Journal of Arabic and Human Sciences | **E-ISSN : 1658-9602 |**
Qassim University, Volume (1), Issue (4), pp. 1-26, (Shawwal 1444 AH / April *2023 (Public Volume 16)*

"دور علامات الإعراب في ترتيب عناصر الجملة في اللغة العربية واللهجة النجدية : مقارنة وتحليل"

د. عبدالرحمن بن عبدالله الثواب
قسم اللغة الانجليزية والترجمة، كلية اللغة العربية والدراسات الاجتماعية، جامعة القصيم
(قدم للنشر في ١٤٤٤/٦/٩ هـ وقبل للنشر في ١٤٤٤/٨/١٠ هـ ونشر في ١٤٤٤/١٠/١هـ)

ملخص البحث :

تعتبر علامات الإعراب أحد مظاهر التباين اللغوي بين اللغات عموما، فتتميز لغات مثل العربية الفصحى واللاتينية والألمانية والروسية بكونها توظف علامات الاعراب لتحديد الحالة الإعرابية لعناصر الجملة. بينما يوجد لغات أخرى لا تستطيع علامات الإعراب فيها أن تقوم بنفس الوظيفة وذلك لأنها لا تشكل جزءًا صرفيا أو اشتقاقيا فيها ومن ذلك اللغة الانجليزية والفرنسية وكذلك اللهجات العربية المتعددة. وقد رُبط توفر وتوظيف علامات الإعراب بخصائص لغوية لعل من أهمها المرونة في ترتيب عناصر الجملة. ونَسبت بعض أبحاث اللغة هذه الميزة اللغوية في العربية الفصحى إلى وجود علامات الاعراب. لكن وكما هو معروف فإن اللهجات العربية والتي ترتبط بالعربية الفصحى تتميز أيضا بأنها تقبل التنوع أو المرونة في ترتيب عناصر الجملة مع أن علامات الإعراب غير متاحة في هذه اللهجات ومن هنا يأتي السؤال الذي سيحاول البحث الإجابة عليه. حيث يسعى هذا البحث إلى معرفة فيما إذا كانت علامات الإعراب هي السبب خلف تميز اللغة العربية بالمرونة في ترتيب عناصر الجملة، ولإجابة هذا السؤال يقارن ما تقوم به علامات الإعراب مع غيرها من العوامل اللغوية ذات

الصلة من ناحية إمكانية تحديد الحالة الاعرابية للمعمول وأثر ذلك في فهم ترتيب عناصر الجملة، ويركز البحث على مقارنة علامات الاعراب مع الخصائص الدلالية للفاعل والمفعول به في هذا الجانب. وتتضمن جوانب هذه المقارنة وصفا لغويا لجمل وأمثلة من العربية الفصحى واللهجة النجدية يتضح من خلالها أن كلتا اللغتين تمتازان بخاصية المرونة في ترتيب عناصر الجملة مع أنهما تختلفان تماماً من ناحية توظيف ووجود علامات الاعراب، ويشمل البحث على تحليل نظري لهذه الأمثلة والمقارنات عن طريق نظرية 'قواعد تركيب الجملة المبنية على العامل' (HPSG) وكذلك 'نظرية الربط' (Linking Theory)، ويمكن تلخيص ما يتوصل إليه البحث بأن علامات الاعراب قد تساعد في تحديد الحالة الاعرابية ولكنها ليست السبب خلف تميز اللغة العربية بالمرونة في ترتيب عناصر الجملة لأن هذه خاصية لغوية عميقة تتداخل مع عدة عوامل وليس من الانصاف أن تُنسب لعاملٍ لغوي واحد.

كلمات مفتاحية: علامات الإعراب، اللغة العربية، اللهجة النجدية، ترتيب عناصر الجملة، نظرية الربط

The Role of Case Marking in the Word Order of Standard Arabic and Najdi Arabic: Comparison and Analysis

Dr. Abdulrahman Althawab
Department of English Language and Translation, College of Arabic and Social Sciences, Qassim University, Saudi Arabia
Received on 9-6-1444 AH Accepted on 10-8-1444 AH Published on 1-10-1444 AH

Abstract
Case marking is a linguistic system through which inflectional marks or endings are added to lexical items like nouns or adjectives. It is widely known that languages differ in this regard. Case marking is overt in many languages, such as Standard Arabic (SA), Latin, German, and Russian. It is employed in such languages to distinguish between syntactic categories, particularly nominal arguments in terms of their grammatical relations (i.e., subjects and objects). Such languages also have a more flexible word order compared with languages that do not have the system of overt case marking, like English and French. A relation of cause and effect has been established here between case marking and the flexibility of word order. Nevertheless, certain varieties of Arabic, such as Najdi Arabic (NA), have a flexible word order too, even though they lack case markings which triggers the main research question of this paper. The paper aims to check whether case marking is really the linguistic factor behind the flexibility of word order in Arabic. To answer this question, the paper compares and contrasts the role of case marking to the role of semantic properties in identifying the grammatical relations of arguments. It also refers to 'agreement' and other relevant linguistic factors. The paper also explores the word order in NA and compares it to that of SA. In addition to presenting and describing fairly sufficient data from these two languages, the paper includes a theoretical analysis of the data within the framework of Head-driven Phrase Structure Grammar (HPSG) and Linking Theory. The conclusion reached in this paper is that case marking may play a role in determining the word order in SA, but attributing the flexibility of word order in Arabic to case marking is possibly a misleading claim.

Keywords: word order, case marking, Arabic, grammatical relations, HPSG

1. Introduction

Various aspects of language variation can be seen, including asymmetry in word order. Some languages follow a restricted scheme or system of word order, while others have more flexibility. Linguistically speaking, word order concerns the linear order of the syntactic constituents within a sentence. Moravcsik (2006) points out that the ultimate goal of the word order theory is to find the right properties that determine the linear asymmetry between constituents. In some languages, word order is relatively fixed, as in English. In contrast, other languages allow for more flexible word order, which can be used to express various linguistic purposes such as topicalization or focus. It should be noted that almost all languages, including those with flexible word order, have one single default or unmarked word order. The other orders are considered marked. Word order in constitutive sentences, as Dryer (2005) argues, is typically identified by a finite verb (V) combined with two arguments, namely, a subject (S) and an object (O). This means that identifying the grammatical relations of arguments results in determining the word order of a sentence.

There is a common observation in the typology of human languages that languages in which the NPs with different grammatical functions are marked by case exhibit more freedom in the order of the constituents than languages in which case marking is not used to indicate the NP's grammatical function (Abdul-Raof, 2013). Since these languages, like Standard Arabic (SA), German, and Russian, use case marking as an integral means to clarify the grammatical relation of nominal arguments, it is widely assumed that case marking is the one behind the flexibility of word order in these languages (Blake, 2001).

Unlike case-inflected languages, some Western European languages, including English and French, have a relatively rigid word order and tend to use word order as a means of grammatical expression in which the subject and object are denoted by their position before or after the verb (Van, 2005).

Some languages, on the other hand, differ in that they have a flexible word order even though identifying the grammatical relation of their nominal arguments is not governed by any case marker but rather by some other linguistic factors or principles. In some languages, for example, the order of words is identified by a semantic property, such as animacy in which subjects display greater animacy than objects. In others, it is identified by pragmatic properties, such as topicality, focus, or information structure. In Navaho, for instance, the order of the subject and object depends on the relative position of each NP on the animacy hierarchy according to the following criteria: HUMAN > ANIMALS > (large > medium > small) > instances > natural forces > plants > inanimate objects > abstract notions (Comrie, 1989).

Similarly, some varieties of Arabic have a relatively free word order with no morphological endings or marks that can distinguish between the arguments. As an example, Najdi Arabic (NA) has a flexible word order even though it lacks a case marking system, which is often linked to the flexibility of word order in SA. If NA is a variety of Arabic that has a flexible word order even though it does not have a case marking system, how can one assume that SA has a flexible word order because of case marking? This raised question touches on the main research question of this paper, which is about whether case marking is the linguistic factor behind the flexibility of word order in Arabic.

To answer this question and understand the role of case marking and its relation to word order in Arabic, the paper compares and contrasts relevant data from SA and NA. It also investigates the role of other linguistic factors in determining word order, such as semantic and pragmatic properties. Moreover, the paper attempts to theoretically account for the data using the frameworks of both Head-driven Phrase Structure Grammar (HPSG) and Linking Theory.

The next sections proceed as follows. Section [2] introduces the theoretical framework of HPSG and Linking Theory. Section [3] presents an overview of word order and the role of case marking in SA in [3.1], whereas it introduces word order in NA in [3.2]. Section [4] discusses the role of semantic properties in identifying the

grammatical relations of arguments and compares it to that of case marking. Section [5] presents an HPSG analysis of examples showing the role of case marking in determining word order in [5.1], while [5.2] attempts to account for the examples illustrating the role of semantic properties here using HPSG and Linking Theory. Finally, section [6] concludes the paper.

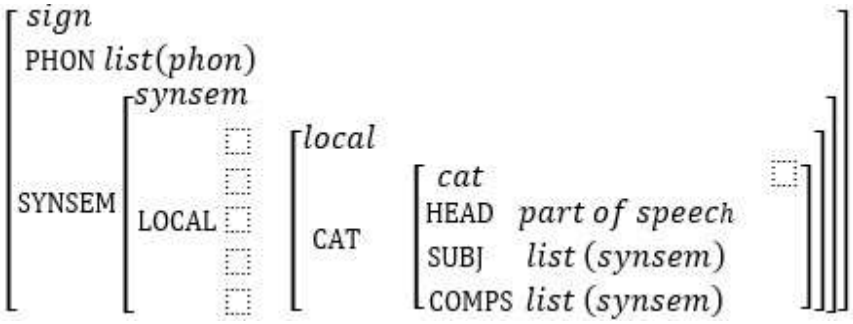
2. Theoretical Framework

2.1 HPSG

Head-driven Phrase Structure Grammar (HPSG) is a non-transformational approach to syntax. It is a constraint-based grammar developed by Carl Pollard and Ivan Sag in the mid-80s. It has a declarative and monostratal structure in which sentences have only one level of syntactic structure, so there are no 'move' processes (Borsley and Müller, 2021). HPSG, in general, is a complex system that includes sets of lexical and phrasal types within which constraints are imposed. Both types of lexical and phrasal signs contain syntactic, semantic, and phonological information. Thus, HPSG takes into account a great number of linguistic phenomena. According to this approach, any linguistic expression is governed by constraints, which are implicational statements. It follows that a linguistic object is well-formed if it adheres to all applicable constraints (Abeillé & Borsley, 2021).

An essential component of HPSG grammar is Attribute Value Matrix (AVM) through which the structure of a linguistic sign, whether lexical or phrasal, is represented. The AVM is made up of features that characterize each lexical entry. These features vary based on the sign type. The general structure of an AVM is partially shown in (1) (Sag et al., 2003).

(1) General structure of an AVM



This partial description of the structure of the AVM in (1) illustrates how each feature describes a type of information. For example, the feature HEAD includes intrinsic information as its value specifies, for instance, the sign's part of speech. The features SUBJ and COMPS are concerned with the head's syntactic selectional properties and the type of subject and complement(s) it requires. Both HEAD, SUBJ, and COMP fall within CAT(egory), which concerns syntax. The semantic features and their values are shown in CONT(ent) as will be discussed in the next section.

2.2 Linking Theory

Linking is the theory that captures the mapping of the semantic roles of arguments to the syntactic functions of the phrases that realize them. Within HPSG, one of the significant theories of linking is proposed by Davis and Koenig (2000). There, the linking properties of a verb depend on the particular semantic content contributed by the verb and on the verb's syntactic selectional properties. For example, the grammar of English indicates that the subject of *drink* fills the drinker role, and the object of *drink* fills the role of the thing drunk. This mapping is usually broken down into two simpler mappings by positing an intermediate representation called ARG-ST (argument structure). The first mapping in the linking theory connects the participant roles shown in the semantic CONTENT to the elements of the value of the ARG-ST feature. The second mapping is argument realization, which connects those ARG-ST list elements to the elements

of the valence lists (i.e., COMPS and SUBJ). The abbreviated lexical description of the verb *drink* in (2) illustrates these two mappings.

(2) A shortened lexical entry of the verb *drink*

$$\begin{bmatrix} \text{PHON} & drink \\ \text{SUBJ} & < [1] > \\ \text{COMPS} & < [2] > \\ \text{ARG-ST} & < [1]\text{NP}i, [2]\text{NP}k > \\ \text{CONT} & \begin{bmatrix} drink-relation \\ \text{DRINKER} & i \\ \text{DRINKEN} & k \end{bmatrix} \end{bmatrix}$$

There are general patterns assumed in the Linking Theory across verbs and languages. The main ones assume that if one argument of a transitive verb in the active voice has an agentive role, the argument will map to the subject. The argument with the undergoer role, on the other hand, will map it to the object (Davis et al., 2021).

3. Word Order and Case Marking in SA and NA

3.1 An Overview of Case Marking's Role in SA Word Order

The sentence word order in Arabic is determined by the subject position. If the subject appears preverbal, it typically gives an SVO order. By contrast, if the subject is placed post verbally, the order is normally VSO. As stated above, Arabic is one of the case-inflected languages in which grammatical relations are encoded in terms of morphological case marking. As can be observed in many languages, the subject of finite clauses in Arabic takes a nominative case. The object, on the other hand, is assigned an accusative case. Typically, the nominative case is overtly realized by the suffix **-u**, whereas the accusative case is realized overtly by the suffix **-a**. Bakir (1980) points out that SA has a free word order due to its rich case marking and inflectional morphemes. In fact, it is also widely believed that word order in Arabic is flexible due to the system of case-marking (e.g., see

Carnie and Fehri, 1996). The following sentences in (3) are examples of the VSO, SVO, and VOS word orders in SA.[1]

(3) a. qaraʔ-a al-muʕalim-**u** ad-dars-*a* (VSO)
read-3.SG.M the-teacher-NOM the-lesson-ACC[2]
"The teacher read the lesson."
b. al-muʕalim-**u** qaraʔ-a ad-dars- *a* (SVO)
the-teacher.M-NOM read-past.3.SG.M the-lesson-ACC
"The teacher read the lesson."
c. qaraʔ -a ad-dars-*a* al-muʕalim-**u** (VOS)
read-3.SG.M the-lesson-ACC the-teacher-NOM
"The teacher read the lesson."

As can be seen in (3), all these sentences are grammatical and imply almost the same meaning but with different word orders. In (3 a-c), *al-muʕalim-**u*** is assigned a nominative case since it is the subject; whereas *ad-dars-**a***, which is the object, receives an accusative case. Consider the examples in (4) to (6), which also illustrate the flexibility of word order in SA, and the NP arguments are of the same type (i.e., they are [+human, +feminine]).

(4) at-tilmiiðat-**u** ʔahda-t al-muʕalimat-*a* (SVO)
the-students.F-NOM (gave a present)-F the-teacher.F-ACC
"The student gave a present to the teacher."
(5) ʔahd-at at-tilmiiðat-**u** al-muʕalimat-*a* (VSO)
(gave a present)-F the-student. F-NOM the-teacher.F-ACC
"The student gave a present to the teacher."
(6) ʔahd-at at-tilmiiðat-*a* al-muʕalimat-**u** (VOS)
(gave a present)-F the-student.F-ACC the-teacher.F-NOM
"The teacher gave a present to the student."

[1] In the examples here and throughout the paper, **boldface** indicates the nominative case marking, and *italics* indicates the accusative one.
[2] Glosses will be added only when are necessary. This means that they may be omitted or shortened here and throughout the paper if they do not play a role in understanding the context.

In (4), *at-tilmiiðat-**u*** appears with the nominative case indicating that it is the subject, whereas *al-muʕalimat-**a*** receives the accusative case because it is the object. The same assignment of cases and meanings applies to (5), but with a different word order. In (6), *at-tilmiiðat-**a*** has the accusative case marking, hence it is the object. The other NP argument *al-muʕalimat-**u*** becomes the subject in (6). Thus, it is assigned the nominative case. Notice that (5) and (6) use the same sequence of words, but the case marking makes the former mean almost the opposite of the latter.

To sum up, this section shows that NP arguments in SA can normally be marked by cases. In addition, case marking can help identify the grammatical relations of arguments and, consequently, the word order of sentences.

3.2 An Overview of NA Word Order

Najdi Arabic, a variety of Arabic spoken in the Najd region of Saudi Arabia, has a flexibility of word order that is similar to that of SA. However, it lacks the system of overt case marking. Since Najdi speakers cannot rely on case marking to identify subjects and objects, one might ask here how they can distinguish the grammatical relations within sentences. This might also suggest that Arabic has another criterion for identifying grammatical relations since understanding such relations is crucial in any language.

To elaborate on this, let us discuss the NA examples in (7), which show that NA has a relatively flexible word order.

(7) a. gara **al-estað** **ad-dars** (VSO)
read_past.M.SG the-teacher.M.SG the-lesson.M.SG
"The teacher read the lesson"

b. **al-estað** gara **ad-dars** (SVO)
the-teacher.M.SG read_past.M.SG the-lesson.M.SG
"The teacher read the lesson"

c. gara **ad-dars** **al-estað** (VOS)
read_past.M.SG the-lesson.M.SG the-teacher.M.SG
"The teacher read the lesson"

Each of these sentences contains a transitive verb with two nominal arguments: *al-estað* 'the teacher' and *ad-dars* 'the lesson'. As indicated earlier in [3.1], SA normally depends on case marking to identify the subject and the object. This system of case marking, however, does not exist in NA. Hence, the available criterion for identifying grammatical relations in these sentences is the semantic properties of the arguments, more specifically 'animacy'. For instance, the subject in (7 a, b, & c) is *al-estað* 'the teacher' since the verb *gara* 'read' requires a human agent as a subject, not *ad-dars* 'the lesson', which is an inanimate object. Using semantic properties as a criterion for identifying the grammatical relations of arguments and how this criterion differs from that of case marking will be the concern of the next section.

4. The Role of Semantic Properties in Identifying the Grammatical Relations

4.1 An Overview

This section discusses employing the semantic properties and, to a lesser degree, the pragmatic properties of arguments to identify their grammatical relations following the linking concept. In short, this linking concept or pattern assumes that the subject of a transitive verb in the active voice, for instance, is an argument playing an agentive role, whereas the object in such constructions is an argument with an undergoer role. The following subsections, which are divided according to some of the semantic selectional properties of verbs, show how the concept of linking can help identify the grammatical relations of arguments.

4.2 [+Animate] Versus [-Animate]

The first case to be discussed is when a sentence contains two nominal arguments (NPs), one of which has the [+animate] feature but the other does not, and the verb requires an agentive [+animate] subject, as shown in (8), (9), and (10).

(8) ʔakala-t at-tifaħeh al-bent (VOS)
ate-F the-apple the-girl
"The girl ate the apple"
(9) ʃaʁʁala-t aʃ-ʃaʁaleh al-meknasah (VSO)
turned on-F the-maid the-vacuum cleaner
"The maid turned on the vacuum cleaner"
(10) Khaled lebas ʃemaʁ (SVO)
Khaled wore shemagh
"Khaled wore shemagh"

To know which one is the subject in the examples (8) to (10) (*a-tufahah* or *al-bent*; *aʃ-ʃaʁaleh* or *al-meknasah*; *Khaled* or *shemagh*), identifying the agent (i.e., the doer of the expressed action, event, etc.) is the first step since the verb in all these three examples requires an agentive subject. From a semantic perspective, it is known that *a-tufahah* 'the apple', as in (7), does not have the ability to eat since it is inanimate. Thus, it is obvious that *al-bent* 'the girl', who has an animate feature, is the agent here and, consequently, it is the subject. Similarly, *al-meknasah* 'vacuum cleaner' and *ʃemaʁ* 'shemagh' do not have an animate feature that enables them to wear or turn on something. As a result, the subjects in each sentence are the other nominal arguments with the animate feature. Given this, VOS is the word order of (8), VSO is the word order of (9), and SVO is the word order of (10).

4.3 [+ Human] Versus [+Animate]

The other case to be considered here is when the two NP arguments have the [+animate] feature. To see how to identify the subject in such cases, let us first discuss the following examples in (11), (12), and (13).

(11) ʔa-kkal al-kalb al-walad (VOS)
fed the-dog the-boy
"The boy fed the dog."
(12) Ahmad ðebaħ al-ʤerði (SVO)
Ahmad killed the- rat
"Ahmad killed the rat."

(13) ʃala-t Salma al-bessah (VSO)
held-F.SG Salma the-cat
"Salma held the cat."

In these sentences, there are two nominal arguments with [+animate] feature (i.e., *al-kalb* and *al-walad*; *Ahmad* and *al-ʤerði*; *Salma* and *al-bessah*). In such cases, another feature is needed to identify the agent because of the semantic selectional properties of the verbs here. The verbs *(ʔakkal* 'fed', *ʃalat* 'held', and *ðebaħ* 'killed') require a higher animate [+ human] NP as an agent. If we look at (11), for instance, the NP *al-kalb* 'the dog' is [- human] and has no control over the action denoted by the verb *(ʔakkal* 'fed'), which requires a higher animate [+ human] NP. Thus, the [+ human] NP *al-walad* 'the boy' will be the agentive subject here. The same also applies to (12) and (13), whose verbs require a [+ human] agentive subject: *Ahmad* and *Salma* respectively. We can also add that the subject of (12), for example, cannot be *al-ʤerði* 'the rat' since it is not reasonable for a tiny creature to kill a human. As stated before, identifying the subject requires first identifying the agent who can do the action or event expressed by the verb since the agent and subject are linked in the active voice as in these examples. Finally, identifying the subject as the [+ human] agents in these examples means that they have the following word order: VOS in (11), SVO in (12), and VSO in (13).

4.4 [+ Human] Versus [+ Human]

Another case that might have a semantic ambiguity is when the two arguments of the verb are [+ human] NPs. In such cases, identifying the grammatical relations may need the help of pragmatic context and world knowledge.[3] To elaborate on this, let us first see the examples (14), (15), and (16).

(14) rawwaʃa-t al-bent al-um (VOS)

[3] Based on Lakoff (1977), we can partially specify the semantic structure of sentences using our world or background knowledge, particularly in the absence of linguistic information.

wash-F the-girl the-mother
"The mother washed the girl."

(15) faħasˁ al-mari:ðˁ ad-dektor (SVO)
checked the-patient the-doctor
"The doctor checked the patient."

(16) χalaf aʃ-ʃertˁi as-sawag (VSO)
issued a ticket the-policeman the-driver
"The police officer issued a ticket to the driver."

By considering what a native speaker knows about notions like authority, typicality, and responsibility of these actions, it is possible to identify the agent even if the two arguments share the basic semantic properties as [+ human]. If we take (14) as an example here, a native speaker may need, in order to identify the agentive subject, to ask himself/herself: who typically takes care of the other? In light of world knowledge, it would be *al-um* 'the mother' who takes care of her daughter and is typically responsible for her; hence, she is the agent here. If we apply the same idea to (15), we will consider who is typically the one checking the other: *ad-dektor* 'the doctor' or *al-mari:ðˁ* 'the patient'. The same can also be applied to (16), and the question here will be who typically issues tickets or penalties: *aʃ-ʃertˁi* 'the police officer' or *as-sawag* 'the driver'.

However, there are cases where it seems implausible to realize the subject in languages that do not employ case marking and have flexible word order like NA, as in (17) and (18).

(17) ʃaf Mohammed Faisal (VSO)
saw Mohammed Faisal
"Mohammed saw Faisal."

(18) gabala-t Hind Amal (VSO)
met-F Hind Amal
"Hind met Amal."

It becomes almost impossible to identify which one is the agent by using our world knowledge in sentences like (17) and (18), where both NPs are of equal information salience as is the case with personal

names of the same gender, which do not include background information (such as age, relationship, etc.). When semantic properties, pragmatic properties, and world knowledge are unable to help identify which argument can be assigned a clear agent or theme role, the last resort here becomes syntax. In such instances, we need the help of syntactic means through which we can understand the grammatical relations of arguments. One of these means is the agreement in person, number, and gender between the verb and its subject provided that this agreement is overtly marked and helps distinguish between the arguments. This occurs when there is a contrast in the arguments in terms of their agreement features, as exemplified in (19).

(19) ʃafa-t Mohammed Farah (VOS)
saw-**F** Mohammed Farah
"Farah saw Mohammed."

Although *Farah* follows *Mohammed* in the word order of (18), the subject-verb agreement makes it obvious that Farah, which is a feminine name, is the subject because the verb has a feminine suffix.

Finally, if there are not any available syntactic, semantic, pragmatic, or world knowledge means through which we can identify the grammatical relations of arguments within a sentence, as in (17) and (18), the flexibility of word order becomes inactive. Hence, we have to assume that such sentences have the unmarked or default word order where the subject precedes the object.

4.5 Case Marking Versus Semantic Properties

After exploring semantic properties as a means of identifying the grammatical relations of arguments in NA, the current section attempts to answer this question: Can case marking override semantic properties in identifying the grammatical relations of arguments in Arabic? Consider the SA example in (20).

(20) !qaraʔ-a al-muʕalim-*a* ad-dars-**u**
read-past the-teacher-ACC the-lesson-NOM
"The lesson read the teacher."!

Following the approach which states that case marking is the means that identifies the grammatical relations, *ad-dars-u* 'the lesson' in (20) would be the subject since it is assigned a nominative case. On the other hand, the other approach, which takes semantic factors into account, considers *ad-dars-u* 'the lesson' as an object. This is so because the verb *qaraʔ-a* 'read' requires an agentive subject which has [+animate] feature, whereas *ad-dars-u* 'the lesson' is [-animate]. Therefore, this approach assumes that *al-muʕalim-a* 'the teacher' is the subject. This also implies that *al-muʕalim-a* in (20) does not have the right inflected ending that marks the nominative case (i.e., it should have the *-u* ending). This example actually shows that relying solely on case marking is not sufficient to identify the grammatical relations of arguments and that case marking can sometimes be misleading. Native speakers of Arabic may agree that slips of the tongue or making mistakes in the case markers is not uncommon, even in formal contexts.

To support this assumption, Abdul-Raof (2013) states that Arab grammarians' account is simplistic and that it is quite misleading to use overt case marking as a criterion of subjecthood in the identification of the underling grammatical relations in Arabic.

Additionally, the other piece of evidence that supports the argument here comes from NA. We have seen many sentence examples in this section as well as the previous section, which show that word order is flexible in NA even though it does not employ the system of case marking.

Another relevant point that should be addressed here is that case marking is not always overt in SA, as in the examples (21) and (22).

(21) qabal-t Ruba Muna
met.F Ruba Muna
"Ruba met Muna."

(22) sˤalla aðˤ-ðˤuħa Musa
prayed the-ðˤuha Musa
"Musa prayed Forenoon/Morning Prayer."

In brief, the two NP arguments in (21), which are female names, end in a vowel, which prevents the case marker from being overt here. Since case marking and the other relevant factors discussed earlier are not available here, we need to assume that the word order is in its default value and that the subject precedes the object. In (22), the situation is almost the same in that the two NP arguments cannot have overt case endings. However, we can easily understand that the word order is marked in that it is VOS. This is so because the semantic properties can help identify the grammatical relations of the two arguments here. The NP immediately following the verb, which is *aðˁ-ðˁuħa* (the Morning Prayer), is [-human] and even [-animate]. Therefore, it cannot pray and be the subject, which is the [+human] Musa. This example is a clear case where semantic properties are the only available factor that can identify the word order and the grammatical relations of arguments.

Given this, it seems that case marking is perhaps not the critical factor for identifying the grammatical relations of nominal arguments in Arabic, and that semantic properties, for instance, may override case marking in realizing these relations. This goes in line with the view of Hallberg and Niehorster (2021), who did an experimental study on Arabic case marking. They consider the markers of case an optional feature that does not directly affect how Arabic speakers identify the grammatical relations of arguments. They also argue that "orthographically marked case is a morphological feature that (a) provides syntactic information that is redundant for comprehension; (b) is only occasionally available; (c) is not represented in speakers' native variety; and (d) is not mastered by most skilled readers" (p. 32).

5. Analysis

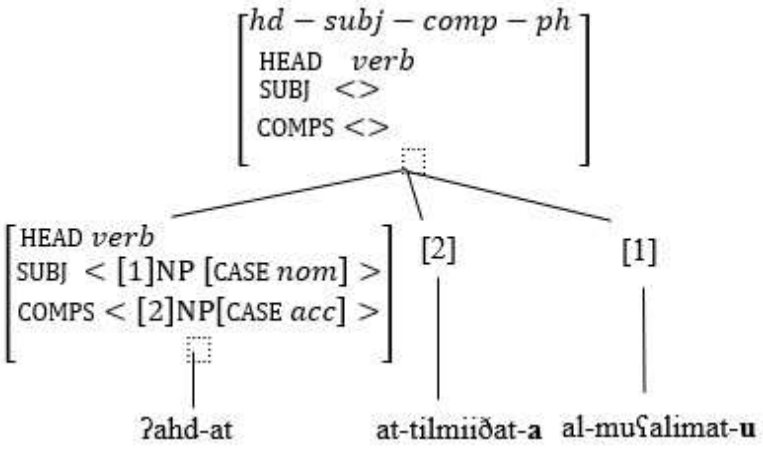
5.1 Case Marking

Arabic has a structural case in which an NP argument typically receives nominative case if it is the subject of a verb and accusative case when it is the object. Applying this in HPSG is quite straightforward. The SUBJ will contain the specification in (23), whereas the COMPS will contain the one in (24).

(23) [SUBJ [CASE *nom*]]
(24) [COMPS [CASE *acc*]]

On this approach, the sentence given earlier in (6) will have the structure in (25) where the SUBJ and COMPS features indicate what sort of subject and object the head requires.[4]

(25)



5.2 Semantic Properties

It is widely assumed in the field of syntax-semantic interface that there are two types of semantic roles for arguments. The first role is actor, and it concerns subject arguments. It is named 'actor' since it describes the role of the argument who is doing or responsible for the action, event, and so forth described by the verb. The actor role includes, among others, the roles of agent, experiencer, and instrument. The second role, which is for object argument, is 'undergoer', which expresses how the object argument receives the

[4] It should be noted here that this sentence is analyzed in HPSG as a head-subject-complement phrase where both the subject and complement are sisters to the head. This analysis is initially adapted from English auxiliary-inverted constructions like *can you leave?*. For more about this, see, for instance, Borsley (1995) and Althawab (2022).

action, or the change of state expressed by the verb. It subsumes roles such as patient, theme, and recipient (Wunderlich, 2006).

Using these two types of roles and their subtypes, HPSG through the linking approach can show how the syntax-semantics interface be activated and employed to illustrate interesting linguistic phenomena like the one under discussion here. In HPSG, the semantic roles and the syntactic or grammatical relations of the arguments can be linked using the relevant attributes and values, as the two constraints in (26) and (27) show (Davis et al., 2021).

(26) $\begin{bmatrix} \text{CONT}|key & [\text{ACT}\,[1]] \\ \\ \text{ARG}-\text{ST} & < \text{NP}[1], \ldots > \end{bmatrix}$

(27) $\begin{bmatrix} \text{CONT}|key & [\text{UND}\,[2]] \\ \\ \text{ARG}-\text{ST} & < \ldots, \text{NP}[2], \ldots > \end{bmatrix}$

The first constraint in (26) shows two values and a constraint on these two values. The first value is a semantic one in that it is presented in the attribute CONT, which concerns the relevant semantic attributes and values. This value identifies the semantic role of the argument tagged [1] as ACT (actor). The second value is shown in ARG-ST (argument-structure), whose value lists the arguments of the head in which the first value is the subject and what follows are the complements. The subject here is also tagged [1]. What concerns us more is that the constraint in (26) links these two values together by using the same tag [1], and the linking here means they are identical (i.e., the NP subject here is the actor). The same explanation applies to the constraint (27), which simply states that the semantic role of the NP object is of the type undergoer. The linking in these two constraints is due to tagging or using the same tag. It should be noted also that, as introduced earlier in [2.2], indexing can be used to

express linking, as exemplified in (28) in which the information of the two constraints is combined.

(28)
$$\begin{bmatrix} \text{HEAD} & verb \\ & \text{SUBJ} & < \text{NP}i > \\ & \text{COMPS} & < \text{NP}k > \\ & & \square \\ \text{CONT} & \begin{bmatrix} verb-relation \\ \text{ACTOR} & i \\ \text{UNDERGOER} & k \end{bmatrix} \\ & & \square \end{bmatrix}$$

Let us now apply these two constraints to an example and see how they can be used as a means for identifying the grammatical relations of arguments. The example that we will use here is the lexical description of the verb *ʃal* 'held' shown in (29).

(29)
$$\begin{bmatrix} \text{PHON} & \int al \\ \text{SUBJ} & < [1] > \\ \text{COMPS} & < [2] > \\ \text{ARG-ST} & < [1]\text{NP}i, [2]\text{NP}k > \\ \text{CONT} & \begin{bmatrix} hold-relation \\ \text{HOLDER} & i \\ \text{HOLDEN} & k \end{bmatrix} \end{bmatrix}$$

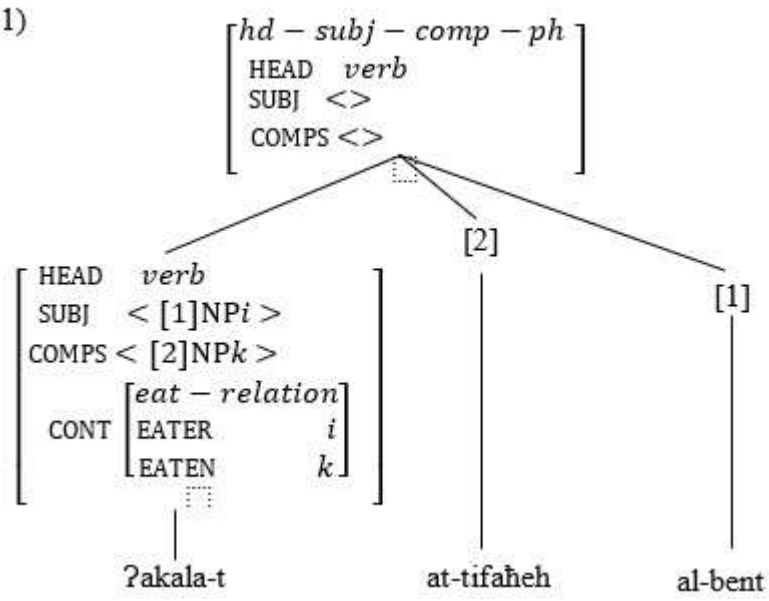
As shown in (29), the semantic role HOLDER, which is an instance of the ACTOR, is linked through co-indexing to the first element of ARG-ST (i.e., the tag[1]). In addition, the semantic role of HOLDEN, which is in turn an instance of UNDERGOER, is linked to the second element in ARG-ST, which is here the NP complement or object of the verb *ʃal* 'held', and it is tagged [2].

As a conclusion of this section, let us apply this linking approach to a sentence example from NA in a marked word order, namely VOS, where agreement does not play a role in identifying the grammatical relations of arguments (i.e., the two arguments share the

same agreement features). Then, we will see how case marking intervenes here when we use the same sentence but this time in SA. Applying the HPSG linking approach to the NA sentence in (30) will give it the structure shown in (31).

(30) ʔakala-t at-tifaħeh al-bent
ate.F the-apple the-girl
"The girl ate the apple."

(31)

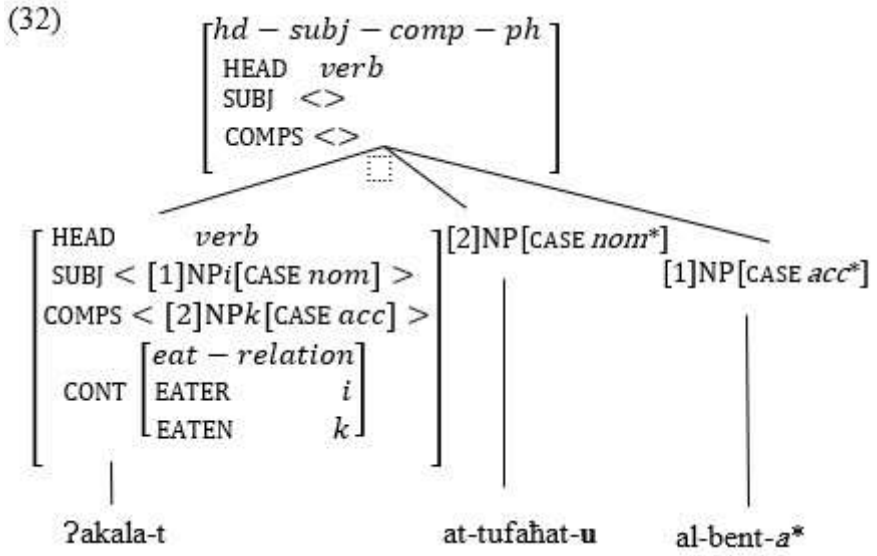


The structure in (31) shows that the subject of the verb *ʔakala-t* 'ate' is the NP *al-bent* 'the girl'. Since it is the subject and the voice here is active, it has the semantic role of EATER. The structure also shows that the object here is the NP *at-tifaħeh* 'the apple' which has the role of EATER. Although the word order here is VOS, which is not the unmarked one in NA, the right matching between the arguments and their semantic roles through the linking approach enables us to identify the grammatical relations of arguments in a clear way.

The structure of the same sentence in (30) but in SA will be the same as (31) except that we may need to add the attribute or feature

CASE to the argument NPs and identify the value as *nom* for the subject and *acc* for the complement. As already discussed, this is so because SA, unlike NA, has overt case marking and, therefore, employs the case in identifying grammatical relations. One might ask here what if we mark the case incorrectly (i.e., the subject has an accusative case marker and the object has a nominative one). If we do so, the structure will be as in (32).

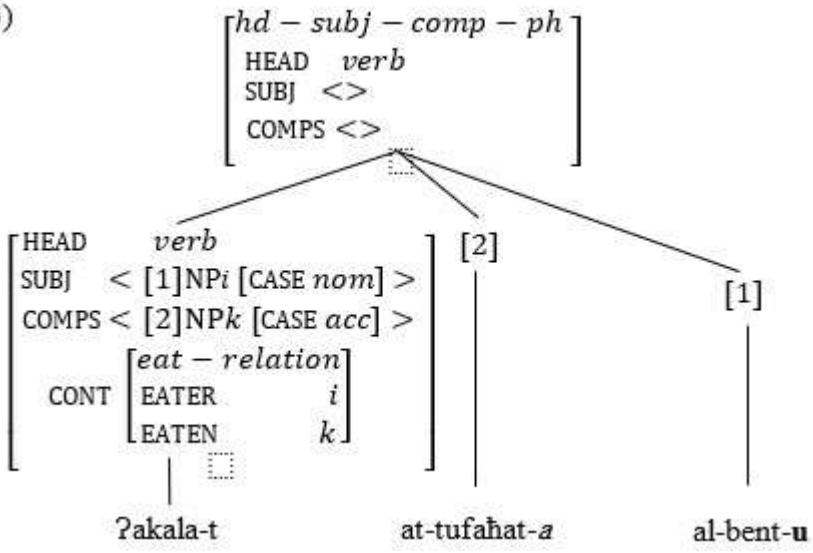
(32)



As shown in (32), the structure of this SA sentence is almost the same as that of the NA given earlier in (31) except the information about case. In HPSG, the notion of under-specification is assumed, which requires the necessary information to appear in the structure and assumes that everything else is in its default value or setting and, for the purpose of succinctness, it is not necessary to include such 'default' information. For example, the information about case is shown in (32) although such information is mostly not included in HPSG structure. It appears here because there is a mismatch between the case requirements shown in the description of the head verb *ʔakala-t* 'ate' and the case of the NP subject and object. The head verb here, as is the case with lexical verbs in general, requires its subject to be nominative and the object to be accusative. Since the case endings

of these two NP arguments do not meet this requirement, the information of their case is marked with an asterisk (*). This implies that resolving this mismatch will fulfill the requirements of the head. In HPSG terms, this will satisfy the constraints listed in the information of the head. Consequently, the sentence will become well-formed in this VOS word order, as in (33).

(33)



The argumentation presented here about the structure of (31), (32), and (33) supports the view that case marking cannot override semantic properties in identifying the grammatical relations of arguments in linguistically well-formed sentences, which, in turn, implies that case marking is not the primary factor behind the flexibility of word order in Arabic. This goes in line with our common sense in that it seems linguistically more reasonable to assume that the structure in (32), for instance, has incorrect case marking than to consider *at-tufaħat-u* 'the apple' as a possible actor or eater here. The semantic properties are intrinsic linguistic properties of the lexical items, while case markers are not.

6. Conclusion

This paper investigates the role of case marking in determining word order in SA and NA. It questions the widely held view that case marking is the linguistic factor behind the distinct flexibility of word order in Arabic. As mentioned throughout the paper, this view is sometimes presented as a fact, especially in introductory and tutorial references to Arabic.

The paper shows in a data-based approach that this view is a bit misleading. Through describing and analyzing various examples within the framework of HPSG and the linking approach, the paper compares the role of case marking to the role of semantic properties in determining the word order of Arabic. Another comparison is also made between SA and NA in this regard to check the credibility of this view.

The conclusion that has been reached here can be summarized in one sentence by saying that case marking is not the reason behind the flexibility of word order in Arabic. Such flexibility is a distinct and complex property of Arabic that cannot be attributed to one factor. In addition to case marking, word order in Arabic interacts with other factors, including syntactic agreement, semantic properties, pragmatic properties, and world-knowledge. Hence, case marking can be described as one of the factors that can help identify word order in SA only (i.e., not in the other Arabic varieties) provided that it cannot be overridden by any other linguistic factors, particularly in sentences which are meant to be linguistically well-formed. Its role becomes more obvious when the other relevant factors are not available. It also seems more reasonable to assume that case marking marks the case of arguments (i.e., through the endings of these nominal arguments) than to claim that it assigns case to this or that argument.

References

[1] Abdul-Raof, H. (2013). *Subject, theme, and agent in Modern Standard Arabic*. Routledge.

[2] Abeillé, A., & Borsley, R. (2021). Basic properties and elements. In S. Müller, A. Abeillé, R. D. Borsley, & J. Koenig (Eds.), *Head-driven phrase structure grammar: The handbook* (pp. 3-45). Language Science Press.

[3] Althawab, A. (2022). The Syntactic Structure of an Introductory PP in Standard Arabic: A Non-Transformational Approach. *World Journal of English Language, 12*(8), 242–251. http://dx.doi.org/10.5430/wjel.v12n8p242

[4] Bakir, M. J. (1979). *Aspects of clause structure in Arabic: A study in word order variation in literary Arabic*. Indiana University.

[5] Blake, B. J. (2001). *Case*. Cambridge University Press.

[6] Borsley, R. D. (1995). On some similarities and differences between Welsh and Syrian Arabic. *Linguistics 33*, 99-122.

[7] Borsley, R., & Müller, S. (2021). HPSG and Minimalism. In S. Müller, A. Abeillé, R. D. Borsley, & J. Koenig (Eds.), *Head-driven phrase structure grammar: The handbook* (pp. 3-45). Language Science Press.

[8] Carnie, A., & Fehri, A. (1996). Issues in the structure of Arabic clauses and words. *Language*, *72*(3), 664-665. https://doi.org/10.2307/416310

[9] Comrie, B. (1989). *Language universals and linguistic typology: Syntax and morphology*. University of Chicago press.

[10] Davis, A. R., & Koenig, J. P. (2000). Linking as constraints on word classes in a hierarchical lexicon. *Language*, *76*(1), 56-91.

[11] Davis, A. R., Koenig, J. P., & Wechsler, S. (2021). Argument structure and linking. In S. Müller, A. Abeillé, R. D. Borsley, & J. Koenig (Eds.), *Head-driven phrase structure grammar: The handbook* (pp. 315-368). Language Science Press.

[12] Dryer, M. (2005). Order of subject, object and verb. *The world atlas of language structures*, 330-333.

[13] Hallberg, A., & Niehorster, D.C. (2021) Parsing written language with non-standard grammar. *Read Writ 34*, 27–48.

[14] Lakoff, R. (1977). What you can do with words: Politeness, pragmatics and performatives. In *Proceedings of the Texas conference on performatives, presuppositions and implicatures* (pp. 79-106). Arlington, VA: Center for Applied Linguistics.

[15] Moravcsik, A. (2006). What can we learn from the collapse of the European constitutional project? *Politische Vierteljahresschrift, 47*(2), 219-241.

[16] Sag, I., Wasow, T., & Bender, E. (2003). *Syntactic theory: A formal introduction* (2nd ed.). CSLI Publications.

[17] Van, R. (2005). Syntactic relations and case marking. In R. Valin (Ed.), *Exploring the Syntax-Semantics Interface* (pp. 89-127). Cambridge University Press

[18] Wunderlich, D. (2006) Argument hierarchy and other factors determining argument realization. In Bornkessel Ina, Matthias Schelesewsky, Bernard Comrie, and Angela D. Friederici (Eds.). *Semantic role universals and argument linking*: *Theoretical, typological, and psycholinguistic perspectives* (pp. 15-52). Mouton de Gruyter.