

Arabic Lexical Phrases

Abdullah S. Al-Dobaian*

Department of English and Literature, King Saud University, Saudi Arabia

Corresponding Author: Abdullah S. Al-Dobaian, E-mail: maliknabi@hotmail.com

ARTICLE INFO

Article history

Received: July 20, 2021

Accepted: September 10, 2021

Published: October 31, 2021

Volume: 12 Issue: 5

Advance access: October 2021

Conflicts of interest: None

Funding: None

Key words:

Lexical Phrases,

Regular Phrases,

Construction Grammar,

Syntax,

Semantics

ABSTRACT

The Arabic traditional grammar as well as Chomsky's mainstream theory may not be able to provide a good analysis of some fixed Arabic phrases. The challenge of such data directly stems from the fact that the general syntactic rules assumed by the two opposing theories cannot explain the syntactic and the semantic aspects of the fixed Arabic data. I argue that the Construction Grammar provides an adequate account that does not rely on syntactic structure alone, as assumed by the mainstream theory or the Arabic traditional grammar, but rather it links phonological, syntactic, and semantic information together in one basic construction by means of some correspondence rules. The Arabic data proves that there is a strong need for a linguistic theory that takes into consideration all data of different range of productivity.

INTRODUCTION

I discuss some Arabic data of irregular phrases that prove to be problematic cases for syntactic rules affecting phrases derived in syntax by means of merger. Such data tend to be empirically challenging data to be explained by either the Arabic traditional syntactic theory or even Chomsky's mainstream theory (1981; 2015). More specifically, the syntactic rules may not account for these specific types of phrases because these phrases are syntactically different from the regular phrases. Moreover, some of these specific lexical phrases are associated with special meaning that cannot be determined by syntax.

In the second section, I introduce the data. The third section explains how the Arabic traditional syntactic theory can account for the data and the problems of this analysis. Then it is shown that Chomsky's mainstream grammar as represented by the Government and Binding (GB) (1981) or the Minimalist theory (MP) (2015) is not able to account for these data due to reasons that we discuss. I think that the data are lexical phrases in section four. I suggest that even though the modern mainstream syntactic tradition generally precludes such analysis, there are studies like Hale and Keyser (1993) that suggest that some syntactic structures may be indeed derived lexically. The fifth section provides a construction grammar as one viable theory that can provide a syntactic as

well as semantic analysis of the data in section two. Finally, a conclusion summarizes the main points of the paper.

DATA

Violation of Regular Case Marking and Agreement Rules in Arabic

The adjacency case marking

The adjective *xarib* picks the case marking of the adjacent noun rather than the further noun it modifies:

- (1) *haaḍa ḡuḥr-u ḍab-in xarib-in.* (Ibn Hishaam, 2002, vol. 6, p. 660)
this is hole.3sm-nom lizard.sm-gen ruined.sm-gen.
This is a ruined lizard hole.

Non-default case marking:

In some restricted sentences, the nominative case can be assigned to the subject while the object gets the nominative case. The examples are given in (2) (examples in (2) are taken from (as-Sayuuṭii 1992, vol.3, p. 8)):

- (2) a. *xaraqa aṭ-ṭawb-u al-mismaar-a.*
penetrated.3sm the dress.3sm-nom the nail.3sm-acc
The nail penetrated the dress.

- b. kasara az-zuḡaaḡ-u a-lḡaḡr-a.
broke.3sm the glass.3sm-nom the stone.3sm-acc
The stone broke the glass.

Ibn ʿasfuur (1982) and alxaliil ibn Ahmad alfarahiidii (1985) among other early Arab syntacticians include similar examples from Arabic poetry where the subject and the object get non-default case marking. In such cases, it is semantics rather than case marking that determines the syntactic grammaticality (azZʿubii, 2012). These cases are lexicalised frozen sentences that are heard from early Arabs and hence tend to be defying regular syntactic case assignment. Therefore, we do not apply non-default case marking to regular syntactic sentences as in (3) (example is taken from azZʿubii, 2012, p. 140):

- (3) *xaraqa al-qimaaš-u al-ʾibrat-a.
Penetrated.3sm the cloth.3sm-nom the needle.3sm-acc
*The cloth penetrated the needle.

Accusative case marking praise or insult

In some restricted usage, the accusative case semantically emphasizes insult or praise depending on the syntactic context. Sibawayhi (1988, Vol 2) designates specific topics for these cases. Let us consider the following examples (Examples are taken respectively from Sibawayhi, 1988, Vol 2, p. 63 and 70):

- (4) a. al-ḡamad-u li-l-lah-i rabb-a al-ʾaalam-iin.
the- praise-3sm-nom to Allah-gen Lord.3sm-acc
the universe-3p-gen
All praise be to Allah the Lord of the universe.
b. ʾataanii Zaid-un al-faasiq-a al-xabiit-a.
came.3sm-1st pro.s.acc Zaid-nom the vicious.3sm-acc
acc the wicked.3sm-acc
The wicked Zaid came to me.

In (4a), *rabbā* is assigned an accusative case to semantically emphasize praise to Allah¹. On the other hand, *alfaasiqa* and *alxabiitā* in (4b) are given accusative case to semantically convey insult as Sibawayhi indicates (1988, Vol 2, p. 70). The fact that the accusative case marking in (4) is triggered by semantics rather than a syntactic rule makes this data lexicalized². *alfaasiqa* and *alxabiitā* in (4b) are given accusative case to semantically convey insult as Sibawayhi indicates (1988, Vol 2, p. 70). The fact that the accusative case marking in (4) is triggered by semantics rather than a syntactic rule makes this data lexicalized³.

One case marking for the dual

The dual in Arabic is usually associated with two case marking: nominative case as in *waladaan* ‘two boys’ and accusative and genitive case as in *waladayn*. However as-Sayuuṭii (1992, vol.1, p. 133) and Ibn ʿaqiil (1980, vol.1, p. 57) among others observe that the dual of the noun has one invariable basic form, i.e., the nominative case as in *waladaan*, in some Arabic languages spoken by Kinaanah, Bani alʿAnbar, Bani alḡaariṭ bin Kaʿab and other Arab tribes.

Fixed case structures

There are some frozen phrases with fixed case markings like:

- (5) a. kull-a šay-in wa laa šatiimat-a ḡurr-in. (Sibawayhi, 1988, Vol. 1, p. 281)
every-acc thing-gen and no insult-acc nobleman-gen
Do everything except the insult of a nobleman.
b. haada wa laa zaʿamaati-ka. (Sibawayhi, 1988, Vol. 1, p. 280)
this and not claim-3pf-acc-2s poss
This is the truth and not your claims.
c. al-kilaab-a ʿala al-baqr-i. (as-Sayuuṭii (1992, vol. 3, p. 20)
the dog.3p-acc on the cow.3p-gen
Send the dogs on cows.
d. waḡh-an li waḡh.
face.3sm-acc to face-gen
(I met him) face to face.

All the examples in (5) are used as fixed syntactic structures with fixed case marking as reported by Arab scholars like Sibawayhi and as-Sayuuṭii among many others. To illustrate, *kulla*, *šatiimata*, *zaʿamaati-ka*, *alkilaaba*, and *waḡhan* are all assigned accusative case despite the lack of overt accusative case assigners. These are puzzling data that we would like to discuss and explain.

Violation of agreement rules in Arabic

It is usually the case that the verb has total agreement in number and gender with the preceding subject as in (6):

- (6) al-awlaad-u ḡarab-uu-nii.
The boys.3p-nom hit-3pro.nom.p-1pro.s.acc
The boys hit me.
Once the verb precedes its subject, a default third singular agreement is marked on the verb. However, the agreement rule may not be maintained as Sibawayhi (1988, Vol. 2, p. 40) indicates:
(7) a. ḡarab-uu-nii qawm-u-ka.
hit-3pro.nom.p-1pro.s.acc The people-nom-2s poss
Your people hit me.
b. ḡarab-aa-nii ʾaxawaa-ka.
hit-3pro.nom.du-1pro.s.acc brother-nom.du-2s poss
Your brothers hit me.

The verb *ḡarab* is marked with third plural and dual inflections even though the verb is not preceded by plural or dual subjects. This violates the regular agreement rules in Arabic.

ARABIC TRADITIONAL SYNTACTIC THEORY

Alxaliil Ibn Aḡmad established the rules of Arabic syntax on the foundations of iʿnmaal theory and then other famous Arab linguists like Sibawaihi, the teacher of Basrii, and Kasaaʿii, the founder of Kuufii school, followed their lead (Shehri, 2006, p. 18). The Arab grammarians, regardless of their linguistic tradition, were busy examining and explaining the inflectional endings of words and how they got the case marking and from what ʿaamil or governor. Sometimes

there was no overt case assigner and the Arab linguist had to resort to *taqdiir* ‘reconstruction’ and *?idmaar* ‘suppression’ in order to salvage *iʕmaal* theory (Versteegh, 1997).

In fact *iʕmaal* theory has a domineering effect in Arabic syntax. To illustrate, we can discuss some aspects of the important role that *iʕmaal* has played in Arabic linguistics.

First, *iʕmaal* theory has controlled the syntactic research. The disagreement among Arab traditional grammarians on different major syntactic issues were related to *iʕmaal* and what governor was responsible for the assignment of case to that word and why (Hamza, 2004, p.18). For example, there has been a major disagreement among Arab grammarians regarding the case assigner of *almubtada?* ‘the topic’ and *ilxabar* ‘the comment’. Kuufiis, for instance, assume that both of them assign case to the other. Basriis, on the other hand, suggest that *almubtada?* is assigned a nominative case from *ibtidaa?*, ‘being used a topic’. They explained *ibtidaa?* case marking as the lack of overt governors (Hamza, 2004, p. 19).

Second, the theory of *iʕmaal* has been used to prove or disprove a syntactic fact. For instance, Arab grammarians differ in terms of whether the definite article in *al* was only *l* or the *hamzat* and *l*. Sibawayhi thought that *l* was the definite article and that the *a* is a *hamzat* was a *hamzat was-sill* ‘glottal soft catch’⁴ while Alxaliil believed that *al* is one letter just as *qadd* and *?inna*. Sibawayhi proved his point by using an *iʕmaal* argument that *l* is one letter because a weak governor such as the preposition *bi* in *bi rrajuli* ‘by the man’ skips *al* to assign case to *rajul*. Had *al* been two letters, the weak governor *bi* would not have skipped it. Linguists in the end adopted Sibawayhi’s view and rejected Alxaliil’s (Hamza, 2004, p. 30-31).

Third, Arab syntacticians classified and arranged grammar topics and the content of their publications according to *iʕmaal*. Thus it was no wonder that sections were assigned based on similar *ʕawamil*. For example, accusative particles like *lan*, *kay*, *?an* are discussed in the same section while jussive particles like *lam*, *lamma*, *lam il?amr* are included together. This type of classification reflected the effect of case marking and disregarded the meaning effect like the different style types of negation, emphasis, and exceptions that were scattered in different chapters (Hamza, 2004, p. 32-33).

Fourth, *iʕmaal* theory changed the Arabic style and altered the meaning of the sentence. That is why Ibn Mad=aa? attacked this theory because, as he viewed it, it distorts the Arabic style. For example, in *ya ʕabd llaahi* ‘O servant [acusative] of Allah’, the vocative *ʕabd llaah* is considered by grammarians to be an object of a deleted underlying verb *?adʕuu* ‘call’ in order to explain the accusative case. Ibn Maddaa? thinks that this is unnecessary and changes the meaning of the sentence from an illocutionary, i.e., vocative, to a propositional sentence. As a result, Ibn Maddaa? strongly believes that *iʕmaal* theory complicates Arabic syntax by its assumptions of *taqdiir*, *?idmaar* and *qiyaas* ‘analogy’ (D=ayf, 1986).

The default case marking rules fail to assign case regularly to the Arabic data in section (2). To illustrate in sentence (1), the case agreement rule determines that the adjective *xaribin* agrees with the head noun *ǧuhru* and not the nearest

noun *ǧabin*. But this is not the case. That is why Ibn Hishaam (2002, vol. 6, p. 660) allocates a specific section in his book to address such problematic data. Likewise, in some irregular examples as illustrated in (2) the subject and the object get non-default case marking. However, this mixing of case marking cannot be possibly applied to other sentences as in (3). Regular rules of case marking cannot account for the contrast of case in (2) and (3). Any account of examples in (2) needs to take into consideration its lexicalised nature. azZʕubii (2012: 140) observes that these such sentences are frozen syntactically and that semantics rather than case marking determines the syntactic grammaticality. In other words, even though the case is mixed and the subject gets accusative while the object is assigned the nominative in sentences in (2), the speaker still figures out the meaning in these lexicalized phrases involving a restricted usage. Semantics plays a role in case assignment as can be further evidenced by the examples in (4) where the accusative case is assigned semantically to emphasize praise as in *rabba* in (4a) or insult as in *alfaasiqa* in (4b). Therefore, accusative case assigning is used to emphasize a semantic notion. For instance, Sibawayhi (1988, Vol 3 and 2 respectively) designates specific topics in his book for the discussion of the semantic assignment of accusative case to emphasizes insult or praise.

In order to explain the data in section (2), adherents to case marking had to assume an underlying governor or case assigner. In fact, Sibawayhi (1988, Vol. 1, p. 281) argues that in example (5) there is an underlying governor *?iti* assigning accusative case to *kulla* and that this governor is obligatorily deleted to avoid repetition. As-Sayuuṭii (1992, vol. 3, p. 19) states the same analysis. However, to assume the presence of underlying governors and then delete them complicates rather than facilitates the understanding of Arabic syntax as Ibn Maddaa? suggested (Dayf, 1986). Furthermore, the data of section (2) involve fixed syntactic structures that are heard from early Arabs and used as lexical memorized chunks. As a result, such data are not regular syntactic structures to which regular case assignment applies (azZʕubii, 2012, p. 142).

CHOMSKY’S MAINSTREAM GRAMMAR

The data in section (2) poses problems for the mainstream grammar not just due to their multiword or the phrasal structure but also the special meaning with which they are associated. To illustrate, regular phrases are derived by means of merging words together. Secondly, syntax according to the Chomskyan theories of GB and MP determines semantics. According to mainstream analysis, the meaning of the phrase is dependent on syntactic structure. Below I explain that lexical phrases as the data in section 2 may not be derived through syntactic merger nor their meaning is dependent upon syntax.

Syntactic vs. Lexical Phrases

In this section, we discuss two important issues. First, how syntactic phrases are built and secondly why the phrases in section (2) are different and thus they can be derived through a different mechanism.

As for the first question, building a syntactic phrase structure has been based on two basic terms *lexical insertion* and *merger*. Jackendoff (1997, p. 13) explains how these terms have evolved in the mainstream grammar. In the beginning, *lexical insertion* was used in the GB model as a mechanism to project lexical items in X-bar structure or a syntactic tree. Then later in the minimalist program *merger* is used to derive phrase structures.

Even though merger readily explains how phrases are generated in syntax as a result of recursively combining words together, it cannot account for fixed phrases. For example, Jackendoff (1997; 2003) observes that idioms like *kick the bucket* are lexical VPs that are derived as one lexical unit. Merger therefore does not work in idioms as they directly come from the lexicon and are accessed in syntax as a multiword structure (Jackendoff 1997, p. 153). In other words, merger, as a process accessing single lexical items and combining them into syntactic phrases, fails to apply in the case of idioms since they already enter syntax as lexical phrases.

Early mainstream tradition lacked the interest in multiword structure. Chomsky has distinguished core from peripheral aspects of language and assumed that only the former is subject to syntactic study (Chomsky, 1986, p. 147). That is why syntacticians working in the mainstream grammar are not interested in addressing fixed structures in general since such structures may not be easily accounted for by regular syntactic rules. Nonetheless Chomsky attempts to provide once and for all a syntactic account for one example of fixed structure: idioms. In a footnote, Chomsky proposes that an idiom as *kick the bucket* can be replaced in *deep structure* by a verb (i.e., die) carrying the idiomatic meaning that is interpreted at *logical form* (Chomsky, 1981, p. 146, footnote 94). Jackendoff (2003, p. 167) reviews Chomsky's analysis and points to some problems. First, there are asymmetric idioms with intervening objects that are problematic to the single verb replacement. For example, Jackendoff provides some problematic examples of idioms like *take NP^s to task*, *show NP the door*, and *take NP for granted* (Jackendoff, 2003, p. 168). These idioms are split by NP and hence the NP causes difficulties for the replacement of the idiom. One way to get around this problem and to have the parts together is to assume the movement of NP to the right position of the idiom forming *take to task NP*. As Jackendoff argues such movement is unmotivated except to keep the idiom parts close to allow for the single verb replacement (Jackendoff, 2003, p. 168). Another problem is that idioms sometimes as observed by Jackendoff can be full sentences as *the jig is up*, *that's the way the cookie crumbles*, and *the cat's got NP's tongue*. If we replace these idioms by single verbs, we will have ungrammatical structures because the sentence structure of these idioms will miss the subject. Thus, Jackendoff concludes that Chomsky's proposal of idiom single verb replacement cannot be maintained. As a result, there is no other way but to accept that idioms are lexical VP (Jackendoff, 2003, p. 169).

Even though fixed expressions are incompatible with productive syntactic rules applicable to the core aspects of

language, they are still part of language that a speaker learns and uses all the time (Jackendoff, 2003, p. 179). Therefore, it would be a grave mistake to ignore such constructions.

Within the mainstream camp, Hale and Keyser (1993), contra Chomsky's assumption, argue that phrases project in the lexicon. To illustrate, a head (i.e., N, V, A, P) develops a full-fledged phrasal structure in the lexicon in which the basic structural relations/arguments of that head are represented. These structural relations of the lexical head involve its specifier and complement. Hale and Keyser suggest that *all* verbs in language are essentially lexically phrasal and they draw their evidence from Igbo and Nwachukwu languages in which the majority of their verbs involve lexical phrases. Furthermore, *light verb* structures in Japanese and English are phrasal lexically (Hale and Keyser, 1993, p. 96). In particular, Hale and Keyser investigate the lexical nature of *denominal verbs* like *shelve*, *saddle*, *bottle*. They propose that such verbs are represented lexically as VP in which the noun (e.g., shelf) starts in the object position and then moves to the head verb deriving the denominal verb (e.g., shelve)⁶. What is interesting in Hale and Keyser's analysis is not only that phrases are developed lexically but also syntactic principles like head movement among others are applicable in the lexicon.

In the reminder of this section, if (syntactic) merger does not work in fixed lexical structures since they are considered as one complex item, then what process can account for their generation? It is evident that the regular process of merger that is used to combine words together forming syntactic phrases as *the old house* may not apply in the case of lexical phrases as *kick the bucket*. I argue that lexical phrases are generated by means of a lexical merger. The Arabic data in section 2 as well as idioms in English are all derived by lexical merger⁷. So merger seems to be a relevant process in building structures in the lexicon and syntax but the end result is different. A syntactic merger creates a syntactic structure with a productive semantics⁸ while lexical merger generates a lexical structure with a specific/unpredictable semantics. Next, I deal with the semantic aspects of lexical phrases.

Syntax and Semantics

As we have explained above, regular phrases merged syntactically are expected to have predictable and regular semantics where the meaning of the phrase as *the old man* is compositionally derived from the components of the phrasal structure. However, in fixed structures as in idioms or the data in section (2) semantics is not dependent on syntax. Therefore, assuming that syntactic structure feeds its semantics is not feasible and it leads to false implications. To illustrate, *kick the bucket* is idiom; hence it is not derived by syntactic merger generating the meaning of someone's kicking a bucket since the meaning is not generated by composing the meaning of the parts of the lexical phrase as the case in a regular phrase. But rather the idiom is treated as a lexical phrase with a specific meaning: *die*. As a result, I argue that the semantics of lexical or fixed structures as in idioms or the data in section (2) may not be drawn or determined by syntax.

Before we get to the discussion of the semantics, it is necessary to briefly review the idiosyncratic properties of the syntactic structure of fixed structure. Unlike the case in regular phrase structure, fixed structure is not constrained by regular syntactic rules. For example, these fixed structures are associated with switched case marking as explained above in examples (2) where the subject and the object get the non-default case marking. Furthermore, there is a violation of the regular agreement rule whereby the adjective as in example (1) agrees in case marking with the nearest noun rather than the far away head. Since these are counterexamples to the general rules of case and agreement in Arabic, The early Arab grammarians had to discuss these special structures in specific sections of their books as we explained above. In fact, Arab grammarians attempted to make some of the fixed structures in section (2) consistent with case theory by resorting to *taqdiir* by assuming an underlying governor or case assigner. To illustrate, let us consider the examples in (5) repeated below for convenience:

- (8) a. kull-a šay-in wa laa šatiimat-a ḥurr-in.
every-acc thing-gen and no insult-acc noble-
man-gen
Do everything except the insult of a nobleman.
- b. haada wa laa zaʿamaati-ka.
this and not claim-3pf-acc-2s poss
This is the truth and not your claims.
- c. al-kilaab-a ʿala al-baqr-i.
the dog,3p-acc on the cow,3p-gen
Send the dogs on cows.

Sibawayhi (1988, Vol. 1, p. 281) argues that there is an underlying governor *?iti* assigning accusative case to *kulla* in (8a) and that this governor is obligatorily deleted to avoid repetition. Sibawayhi also proposes that *zaʿamaatika* in (8b) is assigned accusative case by a deleted governor *?atawahu* ‘I assume’ (1988, Vol. 1, p. 280). According to Sibawayhi, syntax determines the meaning of the structure; thus a specific syntactic structure can be associated with different semantics. Let us consider the following example:

- (9) wa amra?tuhu ḥamaalat-a al-ḥaṭab-i. (Al-Masad, verse 4)
and wife-3sm-poss-nom carrier-3sf-acc the wood-gen
And his wife is the carrier of wood.

Sibawayhi (1988, Vol. 2, p. 70) suggests that *ḥamaalat* may not be only considered as *khabar* to *amra?tuhu* and therefore assigned a nominative case but also *ḥamaalat* may be assigned accusative case to semantically specify insult. To justify the assignment of accusative case syntactically, Sibawayhi assumes an implicit case assigner *?adkur* ‘remember’. As explained in section (2.1.3), Sibawayhi observes that there are semantic restrictions (see footnote 2) for emphasizing a special meaning like praise and as result the noun is assigned an accusative case. Therefore, accusative case of *ḥamaalat* is triggered by semantics rather than syntax or by the assumption of deleted case assigners. Semantics is not dependent on syntax as the meaning of the sentences in (8) are specific or lexical and may not be derived from the syntactic structure. For example, (8a) has the meaning that everything can be tolerated except the insult of a nobleman (Sibawayhi, 1988, Vol. 1, p. 281). (8b) means that this

truth and not your false claims (assuūṭii, 1992, Vol 3, p. 19). Finally, the phrase (8c) does not imply the literal meaning, but that you should stay out of trouble (assuūṭii, 1992, Vol 3, p. 20).

To conclude this section, fixed structure has lexical meaning that cannot be drawn from syntax proving that its semantics is independent of syntax and its rules. Syntax works with semantics in order to make the structure grammatical.

CONSTRUCTION GRAMMAR (CG)

In this section, I argue that CG accounts for the syntactic as well as the semantic structure of fixed syntactic structures in section (2) in ways that the mainstream theory cannot.

Before we get into details of CG analysis of Arabic lexical phrases, a brief discussion of the basic concepts and motivation of CG is in order. The fundamental concept upon which CG is established is the pairing of form and meaning/function (Goldberg, 2006; Booij, 2013 among others). Accordingly, language learners acquire constructions with arbitrary pairings of form and meaning. The mapping of form and meaning can be traced back to Ferdinand de Saussure’s seminal work *Course in General Linguistics*. Saussure (1915, p. 67) proposed that language involves linguistic signs in which the association between form and meaning, or in Saussure’s terms *the signifier* and *the signified*, is arbitrary. This has been a basic principle that linguists have adopted ever since. Therefore, every syntactic theory has attempted to address this basic notion although with different perspective and varying consequences as a result.

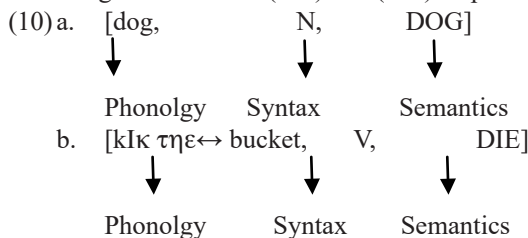
The discussion of form and meaning mapping is at best sketchy and oftentimes inconsistent in the mainstream generative theory. To illustrate, in his *Aspects* book, Chomsky assumed that the *Deep Structure* (DS) derived the meaning of a sentence while the sound was a result of the *Surface structure* (SS). Chomsky, however, changed his view in the *Extended Standard Theory* and thought that meaning was generated from both DS and SS while maintaining the assumption that SS derived the sound structure or the phonological form. Later in the *Revised Extended Standard Theory*, Chomsky believed that SS produced both sound and meaning (Jackendoff, 2003, p. 108-110). From the development of the mainstream theories from the *Aspects* theory to the *Minimalist* program, it is evident that syntax is the only generative engine of grammar while phonology and semantics have only interpretive roles determined by syntax. As a result, syntax is the input deriving both semantics and phonology (Culicover, 1999, p. 13).

Since syntax, in the mainstream generative tradition, has played a major role in the explanation of linguistic data in terms of syntactic rules, it is no wonder that fixed syntactic structures and every data unconstrained by syntactic rules are distinguished from the *core* of language and relegated to the *periphery* of language and therefore ignored.

CG maps syntactic, semantic, and phonological levels together in one basic unit or construction. In such framework, no level, say syntax, has the power to generate other levels: semantics and phonology. Instead, all these levels interact together as Jackendoff argues for a parallel structure (2003). In fact, other linguists like Culicover argue

that syntax cannot account for *peripheral* cases or in Culicover's terms *syntactic nuts* because the rule-governed system of the mainstream grammar does not work (1999). What makes CG theory more appealing is the fact that it does not exclude any syntactic data unlike the mainstream theory. To illustrate, syntactic data can be best analysed in terms of continuum in which different syntactic cases are ranged from the most general rule-governed (e.g., regular grammar) to the less general (irregular grammar) (Culicover 1999) and (Culicover and Jackendoff, 2005). As a result, any syntactic theory must account for both the regular and the irregular or marginal aspects of language (Culicover, 1999). In other words, it is easy to see how this theory works by assuming that there are two mechanisms: one for the regular grammar and another one for the marginal or irregular constructions. These dual mechanisms represent two different learning strategies that a speaker uses to acquire a language (Culicover, 1999, p. 16). The first mechanism is used for core grammar in which general and exceptional rules are used to generate regular syntactic structure. At the same time, the learner of language refers to some other less general rules to account for marginal or fixed syntactic structures that are usually language-specific.

As an illustration of how CG works, the word *dog* and the idiom *kick the bucket* may be represented according to the following constructions (10a) and (10b) respectively:



The word *dog* is one basic construction as the case in *kick the bucket* in which the three levels of phonology, syntax, semantics interact. Each of these levels shares some pieces of information deriving one basic form or construction. Namely, phonology as represented by the phonetic transcription in (10) explains the consonants and vowels that these constructions have. The category is a syntactic information revealing whether the word is a *verb* or *noun* for example. Finally, the semantic information of the idiom is *die* while for *dog* is the semantic as well as the pragmatic connotations associated with DOG in capital letters. These levels have to be controlled by some correspondence rules that we will explain shortly in order to ensure that the grammatical pairing of form and meaning for the linguistic units or constructions is established.

There are two rules linking the syntactic and semantic structures.⁹ These rules are stated as follows (Culicover and Jackendoff, 2005, p. 163):

- (11) a. Head Rule
 A Semantic function F canonically maps to the head H of a syntactic phrase HP.
- b. Argument/Modifier Rule
 The arguments and modifiers of F canonically maps to syntactic constituents of HP.

Before I explain the rules in (11), it might be necessary to briefly discuss the term *function*. This term is used in logic

and mathematics and it is defined as a correspondence relation linking an element from a *domain* set to another element of a *range* set (Allwood, Andersson, & Dahl, 1995, p. 12). The function relation elements are its *arguments*. For example, the fact that a country say, Sweden, has a population of eight million for example can be represented in this function formula: $f(\text{Sweden}) = 8,000,000$, where this relation function has an argument *Sweden* that is linked to the number of the population (Allwood et al., 1995). Linguists, like Jackendoff, use the function concept to represent the correspondence relation between semantics and syntax. To illustrate, the verb predicates, according to Jackendoff (2005, p. 363), are associated with some semantic functions like: (STAY, GO, BE, CAUSE, INCHOATIVE...). Moreover, every syntactic category (e.g., N, Adjective, Adverb...) has a semantic function. For instance, a noun like *boy* is Object, an adjective like *tall* is Property, an adverb like in *class* is Place and so on. The rules in (11) work together coherently connecting the syntactic and the semantic structures. To illustrate, (11a) states that every syntactic head of a phrase is mapped to a semantic function, thus explaining its general semantic field as Jackendoff suggests (2003, p. 357). As for (11b), the syntactic arguments are linked to their semantic arguments and the syntactic modifiers are mapped to their semantic modifiers. In other words, rule (11b) just repeats the theta-theory of the mainstream grammar in which the verb predicate assigns theta roles to its arguments. For example, a verb predicate like *played* in the sentence *Ali played soccer at the stadium* assigns the following theta roles to the NPs or the arguments: an agent role is given to the subject *Ali*, a theme role is given to the object *soccer*, and the PP *at the stadium* is assigned a locative role.

Now we are in a position to explain how CG accounts for Arabic lexical phrases. So, let us consider two examples of the data in section 2.

- (12) a. mararatu biʿabdi Allahi *assaalih-a*. (Sibawayhi, 1988, Vol 2, p. 69)
 passed.3sm-1pro.nom by Abdullah-gen the righteous.3sm-acc
 I passed by Abdullah the righteous.
- b. al-kilaab-a ʿala al-baqar-i. (as-Sayuuṭii, 1992, vol. 3, p. 20)
 the dog.3p-acc on the cow.3p-gen
 Send the dogs on cows.

In the first example (12a), the adjective *assaalih* is assigned an accusative case based on semantic conditions that we have explained in footnote (2) above. Interestingly, Sibawayhi discusses this example as a part of irregular data in a specific topic of his book that he allocates for a semantically based accusative case assignment because this case is not assigned by regular case rules. The CG proposes three parallel structures for the sentences in (12a,b) in (13a,b) respectively:

- (13) a. Phon: maratu: p bI o ʿabdi j llaahI m ↔assaaliha
 Syn: [_{VP} V_p [_{NP} N nom]_i, [_{PP} Prep_o [_{NP} N gen]_j, [_{NP} N gen]_m, [_{AdjP} Adj acc]_n]_q]_r]_s
 Sem: [_{Event} GO_p ([_{Object} I]_i; [_{Path} BY_o ([_{Object} ʾĀABD]_j, [_{Object} ALLAH]_m, [_{Property} S=AAIH]_n, [Def]_r]_q)]]_s

- b. Phon: \leftrightarrow Ikllaab \leftrightarrow \aleph \leftrightarrow l \leftrightarrow \leftrightarrow lb \leftrightarrow q \leftrightarrow r \leftrightarrow j
 Syn: $[_{VP} V_m [_{NP} \text{You nom}]_i; [_{NP} N \text{acc}]_p; [_{PP} \text{Prep } [_{NP} N \text{gen}]_j]_o]_d$
 Sem: $[_{Event} \text{Cause}_m ([_{Object} \text{YOU}]_i; [_{Object} \text{KILAABA}]_q; \text{DEF}_e; \text{P}_c]_p; [_{Event} \text{Inch}_u]_j; [_{State} \text{BE}_n (\text{Path } \text{Ext}_o ([_{Object} \text{BAQAR}]_r; \text{Def}_{n; \text{P}_{s; j}]_a]]]]_d$

Each representation in (13) reveals a three-dimensional structure of information: a phonological, syntactic, and semantic structures. To begin with, the phonological structure represents the vowels and the consonants represented in phonetic transcription. The syntactic level reflects the syntactic information such as the category as well as the case marking and the projection of the phrase structure. The semantic component illustrates how the syntactic structure is represented into semantic functions (GO, CAUSE, BE...), event structure (property, path, situation...), argument structure (object1, object2...). To explain this semantic level in some details, verb *mararatu* in (13a) involves movement and therefore it is associated with GO as a semantic function, while in (13b) the implied verb is *?arsil* ‘send’ as asSayuutih indicates (1992, vol. 3, p. 20). *send* is associated with CAUSE as a semantic function. Furthermore, the verb *mararatu*, represented semantically as Go, has an agent theta-role which is the 1st pronoun, *I*. This agent is an object of GO. As for the verb CAUSE in (13b), it is associated with two theta-roles: agent rule and a theme. The agent is object *you* and the theme is the object *alkilaaba*. The verb in (13a) involves a path trajectory represented by BY and the preposition has an object theta-role assigned to \dot{A} ABD which in turn has an object theta-role given to ALLAH. The adjective *s=aalih* is represented semantically as a property of the noun \aleph abd. However, the *Cause* verb in (13b) involves an inchoative event that takes a state *Be* event with an extension path trajectory. The phonological, syntactic, and semantic levels are linked together by means of correspondence rules in (11). To illustrate, the Head rule maps every syntactic head to a semantic function. For example, in (13a), the syntactic heads: V, N (pronoun *I*), P, N (\dot{A} abd), N (Allah), Adj are mapped respectively to these semantic functions: GO, Object, BY, Object, Object, Property. Moreover, the Argument/Modifier rule applies as the syntactic constituents map to semantic arguments and modifiers. So, the pronoun *I*, N (\aleph abd), N (Allah) are the semantic arguments of GO, BY, and \aleph abd respectively. The syntactic PP *bi \aleph abdi llaahi assaaliha* is a modifier and is mapped to a semantic modifier or a path. In addition to these two rules, co-indexation rules connect the syntactic and semantic levels. More specifically, every lexical item in (13) has a coindexed bits of syntactic, phonological, and semantic information represented by a co-indexation subscript symbol so that all these levels are mapped together in one lexical item (Jackendoff 2018; Culicover and Jackendoff 2005). To take one example, as we can observe in (13a), the head verb has a co-indexation symbol *P* for all levels: $m \leftrightarrow r \leftrightarrow r_t u$: p (at the phonological level), V_p (at the syntactic level), GO_p (at the semantic level). This can be also extended to other lexical categories in (13).

Now regarding the (irregular) accusative case assignment for the adjective *assaalihas*, we can simply consider

the accusative case as part of the syntactic information structure without the need to impose any implicit case assigner as has been proposed by Arab grammarians¹⁰. Interestingly, the adjective *assaaliha* can also be assigned a genitive case¹¹. The question that we need to ask here is about the nature of the genitive case. That is, is the genitive case regularly assigned by means of syntactic rules? Or is it irregularly assigned as the accusative case? To answer these questions, we can suggest that the genitive case, unlike the accusative case, is regularly assigned by a syntactic case rule. As a result, the adjective *assaalih* is an appositive and thus follows its antecedent \aleph abdi in example (12a) and hence picks up its genitive case marking. Now do we have to specify this regular case rule as part of the syntactic information for the sentence? No, we do not have to list this syntactic information as it can be derived syntactically by a regular rule. As for *alkilaaba* in (12b) it is assigned an accusative case despite the lack of an overt case assigner. Again, we assume that such case is part of the lexical syntactic information with which this sentence is specified for. The CG theory therefore can handle the fixed syntactic structures in ways that the mainstream theory cannot. In one hand, CG provides a natural explanation of the irregular case assignment that is semantically based while the mainstream theory sweeps such data under the rug and ignores them. Moreover, CG can also explain that the phrase is not derived through syntactic merger. Instead, the fixed phrase structure is lexical and hence is associated with specific meaning. For example, we may not produce a structure in Arabic like *the dogs on horses* equivalent to *the dogs on cows*. On the other hand, CG does not marginalize any syntactic data whether regular or not. CG recognizes “core” data constrained by general rules of the mainstream theory. At the same time, the CG addresses semi-productive data by another set of rules. That is why Culicover (1999) argues that the acquisition of human language requires two different learning systems, one for the regular data and another for the irregular data.

CONCLUSION

This paper discussed the problems of fixed lexical phrases to the traditional Arabic syntactic theory as well as the Chomsky’s mainstream grammar as represented by the Government and Binding (GB) (Chomsky, 1981) and the Minimalist theory (MP) (Chomsky, 2015). What makes the Arabic data in section (2) problematic is that the general rules of syntax assumed in either the traditional Arabic syntactic theory or the mainstream theory cannot provide an adequate explanation. I argued that the Arabic data are lexical phrases and I have explained that such analysis should not come as a surprise since there have been studies within the mainstream tradition like Hale and Keyser (1993) that has suggested that some syntactic structures are indeed derived lexically. I also argued that the Construction Grammar provides a solution that does not rely on syntactic structure alone but rather it links phonological, syntactic, and semantic information together in one basic construction by means of some correspondence rules. The Arabic data proved that there is a strong need for a linguistic theory that takes into consideration all data of different range of productivity.

Abbreviations

F	feminine
M	masculine
S	singular
Du	dual
P	plural
Nom	nominative
Acc	accusative
Gen	genitive
Per	person
1	first
2	second
3	third
Poss	possessive
Pro	pronoun
Ext	extension
Inch	inchoative
Def	definitive
Adj	adjective
N	noun
V	verb
Prep	proposition

ENDNOTES

- Sibawayhi (1988, Vol 2, p. 62-63) observes that *rabb* can be syntactically associated with two more different case markings: a genitive case as a modifier adopting the genitive case of its antecedent *Allahi* and a nominative case as *mubtada*.
- Sibawayhi suggests that there are semantic conditions for praise and hence the noun or the adjective can be assigned an accusative case. That is, not every situation or quality is worthy of praise but only those that are proven to be. For instance, Sibawayhi provides the following sentence (1988, Vol 2, p. 69):
 - mararatu bi-ʔabdi Allahi assaalih-a. passed. 3sm-1st pro.nom by Abdullah-gen the righteous.3sm-acc I passed by Abdullah the righteous.*
This sentence is acceptable only if Abdullah is known to people by his righteousness; otherwise, if he is not, then no accusative case is assigned to *assaalih-a*. Therefore, case marking is not determined by a regular syntactic rule of case but it is rather semantically-based.
- Sibawayhi suggests that there are semantic conditions for praise and hence the noun or the adjective can be assigned an accusative case. That is, not every situation or quality is worthy of praise but only those that are proven to be. For instance, Sibawayhi provides the following sentence (1988, Vol 2, p. 69):
 - mararatu bi-ʔabdi Allahi assaalih-a. passed. 3sm-1st pro.nom by Abdullah-gen the righteous.3sm-acc I passed by Abdullah the righteous.*
This sentence is acceptable only if Abdullah is known to people by his righteousness; otherwise, if he is not, then no accusative case is assigned to *assaalih*. Therefore, case marking is not determined by a regular syntactic rule of case but it is rather semantically-based.
- This is the translation of Abdelkader Fassi Fehri (2009, p. 120).
- NP here stands for the object that the verb selects, e.g., *take Ali for granted*.
- It is not my intension in this paper to evaluate Hale and Keyser's analysis of denominal verbs as much as to report their view of lexical phrases.
- Lexical phrases in general is a cover term that applies to morphologically complex words like collocations and fixed expressions. As argued in (Al-Dobaian, 2017), collocations are derived by a lexical merger.
- The meaning of a syntactically merged phrase is drawn from the combined units of the structure.
- I will only consider the semantic-syntactic interaction since it is the focus of this paper.
- In order to justify the accusative case assignment to the adjective *assaalih* in (12), Sibawayhi suggests an implicit case assigner *ʔadkur* that assigns the case to the adjective (1988, Vol 2, p. 70).
- See Sibawayhi (1988, Vol 2, p. 77) for a similar discussion.

REFERENCES

- Al-Dobaian, A. (2017). Collocations in Generative Theory. *Journal of Arts, Literature, Humanities, and Social Sciences*, issue 14, 38-61.
- Alfarahiidi, Kh. A. (1985). *Kitaab aljurnal fii annahw. taḥqiiq* by Fakharuddin Qabaawah. Beirut, Lebanon: Alrisaalat Publishing Company.
- Allwood, J, L.-G. Andersson, & o_. Dahl. (1995). *Logic in Linguistics*. Cambridge: Cambridge University Press.
- Assyuuttii, J. (No Date). *Huma Alhuma*. Abdulhamiid Hendaawi (ed.), vol. 1 and vol. 2, Egypt: alMaktabatu al-Tawfiqiyyah.
- AzZubii, A. (2012). Attaraakiib althaabitah fii allughati alarabiiyyati alfusha fii baabi almafaiil bayana alnithami al-lughwayy wa althakirati allughwayyati. *Journal of Damascus University*, Vol. 28, no. 1, 133-172.
- Booij, G. (2013). *Construction Morphology*. Oxford: Oxford University Press.
- Chomsky, N. (1981). *Lectures on Government and Binding*. Dordrecht: Foris.
- Chomsky, N. (1986). *Knowledge of Language*. New York: Praeger.
- Chomsky, N. (2015). *The Minimalist Program. 20th Anniversary Edition*. Cambridge, MA: MIT Press.
- Culicover, P. (1999). *Syntactic Nuts: Hard Cases in Syntax*. Oxford: Oxford University Press.
- Culicover, P. and Jackendoff, R. (2005). *Simpler Syntax*. Oxford: Oxford University Press.
- Daif, Sh. (1986). *Tahqiiq kitaab arradd ʔalaa annuḥaat*. Cairo: Dar almaʔarif.
- Fassi Fehri, A. (2009). *A lexicon of Linguistic terms*. Beirut: Dar alkitab aljadiid.
- Goldberg, A. (2006). *Constructions at Work. The Nature of Generalization in Language*. Oxford: Oxford University Press.