different functional structures is maintained by some mapping principles. Bresnan (2001) and Falk (2001) provide detailed description of the c-structure, f-structure, and a-structure; King (1997) provides details of the i-structure and s-structure. The following section highlights these LFG structures.

2.1. Lexical structure (l-structure)

LFG lexical structures refer to the meanings, thematic structures, and grammatical functions of the lexical items. Examples of grammatical functions such as SUBJECT, OBJECT, PRED etc. are associated with their appropriate arguments. For example, the verb *katab* ‘wrote’ has a predicate argument structure which is made up of an agent argument associated with the SUBJ function and a theme argument associated with the OBJ function. This is illustrated in (1).



Neidle (1994) argues that “grammatical functions are universal primitives within this framework [LFG], and since they are associated both with lexical and with syntactic positions – by means of annotated phrase structure rules – they mediate between lexical and constituent structure representations”. Although these grammatical functions play an essential role in LFG, they have no intrinsic significance and are situated at the interface between the lexicon and the syntax: “LFG imposes the restriction of Direct Syntactic Encoding, which prevents any syntactic process from altering the initial assignment of grammatical function” (Ibid, emphasis in original).

Kaplan & Bresnan (1982) show that grammatical functions are associated with lexical items, which have functional schemata, and syntactic positions by means of annotated PS-rules. These annotated PS-rules function as an interface between lexicon and syntax. Example (2) provides an example of schematized formats of LFG Lexical Items. The name *?aHmad* ‘Ahmed’, for instance, comes with grammatical data such as gender. As illustrated in (2), a lexical rule takes a lexical item as input and returns a new lexical item. It is defined over a whole class of items.

- (2) *?aHmad* ‘Ahmed’, Noun
 (↑ PRED) = '{meaning of 'Ahmad'}'
 (↑ GEND) = MAS
 (↑ NUM) = SING

The variable ‘↑’ in this representation refers to the lexical item under which this entry is found, here ‘Ahmad’.

2.2. Constituent structure (C-structure)

Sadler et al. (2000:4) point out that in LFG “the correspondence between functional and constituent structures is partly defined in terms

of annotations associated with c-structure nodes”. They also argue that these used annotations follow “universal and language specific principles which involve partial PS configurations and apply them to all Context-free grammatical rules that meet such relevant partial configurations” (Ibid. 4). Functional annotation used in LFG is functional schemata transferred into tree and is interpreted as deriving data about the functional structure. The LFG c-structures have also the form of context-free phrase structure trees.

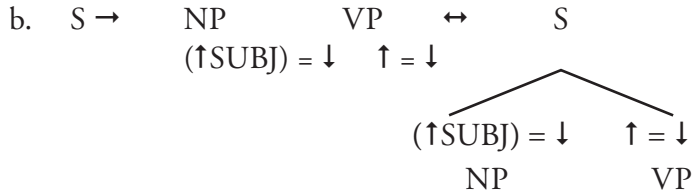
All that is needed to create c-structure in LFG is context-free phrase-structure rules, PS-trees, and the insertion of functional schemata. Language specific annotation of PS-rules helps to identify the grammatical functions that occur in specific syntactic positions, i.e. word order whether it is SVO, VSO, or OVS.

Consider the sentence in (3):

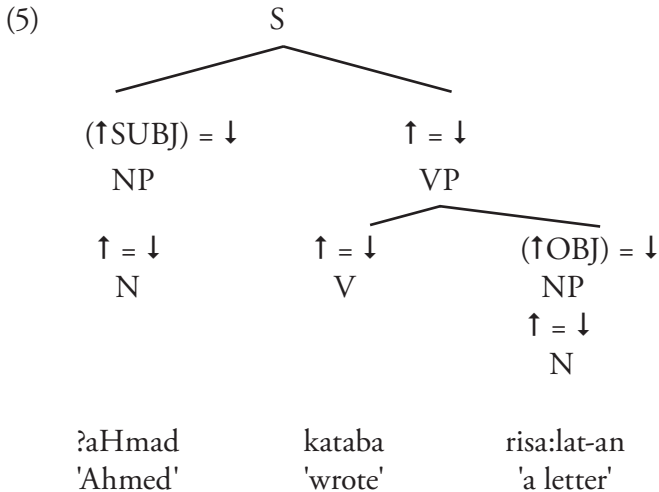
- | | | | |
|-----|-------------------------|---------------------------|---------------------------------------|
| (3) | ? <i>aHmad</i>
Ahmed | kataba
write-3MAS-PAST | risa:lat-an
letter-ACC-SG-
INDF |
|-----|-------------------------|---------------------------|---------------------------------------|
- 'Ahmed wrote a letter.'

In order to consider this sentence, one has to consider its PS-rules, and then construct a tree with annotations that prescribe these rules. The PS-rules in (4.a) illustrate this point. Thus (4) shows a relation between rules and annotations in the tree from an LFG perspective.

- | | | | | |
|-----|----|------|---------|-----------|
| (4) | a. | S → | NP | VP |
| | | NP → | (Det) N | (AP) (PP) |
| | | PP → | P NP | |
| | | VP → | V (NP) | (NP) (PP) |



The up-arrow (\uparrow) refers to the f-structure of the mother node. The node that immediately dominates the constituent under which the arrow is placed substantiates it. On the other hand, the down-arrow (\downarrow) refers to the f-structure of the current node. Thus from the PS rule $S \rightarrow NP VP$, the equation states that the NP is the (SUBJ) of S that dominates it. The $(\uparrow = \downarrow)$ equation beneath the VP indicates that the features of that node are shared with higher nodes, making all functional information carried by this node also direct information about the mother's f-structure. Finally, the $(\uparrow\text{SUBJ}) = \downarrow$ means that all functional information carried by this node goes with the subject part of the mother's function. Thus the sentence in (3) above can be represented in a tree such as the one shown in (5).



By introducing the annotations specified by the lexical entries for *?Hmad kataba risa:lat-an*, the sentence is complete. This is achieved by consulting the lexical items in the tree for their functional structure.

2.3 Functional structure (*F-structure*)

In LFG, functional structures are *unification-based* grammars and encode information about the various reactions between parts of sentences. Sadler et al (2000:3) maintain that “the correspondence between s-structure and f-structure follows from linguistically determined principles which are partly universal and partly language specific (Bresnan 2001), (Dalrymple 1999)”. These f-structures are set of pairs of attributes and values. They are themselves functions from attributes to value, where the value for an attribute can be configured in (6) as:

- (6) a. Atomic symbols: e.g. [GEND FEM]
- b. Semantic forms: e.g. [PRED 'qa:lai (↑S (↑OBJ)i)]
- c. f-structure: e.g.:

$$\left[\text{SUBJ} \left[\begin{array}{ll} \text{PRED} & \text{'Ahmed'} \\ \text{NUM} & \text{SING} \\ \text{GEND} & \text{MAS} \end{array} \right] \right]$$

Schematized or graphical f-structures in LFG are represented as information enclosed in large square brackets as shown in (6). A basic assumption in LFG is that some f-structure should be associated with each node in the c-structure tree. Thus the lexical information is combined with the structure information available from the c-structure tree to get the f-structure.

2.3. F-structure well-formedness conditions

The LFG theory posits some well-formedness conditions on the f-structure, which is to be valid according to such well-formedness conditions. These well-formedness conditions filter out overgeneration in c-structures. These well-formed conditions are Functional Uniqueness (i.e. each attribute in the matrix should have a unique value), *Completeness* (i.e. an f-structure is ill-formed if it does not contain values for all the grammatical functions that are subcategorized by the predicate), and *Coherence* (i.e. every meaningful semantic form is required to be a grammatical function stated in the predicate argument structure of a predicate in its clause). Thus f-structures in LFG should be *unique*, *complete*, and *coherent*. The functional uniqueness ensures that each attribute has a unique value. In other words, functional uniqueness refers to or is equivalent to consistency. For example, the f-structure in (7.a) below refers to an example of consistent structure in the f-structure whereas the example in (7.b) refers to an inconsistent structure in the f-structure.

- (7) a. $\begin{bmatrix} \text{NUM} & \text{SING} \\ \text{NUM} & \text{PL} \end{bmatrix}$ [NUM]
- b. $\begin{bmatrix} \text{GEND} & \text{FEM} \\ \text{GEND} & \text{MAS} \end{bmatrix}$
- c. *kataba*

For the f-structure to be complete, the Completeness Condition makes sure that the subcategorization requirements are met. Hence, an f-structure is not well formed if it lacks *values* for grammatical functions that are subcategorized by the predicate. The example in (7c) lacks a value for the (SUBJ) and is therefore considered incomplete.

Finally, the Coherence Condition on f-structures ensures that every argument is the argument of some predicate. The f-structure in (8) represents the sentence in (3).

(8) F-structure of (3):

$$\left[\begin{array}{ll} \text{PRED} & \text{'kataba } \langle (\uparrow\text{SUBJ})(\uparrow\text{OBJ}) \rangle' \\ \text{TNS} & \text{PAST} \\ & \left[\begin{array}{ll} \text{PRED} & \text{'Ahmed'} \\ \text{PERS} & 3 \\ \text{NUM} & \text{SING} \\ \text{GEN} & \text{MAS} \end{array} \right] \\ \text{SUBJ} & \\ \text{OBJ} & \left[\text{PRED} \quad \text{'risa : lat - an'} \right] \end{array} \right]$$

Frank (2000:2) suggests that the f-structure annotation principles provide by themselves a principled-based, modular description of the LFG c-structure/f-structure interface. They define characteristic functional correspondences between partial c-structure configurations and their f-structure projections. By abstracting from away from irrelevant c-structure context, these principles are highly general and modular, and therefore apply to previously unseen tree configuration.

2.4 Argument structure (A-structure)

A-structure is a representation of participants in the action expressed by the PRED. Bresnan (2001: 307) pointed out that “[a]n a-structure consists of a predicator with its argument roles, an ordering that represents the relative prominence of the roles, and a syntactic classification of each role by a feature”. She added that the “relative prominence of the roles is indicated by their left-to-right order and reflects a thematic

hierarchy”, which is stated in (9):

- (9) Thematic Hierarchy:
agent > beneficiary > experiencer/goal > instrument >
patient/theme > locative
 (Ibid.:307)

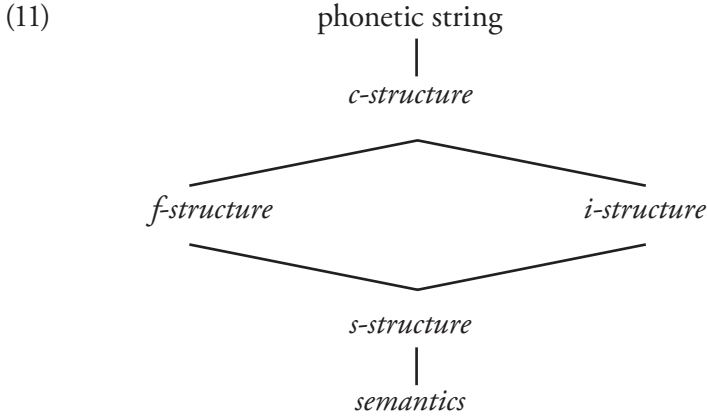
The a-structure of sentence (3) is represented in (10):

- (10) *kataba* < Agent, Theme >

The structure in (10) shows that there are two arguments for the verb *kataba* ‘wrote’, which include an agent and a theme. According to the Thematic Hierarchy, the agent precedes the theme. This thematic hierarchy will have influence on the mapping between a-structure and f-structure that is governed by the Lexical Mapping Theory (see Bresnan 2001; Falk 2001).

2.5 Information structure (*I-structure*)

This study makes use of *i-structure* that was proposed by King (1997). King (1997: 7) proposes that in any complete analysis of a sentence, “all lexically substantial items have a discourse function role (topic, focus, background, etc.) based on the utterance context”. She adds that “[s]ince the discourse function constituents do not always overlap with the f-structure constituency, information should be projected from the c-structure into a distinct *i(nformation)-structure*, which can be accessed by the *s(ematic)-structure*.” The following diagram in (11) shows this (King 1997: 8):



According to (11), the *i-structure* of (3) is presented in (12):

(12) I-structure of (3)

(4)
$$\left[\begin{array}{ll} \text{DISCOURSE TOPIC} & \{\text{Ahmed}\} \\ + \text{New} & \left\{ \begin{array}{l} \text{kataba} \\ \text{risa : lat - an} \end{array} \right\} \end{array} \right]$$

The *i-structure* in (12) specifies that *Ahmed* is the discourse topic whereas *kataba* and *risa:lat-an* are [+New] information. The presentation in (12) is just one way of the possible *i-structures* of (3). To decide the information distribution of a construction, the context has to be taken into account.

3. TOPIC and FOCUS constructions in Arabic

This section illustrates how the TOPIC and FOCUS discourse functional constructions in Arabic are illustrated in LFG.

3.1. Topicalization: TOPIC and FOCUS definitions

Topicalization is a syntactic phenomenon that has been used and given a variety of definitions (cf. Gregory & Michaelis 2001; Krapova 2004; Lambrecht 1994; LaPolla 1998; Mathews & Yip 1994; Rosen 1998; Mimura 2009; Speyer 2010, etc.). For example, Rosen (1998:184) defines the topicalization phenomenon as “a construction in which a leftmost constituent is understood as filling a missing constituent in the sentence”.

In his definition of topicalization, Crystal (1997:392) points out that “topicalization takes place when a constituent is moved to the front of a sentence, so that it functions as a topic”. This is to say that a constituent is left-dislocated or preposed to the front of a sentence to function as the FOCUS or TOPIC of the sentence.

Gregory and Michaelis (2001:1668) differentiate between topicalization and left-dislocation when they say that a TOPIC and Left-Dislocation(LD) differ formally in the following way:

TOP [TOPIC] contains a gap in the clause which corresponds to an argument position that the preclausal NP can be construed as filling, whereas left-dislocation contains an argument-position pronoun which is coreferential with the preclausal NP... Since LD sentences contain no gaps, they are compatible predications with or without the left-dislocated NP. In other words, the detached NP is nonsyntactic at least in the sense that it does not participate in the predicate-argument structure of the clause.

Mimura (2009:287) in his discussion of topicalization argues that Heavy NP Shift (HNPS) construction in English involves “Focus-Topicalization of the heavy NP and Topic-Topicalization of the remnant TP, based on some similarities between HNPS and focus-Topicalization in the language”.

Snider and Zaenen (2006:326) also use Topicalization in English

to refer to “a syntactic construction without implying a link to pragmatic topichood... We refer to sentence initial constituents in these constructions as the *fronted elements* and to the subcategorization relation between the fronted element and the verb as the in-situ grammatical function (emphasis in original)”. This last approach discusses topicalization from a syntactic and pragmatic perspective. Workman (2009), on the other hand, discusses Topicalization in Malagasy from a pedagogical point of view, which is not the focus of current study.

Lambrecht (1994:127) characterizes a topic as a relation to a proposition by saying that “a referent is interpreted as the topic of a proposition if in a given discourse the proposition is construed as being ABOUT this referent”. Grundel’s (1988:210) definition of topic is also compatible with Lambrecht’s, and provides a clear illustration of the *aboutness* relation: “TOPIC. An entity E is the topic of a sentence S, iff in using S the speaker intends to increase the addressee’s knowledge about, request information about, or otherwise get the addressee to act with respect to E.”

Prince (1997:128) defines topicalization from a pragmatic point of view by pointing out that “topicalization triggers influence on the part of the hearer that the entity represented by the initial NP stands in salient partially-ordered set relation to some entity or entities already evoked in the discourse-model”. Givón (1983:11) points out that the encoding of 'topics' is determined by their relative 'availability to the hearer', by which he means, “what may reasonably affect the *degree of difficulty* that speakers/hearers may experience in *identifying* a topic in discourse” (emphasis in original).

In his treatment of Topicalization and Stress Clash Avoidance, Speyer (2010: 1) argues that “topicalization is an exemplary case for demonstrating this interaction [of phonology and syntax] and the power of the Clash avoidance requirement, and therefore much of this [his] text is devoted to a discussion of topicalization in the history of English”. Fung (2007a, 2007b) discusses topicalization in Cantonese from an OT-LFG account. She starts the discussion of topicalization

in Cantonese “by looking at the terms ‘TOPIC’ and ‘FOCUS’, which refer respectively to topicalized discourse topic and topicalized discourse focus conforming to the definition of topicalization in LFG”.

As mentioned earlier, topicalization involves a preposing of the predicate in a nominal sentence or the object in a verbal sentence in order to function as the TOPIC/FOCUS of the sentence. This topicalized element can also be separated from its clause mates by another clause. Topicalization observes syntactic constraints upon long-distance dependencies. To account for these dependencies by appealing to non-structural architectural assumptions responsible for the representation of grammatical functions within LFG, Kaplan and Zaenen (1989) develop a formal apparatus called Functional Uncertainty (FU). The current paper explores the interaction between FU apparatus in LFG and the topicalization paradigms in Arabic.

3.2 Discourse TOPIC/FOCUS vs. Syntactic TOPIC/FOCUS

TOPIC and FOCUS are sometimes confusingly used to refer to both the discourse and syntactic notions. To resolve such confusion, scholars use different terms that differentiate between the syntactic and discourse concepts. For instance, Bresnan & Mchombo (1987: 746) point out that “grammatical topics-constituents that bear the TOP[IC] function-designate discourse topics; but not all discourse topics are grammatically marked”. According to Bresnan and Mchombo topicalized topics, TOPICs, designate discourse topics whereas discourse topics do not have to be topicalized. The same can be said to FOCUSes. Topicalized focuses, i.e. FOCUSes, designate discourse focuses whereas discourse focuses are not necessarily topicalized. Following Bresnan and Mchombo (1987) and the general practice in LFG which represents grammatical functions in small capitals, ‘TOPIC’ and ‘FOCUS’ are used in this study to refer to topicalized discourse topics and topicalized discourse focuses respectively.

4. *Arabic TOPIC and FOCUS constructions in LFG*

The LFG framework provides a systematic account of topicalization. At the beginning, a phrasal structural framework was introduced for the topicalized TOPIC and FOCUS phenomena (Kaplan & Bresnan 1982, Zaenan 1980, 1983 cited in Kaplan & Zaenan 1989). Later, a functional approach was fully developed and suggested by Kaplan & Zaenan 1989. King (1997) suggested an additional information approach for representing discourse information. Marfo & Bodomo (2005:185) also argued that “the function of the FOC [FOCUS] in focus constructions...is a semantic one. That is, it has discourse information alteration significance and, for that matter, it induces semantic contrast”.

This section discusses the suitability of adopting only a functional approach for accounting for the Arabic TOPIC and FOCUS discursal functions. Later, the informational approach will be investigated to see how it complements the functional approach to suitably account for the topicalized constructions in this study.

4.1 A Functional approach to TOPIC and FOCUS constructions

In the LFG framework, topicalization is considered a kind of long distance dependencies. Kaplan & Zaenan (1989) proposed the Functional Uncertainty (FU) apparatus as a formal device for accounting for long distance dependencies. Below is a brief description of the FU. The functional uncertainty equations for Arabic topicalized constructions are provided.

4.1.1 Functional Uncertainty

Functional Uncertainty is the formal apparatus in LFG for configuring long distance dependencies. It is a device that is based in functional

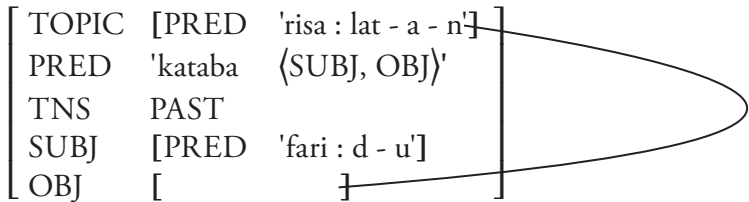
terms, rather than in phrase structural terms. Kaplan & Zaenan (1989: 297) point out that the FU “permits a functional statement of constraints on unbounded dependencies”.

In order to account for long-distance dependency phenomena such as topicalization, the FU developed by Kaplan & Zaenan (1989) will be used. This FU represents a departure both from the original analysis of long distance dependencies within LFG, which was structure-based, and from the standard appeal to structure within generative proposals. To see how this works, consider the sentence in (13).

- (13) [risa:lat-a-n]_i fari:d kataba Ø_i
 a letter- ACC-INDEF Farid wrote-3MAS-SG-PAST
 ‘A letter, Farid wrote./ Farid wrote a letter.’

Example (13) is an instance of topicalization in Arabic where *risa:lat-an* ‘a letter’, the TOPIC, is identified with the OBJ argument absent from the c-structure headed by the verb *kataba* ‘wrote’. The *f-structure* for sentence (13) is provided in (14).

- (14) Schematized *f-structure* representation for a TOPIC & OBJ



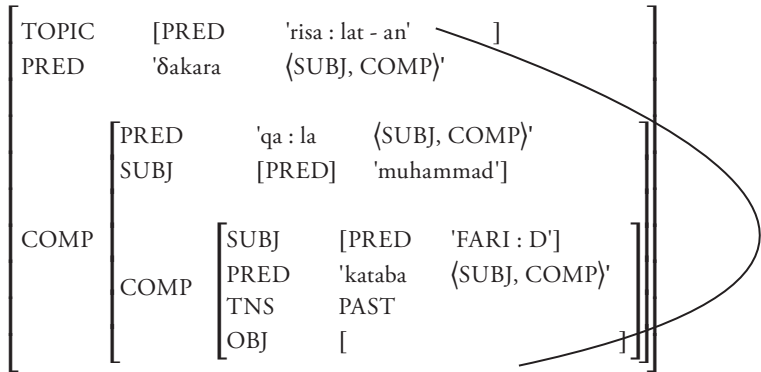
According to LFG, the missing OBJ argument is not represented as an empty or null category in c-structure, but is rather represented only in the *f-structure* representation. The line connecting the TOPIC function and the OBJ function argument in (14) indicates that the value for the

TOPIC is identical to the value for the OBJ attribute: they both have the value *'risa:lat-a-n'*. To put it differently, the TOPIC function provides the value for some grammatical functions, which would otherwise have none. According to the well-formedness conditions on f-structure discussed before, LFG demands that all functions associated with an argument taking entity must have a value, thus the OBJ function must bear a relation to some nominal in order for the sentence to be well-formed. Example (13) presents an instance in which the OBJ is an argument of the matrix predicate and thus bears a 'local' relationship to the TOPIC. On the other hand, the TOPIC can provide a value for a non-local argument as in example (15).

- | | | | |
|------|--|----------------------|------------------------------|
| (15) | [risa:lat-an] _i
a letter-ACC | ?aHmad
Ahmed | δakara
stated-3MAS-PAST |
| | ?anna
that | muHammad
Muhammed | qa:la
said-3MAS-SG-PAST |
| | ?inna
that | fari:d
Farid | kataba
wrote-3MAS-SG-PAST |
| | ∅ _i | | |
- 'A letter, Ahmad mentioned that Muhammad said that Farid wrote.'

The f-structure for (15) is presented in (16). Example (15) above consists of three clause nuclei; the TOPIC *risa:lat-an* 'a letter' corresponds to an OBJ argument embedded two clauses deep. In (16), the relationship between the TOPIC and OBJ is presented by the f-structure. The line in (16) conveys that the missing value for the OBJ attribute, after two clauses deep, is identified with the value for the TOPIC. Thus, all the attributes in the f-structure have values and the f-structure as a whole is well formed.

(16) TOPIC and OBJ relationship in an f-structure



Kaplan & Zaenan (1989) found out that the relationship between a gap and a filler can be captured by an equation. The FU constrains the relationship between fillers and gaps in local and non-local dependencies. Kaplan & Zaenan (1989) propose the following FU schema in (17).

(17) Schema for Functional Uncertainty (=FU)

$$S' \rightarrow \Omega \quad \Sigma$$

$$(\uparrow \text{DF} = \downarrow) \quad (\uparrow \text{DF} = (\downarrow \text{BODY}^* \text{BOTTOM}))$$

Where DF = discourse functions, the values for BODY and BOTTOM is some (set of) GF(s) and $\Omega = \text{XP}$ and $\Sigma = S$

The FU schema in Figure (17) postulates that the value of a maximal phrasal category, Ω , bearing a discourse function, e.g. in the present discussion a TOPIC, be identified with the value for a grammatical function (GF) attribute of some element (=BOTTOM) an indeterminate distance away (= BODY*). Kaplan & Zaenan (1989: 27) devise FU that is able to denote a set of strings:

- (18) $(fa) = v$ holds if and only if
 $((fs) \text{ Stuff}(s, a)) = v$ for some symbol s ,
 where a is a (possibly infinite) set of strings and $\text{Stuff}(s, a)$
 is the set of suffix strings y such that $s y \in a$

By applying the Schema for Functional Uncertainty in (17 and 18), the schematic representation for example (15) will be like the following one in (19).

- (19) $(\uparrow \text{ TOPIC }) = (\uparrow \text{ COMP}^* \text{ OBJ})$

The Kleene star operator ‘*’ indicates that the OBJ can be embedded in any number of COMPs, including zero. This equation is specifically relevant for the topicalization of the OBJ function. According to (19), FU identifies *risa:lat-an* ‘a letter’ as the value for both the TOPIC function and the OBJ function by tracing it through a path of an indeterminate number of COMPs. Thus, FU relates a many-to-one function of attributes into a single value. This function-based apparatus ensures that only maximal categories (XP) can be topicalized, whether these maximal categories are local as shown in example (13) or non-local as shown in example (15).

Kaplan & Zaenan (1989: 32) provide a more general equation for long distance dependencies in (20):

- (20) $(\uparrow \text{ DF}) = (\uparrow \text{ body bottom})$
 where DF is the discourse function, body is the uncertainty path, and bottom is the end of the path.

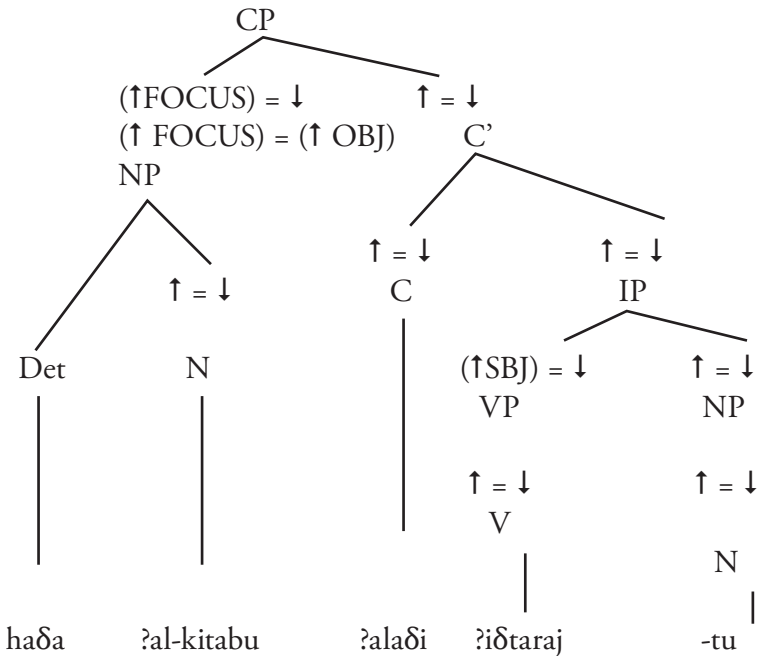
The relevant DFs for topicalization are TOPIC and FOCUS. As an illustration, the functional uncertainty equation for the Arabic example in (21) is shown in (22).

- (21) haḍa ?al-kitabu ?alaḍi ?iḥtarajtu
 This the-book that bought-I
 ‘This is the book that I bought.’

(22) (\uparrow FOCUS) = (\uparrow OBJ)

The equation in (22) captures the relationship between the leftmost FOCUS and the missing OBJ in (21). It indicates that the FOCUS of the clause *haḍa ?al-kitabu* is at the same time the OBJ of the clause. This equation is annotated to the c-structure node of the FOCUS in order to specify the unification of the topicalized function and the missing function as in (23).

(23) Annotated *c-structure* of (21)



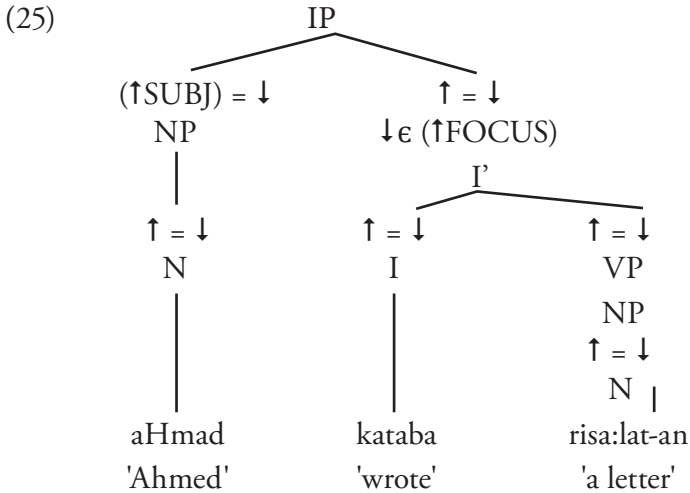
‘(↑FOCUS) = ↓’ above specifies the NP *haḍa ?al-kitabu* specifies that the FOCUS of the sentence is the present NP. ‘(↑ FOCUS) = (↑ OBJ)’ indicates that the FOCUS of the sentence is at the same time the OBJ of the sentence.

4.1.2 An I-Structure for discourse functions

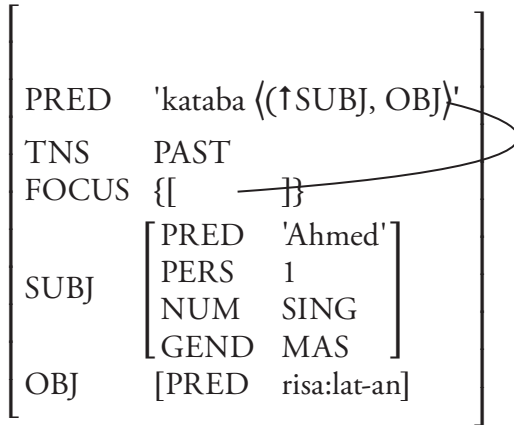
The use of standard FU is limited in characterizing discourse functions. King (1997) postulates an independent i-structure for accommodating such discourse information. The use of functional annotations posits problems for licensing discourse focus. Scoping problems will result if the element designating the discourse focus is an f-structure head. King (1997:7) illustrates this problem with a Russian example where the PRED is the discourse FOCUS. By the same token, the same scope problem results in Arabic FOCUS. The example in (3), repeated here as (24) for convenience, can either be interpreted as focusing just the object *risa:lat-an* ‘a letter’ or the verb and its object *kataba risa:lat-an* ‘wrote a letter’.

- (24) ?aHmad kataba risa:lat-an
 Ahmed write-3MAS-PAST letter-ACC-SG-INDF
 'Ahmed wrote a letter.'
 FOCUS = ‘wrote a letter’
 (in an answer to: what did Ahmed write?)
 [also can have: focus = [‘letter’]

The illustrations in (25) and (26) show how to assign this type of focus throughout annotating a clause final constituent ↓ε (↑FOCUS).

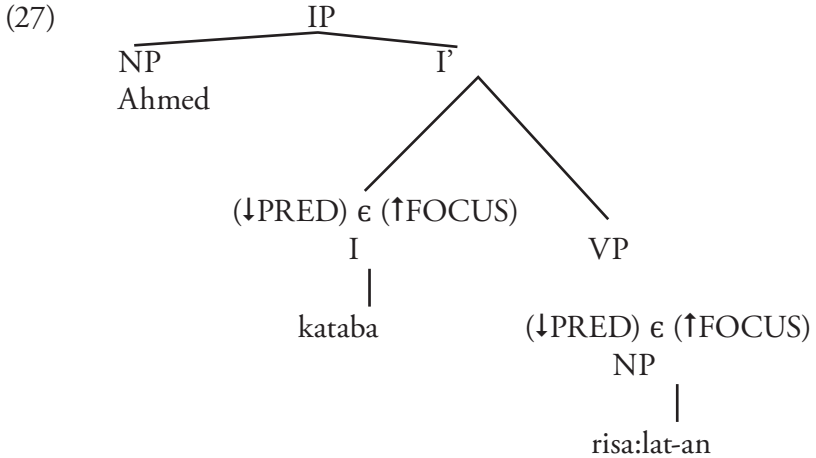


(26) *F-structure* for (25):



The configuration in (26) results in the entire f-structure being the focus of the clause, not just the verb and its object. Thus the annotation of the PRED with ↓ε (↑ FOCUS) leads to the wrong conclusion that the whole f-structure is the discourse focus since the annotation to the PRED is passed to the entire f-structure.

Another possibility is to annotate the clause-final PRED which is the discourse focus with the annotation (\downarrow PRED) ϵ (\uparrow FOCUS). This results in the illustrations in (27) and (28).



(28) F-structure

[PRED	'kataba <math>\langle\langle\uparrowSUBJ, OBJ)>>'
]	TNS	PAST
]	FOCUS	{ [] }
]	SUBJ	[PRED 'Ahmed' PERS 1 NUM SING GEND MAS]
]	OBJ	[PRED risa:lat-an]

Because the PRED of the verb contains its arguments, here the SUBJ and the OBJ, and as a result will incorrectly include the subject in the focus. King (1997) argues that the problem is that “focusing the

head of the *f*-structure results in the whole *f*-structure being focused, including the arguments of that head”.

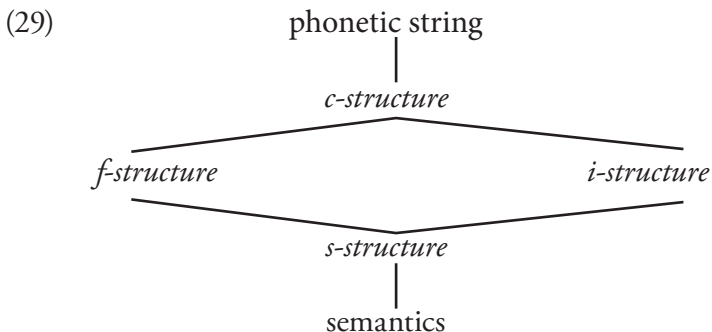
5. *Solution to the limited functional account*

To solve the limitations of the *f*-structure as the only way to account for the TOPIC and FOCUS discourse functions, an *i*-structure projection should be introduced and should be also distinct from that of the *f*-structure.

5.1. I-Structure

To completely analyze a sentence, all lexically substantial items have some discourse function role (e.g. topic, focus, background, etc.) that is based on the utterance context. An *i*-structure projection should be adopted in order to overcome all the scoping and mismatch problems that result from adopting only a functional approach.

King (1997) states that “since the discourse function constituents does not always overlap with the *f*-structure constituency, information should be projected from the *c*-structure into a distinct *i*-structure, which can be accessed by the *s*-structure”, as represented in (11) above and repeated here as (29).



Consider the question-answer pair in (30):

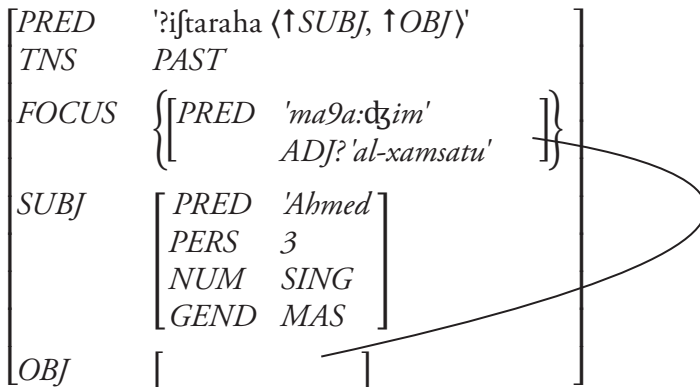
(30) Q: ma: hiya ?al-xamsat kutub
 what are the-five books
 ?alati: ?iftara:-ha ?ahmad
 that bought-it Ahmed

‘What are the books that Ahmed bought?’

A: ?in-ha ?al-xamsatu ma9a:ɗim ?alati:
 it-is the-five DICTIONARIES that
 ?iftara:-ha: ?ahmad
 bought-it Ahmed

‘They are the five DICTIONATIES that Ahmed bought.’

(31) *F-structure* for (30A):



In order to solve the scoping focus problem of PRED in the f-structure in which the argument of the verb was included in the scope of focus when the verb was focused. To focus just the basic meaning of the PRED instead of the whole PRED – including its argument structure

–, Kaplan & Maxwell (1996) and King (1997) refer to this value as PRED FN. For example, King (1997: 9) provides the verb ‘read’ that has the PRED and the PRED FN values in (33).

(32) The *i*-structure of (30A):

<i>DISCOURSE TOPIC</i>	{ <i>Ahmed</i> }	
+ <i>NEW</i>	{ <i>ma9a:ɗim</i> }	
- <i>NEW</i>	{ <i>?iʃtaraha</i> }	
	{ <i>?alxams</i> }	

(33) PRED ‘read <(SUBJ, OBJ)>
 PRED FN <read>

Thus by stripping off the arguments of the PRED and hence not projecting the argument structure into the *i*-structure, many of the problems for the TOPIC and FOCUS discourse and information functions can be solved.

6. Conclusion

This paper investigates the TOPIC and FOCUS functional discourse constructions in Arabic. It answers these questions: 1) how should Arabic topicalized/focalized constructions be syntactically/functionally defined? 2) How should Arabic TOPIC and FOCUS constructions be configured via the LFG Functional Uncertainty (FU) apparatus? 3) To what extent does the LFG framework suffice to account for the TOPIC and FOCUS functional constructions in Arabic? And 4) How could the incorporation of *i*-structure configuration help defining the FOCUS function and overcome the scoping problems that existed in the topicalization of the PRED of the matrix *f*-structures? The study provides a non-derivational analysis of the syntax-discourse interface

of Arabic TOPIC and FOCUS constructions through integrating both the LFG-FU and the *i-structure*.

The Arabic TOPIC and FOCUS constructions can be syntactically/semantically accounted for by using the LFG-FU and its functional equations that specify the grammatical functions to which TOPIC and FOCUS can be identified. The suggested incorporation of an *i-structure* for the FOCUS function proves that the [+New] feature is what licenses FOCUS and considers FOCUS a topicalized construction that contains some [+New] information. Because of the occurring mismatches of TOPIC and FOCUS in the *f-structure*, the adoption of information framework provides an independent *i-structure* that is responsible for representing information that defines the elements in the FOCUS that are [+New] and the ones that are [-New].

The study concludes that although the Arabic TOPIC and FOCUS bear discourse information, they should be also included in the LFG *f-structure*. It specifies how a construction with only a topicalized PRED should be presented in the *f-structure*. The incorporation of both the *f-structure* and *i-structure* in the suggested functional/informational framework has solved many of the resulted scoping semantic/discoursal problems that existed in the topicalization of the PRED of the matrix *f-structures*. The informational framework and its incorporated *i-structure* complement the *f-structure*, which includes an additional *i-structure* that independently captures the [+New] and [-New] features of the Arabic FOCUS functions.

Department of English
Gulf University for Science and Technology
P.O. Box 7207
Hawally 32093, Kuwait

References

- Bresnan, Joan. 2001. *Lexical Functional Syntax*. Oxford: Blackwell Publishers.
- Bresnan, Joan. & Jonni Kanerva. 1989. Locative inversion in Chichewa: A case study of factorization in grammar. *Linguistic Inquiry* 20. 1-50.
- Bresnan, Joan & Sam Mchombo. 1987. Topic, pronoun, and agreement in Chichewa. *Language*, 63 (4). 741-782.
- Crystal, David. 1997. *A Dictionary of Linguistics and Phonetics* 4th Edition. Cambridge, Massachusetts: Blackwell Publishers.
- Dalrymple, Mary. 1999. *Lexical-Functional Grammar*. In Rob Wilson and Frank Keil (eds.), *MIT Encyclopedia of the Cognitive Sciences*. Cambridge: The MIT Press. 464-467.
- Falk, Yehuda. 2001. *Lexical-Functional Grammar: An Introduction to Parallel Constraints-Based Syntax*. Chicago: CSLI Publications.
- Frank, Anette. 2000. Automatic F-Structure annotation of Treebank Trees. *Proceedings of the LFG00 Conference*. Berkeley, CA: CSLI Publications. 1-21.
- Fung, Cauvis. 2007a. OBJECT topicalization in Cantonese. Paper presented at the Fifth Cambridge Postgraduate Conference in Language Research. The University of Cambridge, March 20-21.
- Fung, Cauvis. 2007b. Topic and Focus in Cantonese: An OT-LFG Account. Unpublished M.A. Thesis, The University of Hong Kong. Online: http://www.hku.hk/linguist/staff/TopicFocus_Cauvis.pdf
- Givón, Talmy. 1983. *Topic Continuity in Discourse*. Amsterdam: Benjamins.
- Gregory, Michelle L., & Michaelis, Laura A. 2001. Topicalization and Left-Dislocation: A Functional Opposition Revisited. *Journal of Pragmatics* 33:11. 1665-1706.
- Grundel, Jeanette K. 1988. Universals of topic-comment structure. In Michael Hammond, Edith Maravcsik, and Jessica Wirth (eds.), *Studies in Syntactic Typology*. Amsterdam: John Benjamins. 209-239.
- Jackendoff, Ray. 1987. *Consciousness and the Computational Mind*. Cambridge, Massachusetts: MIT Press.
- Jackendoff, Ray. 1997. *The Formal Architecture of Lexical-Functional Grammar*. Cambridge, Massachusetts: MIT Press.
- Kaplan, Ronald M. & Annie Zaenan. 1989. Long Distance Dependencies, Constituent Structure and Functional Uncertainty. In Mark Baltin and Anthony Kroch (eds.), *Alternative Conceptions of Phrase Structure*.

- Chicago: Chicago University Press. 17-42.
- Kaplan, Ronald M. & Joan Bresnan. 1982. Lexical-Functional Grammar: A Formal System for Grammatical Representation. In Joan Bresnan (ed.) *The Mental Representation of Grammatical Relations*. Cambridge, Massachusetts: MIT Press. 173-281.
- King, Tracy H. 1997. Focus domain and information-structure. In Miriam Butt & Tracy H. King (eds.), *Proceedings of the LFG97 Conference*. Stanford: CSLI Publications.
- Krapova, Iliyana. 2004. Word order in Topic-Focus structures in the Balkan languages. *Romanio Orientale* XVII. 139-162.
- Lambrecht, Knud. 1994. Information Structure and Sentence Form: Topic, Focus and the Mental Representations of Distance Referents. Cambridge: Cambridge University Press.
- LaPola, Randy. 1998. Topicalization and the question of lexical passives in Chinese. Paper presented at the Third Annual Ohio State University Conference on Chinese Linguistics, Ohio State University, May 13-14.
- Marfo, Charles & Adams Bodomo. 2005. Information structure in Akan question-word fronting and focus constructions. *Studies in African Linguistics*, 34:2. 179-208.
- Mathews, Stephen & Virginia Yip. 1994. *Cantonese: A Comprehensive Grammar*. London: Routledge.
- Mimura, Takayuki. 2009. "HNPS" as a combination of focalization and topicalization to the left periphery. *Interdisciplinary Information Sciences* 15:2. 273-289.
- Neidle, Carol. 1994. Lexical-Functional Grammar. In *Encyclopedia of Language and Linguistics*. New York: Pergamon Press.
- Prince, Ellen. 1997. On the functions of left-dislocation in English discourse. In Akio Kamio (ed.), *Directions in functional linguistics*. Philadelphia/Amsterdam: John Benjamins. 117-144.
- Rosen, Victoria. 1998. *Topics and Empty Pronouns in Vietnamese*. Doctoral Dissertation. Department of Linguistics, University of Bergen.
- Sadler, Louisa, Josef van Geabith & Andy Way. 2000. Automatic F-structure Annotation from the AP Treebank. In Miriam Butt and Tracy H. King (eds), *Proceedings of the LFG00 Conference* CSLI On-Line Publications: <http://www-csl.stanford.edu/publications>
- Saddock, Jerrold. 1991. *Autolexical Syntax*. Chicago: CUP.
- Snider, Neal & Annie Zaenen. 2006. Animacy and syntactic structure: Front-

- ed NPs in English. In Ronald Kaplan, Miriam Butt, Mary Dalrymple, & Tracy King (eds.), *Intelligent linguistic architectures: Variations on themes*. Stanford: CSLI Publications. 323-338.
- Speyer, Augustin. 2010. *Topicalization and Stress Clash Avoidance in the History of English*. Berlin/New York: De Gruyter Mouton.
- Webelhuth, Gert & Farrell Ackerman. 1999. A Lexical-Functional Analysis of Predicate Topicalization in German. *The American Journal of Germanic Linguistics and Literature*, 11:1. 1-65. Also Online: <http://idiom.ucsd.edu/~ackerman/glac.pdf>
- Workman, Jeremy. 2009. *Topicalization in Malagasy: Effects of teaching Malagasy as a topic language*. Unpublished dissertation, Brigham Young University, Department of Linguistics and English Language.
- Zaenan, Annie. 1980. *Extraction rules in Icelandic*. Ph.D. Dissertation. Harvard University. Reprinted by Garland Press, New York, 1985.