Subject-Verb Agreement in Arabic from a Sign-Based Construction Grammar Perspective: A Sentence Types View

DOI: https://doi.org/10.33806/ijaes.v24i2.510

Rukayah Alhedayani
King Saud University, Saudi Arabia

Received: 23.8.2023 Accepted: 15.1.2024 Published: 6.6.2024

Abstract: Arabic sentences exhibit both SVO and VSO word orders, and these word orders show different agreement patterns between the subject and the verb. This agreement asymmetry has been the subject of research for a long time. This paper adds to the body of literature regarding this topic by presenting a Sign-Based Construction Grammar account of the phenomenon. It differs from previous research in that it accounts for all sentence types in Arabic and not the ones that include a verb. To achieve this goal, a sentence type view is adopted instead of a word order view. This paper shows that adopting a sentence type view instead of a word order view allows for a better understanding of Arabic sentences and paves the way to accounts of other languages exhibiting the same phenomenon.

Keywords: agreement, Arabic, sentence types, Sign-Based Construction Grammar, word order

1. Introduction

Agreement between the verb and the subject in Arabic sentences is affected by many factors, such as word order and the type of the subject (Jarrah, Rayyan, Al-Shawashreh and Zuraikat 2020). Arabic sentences exhibit both SVO and VSO word orders. In SVO word order, as in sentence (1) below, the verb shows full agreement with the subject in gender and number. However, in VSO word order, as in sentence (2) below, the verb shows gender agreement only.

(1) al-walad-u ʔakala ʔa-ttuffāħat-a
the-boy.S.M-nom. ate.S.M the-apple-acc.
The boy ate the apple.

(2) akala ʔal-ʔawalād-u ʔa-ttuffāħ-a
ate.S.M the-boys.P.M-nom. the-apples-acc.
The boys ate the apples.

Subject-verb agreement in Arabic is widely researched (Aoun, Benmamoun and Sportiche 1994; Aoun, Benmamoun and Choueiri 2009; Larson 2013; Albuhayri 2019; Jarrah et al. 2020) among others. Most linguists follow the Greenberg view of word order (Greenberg 1966). However, agreement in Arabic is affected by more than just word order (Peled 2008). The type of the constituent in the sentence plays a role in agreement, too; for example, whether the subject is a conjoined phrase or not and what type of phrasal structure the sentence has. Therefore, a wider view would be one that includes all factors affecting agreement looking at sentence types not word order. There are two widely studied sentence types in Arabic[1]. These two

sentence types are often represented by sentences similar to the ones in (1) and (2) above. However, it is important to mention that Arabic allows verbless sentences which are traditionally included with sentence (1) as of the same type. Therefore, Arabic sentences can be divided into two main types, one that begins with a verb and one that begins with a noun regardless whether it contains a verb in the predicate or not. Most work on agreement in Arabic was done following Government and Binding and the Minimalist Program theories (Chomsky 1957, 2005). The present paper, however, uses a novel theory to account for this phenomenon, namely Sign-Based Construction Grammar (SBCG) (Sag 2012).

A quick search in the arTenTen12 corpus (Arts, Belinkov, Habash, Kilgarriff and Suchomel 2014) revealed that the VSO structure is three times more frequent than the SVO structure. Verb initial sentences in the arTenTen12 tagged by the Stanford tagger (Toutanova, Klein, Manning and Singer 2003) yielded over eleven million hits. However, a search of noun initial sentences - that contained a verb - in the same corpus yielded a little over three million hits. This shows how frequent the VSO word order is compared to the VSO word order. The table below shows the number of hits and percentages per million tokens.

Frequency of VSO/SVO sentences in the arTenTen12 corpus

Sentences	No. of hits	Per million token	percent
VSO sentences	11,437,417	1,374.34	0.14
SVO sentences	3,170,133	380.93	0.038

This search shows that when there is a verb in the sentence in Modern Standard Arabic (MSA), it is preferred to be at the beginning (i.e. the head) and not as a complement. However, this is not the case in Arabic dialects. Edwards (2010) shows that Egyptian Arabic is a SVO dialect, and Owens, Dodsworth and Rockwood (2009) argue that the dialects of the Arabian Peninsula are SV/VS dialects. The dialects of Arabic differ from MSA in that they show tendencies towards certain structures found in MSA and that they sometimes allow structures that are considered ungrammatical in MSA. For example, Aoun et al. (1994) discusses the sentence ʤaw marwan wa karīm in which the sentence initial verb agrees with the following subject in number. This agreement is allowed in Moroccan Arabic but not MSA. In this paper, I present a treatment of simple sentences in MSA regarding their constituent structure, agreement, and their pragmatic usage.

I mainly present a Sign-Based Construction Grammar (SBCG) (Boas and Sag 2012) account of agreement in Arabic sentences. A grammar like SBCG is suitable for accounting for this phenomenon because SBCG does not presuppose the existence of an underlying structure. Therefore, there is no need to link different structures to each other. The different structures in which this phenomenon appears can be accounted for without assuming derivational, merge, or deletion rules.

SBCG allows each construction to be described individually, including its syntactic and pragmatic features, and in relation to each other, as well.

Moreover, Construction Grammar in general 'lacks cross-constructional generalizations' (Michaelis 2012). Therefore, each construction reflects a specific form-meaning pairing. Sign-Based Construction Grammar is also a licensing-based model as opposed to the Minimalist Program and Government and Binding which are suppression-based syntactic theories. As a licensing-based model, SBCG offers greater psychological plausibility and descriptive precision (Michaelis 2012). Michaelis (2012) also argues that "suppression-based approaches fail to account for speaker's structural preferences, as revealed by speech corpora". This is important for our purposes in this paper because agreement in Arabic exhibits a number of structures that are used by different speakers. For example, different dialects of Arabic have different preferences in terms of word order and sensitivity of the verb towards the agreement of the two conjuncts.

In SBCG, constraints are posited on linguistic expressions which require them to have certain properties. These constraints state that a linguistic 'expression is syntactically well-formed if its phonological form is paired with its semantics as an instance of some syntactic construction' (Zwicky 1994). This paper assumes that each Arabic structure that exhibits a different agreement between the verb and its subject is a separate construction with its own phonological and semantic properties. The paper is structured as follows: The following section explains the difference between the three sentence types in Arabic regarding their syntax and pragmatics which shows why the traditional view of sentence types is favoured over the Greenbergian view of word order. Section 3 discusses the theory of SBCG and explains its main operations and how syntactic derivations are computed. Section 4 presents the SBCG account of sentence types in Arabic. Section 5 concludes the paper and discusses areas for future research.

2. Sentence types in Arabic

This section discusses Arabic simple sentences based on sentence types rather than word order. It shows how Type 1 sentences, with a VSO word order, and Type 2 sentences, with an SVO word order, in Arabic are fundamentally different from each other and that none of them is the result of movement. Of course, whether a verb is present in the sentence or not is not the main issue in Arabic sentences. It is how the constituents behave in the sentence and what restrictions are posited on them. The following subsection (2.1) discusses the difference between the three sentence types in terms of their syntax. Subsection 2.2 discusses how these sentence types differ in their pragmatic properties.

2.1 Syntax

The simple sentences in Arabic are of three types: Type 1 (T1) sentences begin with a verb and has two variations according to its constituents, VSO or VO. Type 2 (T2) sentences, however, begin with a noun and have the main structure of subject-complement. The complement can be an NP, AjdP, PP or a VP. Type 3 (T3) sentences consist also of two parts. The first part is a preposition phrase and the

second one is an indefinite noun phrase. The following sections explain the three types of sentences and discuss their idiosyncrasies.

2.1.1 Type-1 sentences (T1)

Type 1 sentences, mainly, begin with a verb which is followed by an agent and a complement. This type of sentences feature a verb that always carries singular marking regardless of the number of its subject. Examples for this type include the following:

(3) a. qaraʔa ʔal-ʔawlād-u al-kıtāb-a
read the-boys-nom. the-book-acc.
The boys read the book.
b. qaraʔa ʔal-wlad-u al-kıtāb-a
read the-boy-nom. the-book-acc.
The boy read the book.
c. qaraʔat al-bınt-u al-kıtāb-a
read.F the-girl-nom. the-book-acc.
The girl read the book.

Sentence (3a) shows the verb 'read' which agrees with its subject in gender (they are both masculine). However, it does not agree with it in number (the verb carries singular marking while the subject is plural). Sentences (3b) and (3c) show the verb agreeing in gender with its subject but not in number regardless of the number of its subject. Another variation of Type 1 sentences in Arabic is when the subject of the verb is not overtly mentioned in the sentence. This includes sentences that begin with a verb and the subject is represented morphologically similar to the sentence in (4).

(4) (ʔanā) qaraʔ-tu al-kıtāb-a
(I) read-I the-book-acc.
I read the book.

In sentence (4), the subject which is a pronoun is put between parenthesis, because it is not needed unless further emphasis is required which is different from English in which the pronominal subject is obligatory. In both cases, the verb carries agreement to its agent. This sub-type depends on context for its subject.

2.1.2 Type-2 sentences (T2)

Sentences in this type consist of a sentential subject which is always a noun phrase and a predicative phrase that can be an NP, AjdP, PP or a VP. Sentence (5) is an example of this type in which the head noun phrase is followed by another noun phrase.

(5) Fatımat-u mudarrısat-un
Fatimat-nom. teacher.F-nom.
Fatimah is a teacher.

The sentence in (5) consists of two noun phrases: 'Fatimah' is the subject of the sentence. It is nominative because subjects of sentences are always nominative in Arabic - marked by a -u suffix. The second noun phrase is indefinite and it is also marked by a -u suffix because predicates of Arabic sentences are also nominative.

This sentence type allows for variation. The second constituent in this sentence type can be an adjective phrase instead of a noun phrase as in sentence (6a) or a preposition phrase as in (6b).

(6) a. Fatımat-u ʤamīlat-un
Fatimat-nom. beautiful.F-nom.
Fatimah is beautiful.
b. Fatımat-u fī al-bayt-ı
Fatimat-nom. in the-house-gen.
Fatimah is in the house.

The AdjP in (6a) is also nominative because it is the predicate of the sentence. Another extension of this sentence type is the verbal predicate. A sentence like the one in (7) repeated from (1) above consists of a sentential subject NP and a predicational verbal complement VP.

(7) ʔal-walad-u ʔakala ʔa-ttuffāħat-a
the-boy.S.M-nom. ate.S.M the-apple-acc.
The boy ate the apple.

The subject 'ʔalwaladu' is the subject of the sentence and not the verb. Both subjects of the verb and of the sentence are nominative in Arabic. This similarity in case marking has sparked a long debate as to whether the subject is assigned nominative case because it is the subject of the sentence or the subject of the verb. The idea that the S in a SVO sentence is treated as the subject of the sentence (topic) not the subject of the verb is rooted in the Arabic Linguistic Tradition (Peled 2008).

The view adopted in this paper is that the subject in an SVO sentence is assigned nominative case because it is the subject of the sentence not the verb[2]. It follows then that the clausal verb in VO acts like a predicate for the subject of the sentence and that the subject (of the sentence) is the head of the clause. Badawi, Carter and Gully (2004) explain that 'what appears to be inversion of the agent and verb is actually a variety of topic-comment sentence, in which the topic of the comment[3]-verb and the binding pronoun all happen to be identical (co-referential)' (p. 346). The sentences in (8) below show the verb carrying a marker that agrees with the subject of the sentence regardless whether the subject is a doer or a receiver of the action.

(8) a. ʔal-ʔawlād-u ʔakal-ū a-ttuffāħ-a
the-boy.P ate-they the-apple.P-acc.
The boys ate the apples.
b. ʔal-kıtāb-u qaraʔ-tu-hu bı-ʃaʏaf
the-book.S.M-nom. read-I-it in-love
I read the book keenly. lit. The book, I read it keenly.

In sentence (8a), the subject of the sentence is 'the boys' which is nominative, masculine, and plural. The verb carries morphological marking to indicate that its subject is masculine and plural. This marking is co-indexed with 'ʔalʔawlādu'. The subject of this sentence provides reference to the agent but does not act as one. On the other hand, sentence (8b) features a sentential subject 'the book' which is

nominative, singular, and masculine. The verb carries morphological marking that shows that its agent is singular and also carries a clitic that is co-indexed with the subject of the sentence. However, the subject in this sentence refers to the receiver of the action. The two sentences show that the verb is inflected to indicate the number, gender, and person of the agent. The verb is also co-indexed with the subject regardless of whether this subject provides reference to the verb's agent or not. The subject is always nominative even if it is the receiver of the action, which is usually accusative and marked with the '-a' vowel suffix in Arabic.

2.1.3 Type-3 sentences (T3)

Sentences in Type 3 are always verbless sentences. They consist of a prepositional phrase which is followed by a noun phrase. This type can be distinguished from T2 in that it allows an indefinite head and that the word order is less flexible. Consider the following examples:

(9) a. fī a-ddār-ı radʒul-u-n
in the-house-gen. man-nom.-indef.
A man is in the house.
b. *radʒul-u-n fī a-ddār-ı
man-nom.-indef. in the-house-gen.
A man is in the house.

In sentence (9a), the order of the two phrases is obligatorily reversed. Beginning with the noun phrase is ungrammatical as in sentence (9b)[4]. A sentence of this type begins with the complement PP followed by the subject NP. The reason behind this inversion is that the subject of the sentence, and therefore the head, cannot be indefinite in Arabic.

This section discussed the difference between T1 sentences (VSO) and T2 sentences that have a verb as their complements (SVO) regarding their syntax. The two sentence types differ in agreement. Type 1 agrees with the overt subject in gender and person. However, it does not agree with it in number, unless this subject is deleted. On the other hand, Type 2 sentences feature a sentential subject followed by either an NP, PP or a VP. The verb phrase is treated as a verbal complement to the subject of the sentence. It always carries inflection that agrees with its subject whether this subject is compatible with the subject of the sentence or not. The following section explains the difference between these sentence types regarding their pragmatics.

2.2 Pragmatics

The syntactic differences between SVO and VSO sentences in agreement and case assignment are not the only differences that can be found between them. The two types that these two word orders belong to differ also in their use. Arabic sentence types differ in what the sentence predicates about. The most prominent comes first in a sentence. Therefore, Type 1 sentences predicate about the subject and Type 2 ones predicate about the action. This aboutness is reflected by the topic of the sentence. Samek-Lodovici (2002) investigated a number of languages that exhibit

agreement impoverishment under I° and C-command (agreement asymmetry), and found that the VSO word order features pragmatically neutral subjects. This is in agreement with the treatment presented here as the predication in this type of sentences is about the action not the subject[5]. Because VSO word order is the discourse neutral of the two sentence types in Arabic, many linguists assumed it the basic word order from which the SVO word order is derived (Soltan 2007). This is because most studies investigated VSO and SVO word orders and do not take into account sentences that do not include a verb. In T2 sentences, for example, there is a number of variation. However, what brings these variations together is the fact that they begin with a definite subject that is the topic of the sentence followed by a predication be it a nominal one or a verbal one or otherwise.

The difference in use among the three sentence types depends on the topic and focus of the sentence. The two notions of topic and focus are very important for the structure of Arabic sentences (Peled 2008). Topic refers to what the sentence predicates about. It is old information that comes first in the sentence. Because it is shared information, it is definite. For this reason, subjects (words that come first in the sentence) are always definite in Arabic. Focus, on the other hand, refers to new information that is assumed by the speaker not to be shared by the hearer (Jackendoff 1972). In other words, Arnold, Wasow, Losongco and Ginstorm (2000) argue that items which have been recently mentioned tend to be accessible to both speaker and hearer, and require less complex descriptions than items which are not. Therefore, items that are new to the discourse tend to be complex, and items that are given tend to be simple (Arnold et al. 2000: 34).

The sentences in (10) and (11) repeated form (1) and (2) above are separate sentences from two different sentence types. Sentence (11), because it features a VSO word order, is pragmatically neutral. The change that occurs to this word order is pragmatically motivated (Bakir 1979; Soltan 2007).

(10) ʔal-walad-u ʔakala ʔa-ttuffāħat-a
the-boy.S.M-nom. ate.S.M the-apple-acc.
The boy ate the apple.

(11) ʔakala ʔal-ʔawalād-u ʔa-ttuffāħ-a
ate.S.M the-boys.P.M-nom. the-apples-acc.
The boys ate the apples.

Sentence (10) predicates about ‘the boy’ and is followed by a clausal predicate that is co-referential with the subject of the sentence. The treatment presented here is in line with the one presented by Badawi et al. (2004) in which they explain that T1 sentences in Arabic represent a basic structure. ‘They are not the result of any movement, fronting or extraction, nor are they an inversion of the kind “that film I have seen before”’ (Badawi et al. 2004: 326-327).

This inversion mentioned by Badawi et al. (2004) can occur in T2 Arabic sentences when the focus of the sentence is changed. However, the topic remains the same. It is the focus of the sentence that changes. For example:

(12) a. ʔal-ʔawlād-u muhandıs-ūn
the-boys-nom. engineer-P.

The boys are engineers.
b. muhandıs-ūn ʔal-ʔawlād-u
engineer-P.nom. the-boys-nom.
Engineers are the boys.

In sentence (12a), the word for 'the boys' is still the subject of the sentence. It has been inverted for focus purposes only. This is because subjects in Arabic are definite in T2 sentences. This is similar to inversion in English that is done for stylistic and pragmatic reasons. The possibility of inversion of the subject and predicate in Arabic sentences is what lead Brustard (2000) to consider Arabic as both topic and subject-prominent language. She explains that Arabic is a language with topic-introduced sentences with the possibility of inversion. This, in effect, means that the order of constituents in Arabic sentences is determined based on their pragmatic relativity and not on their syntactic features such as agreement or case.

2.3 Concluding remarks

The simple sentence in Arabic can be of three types. T1 sentences feature two variants, both beginning with a verb. In one variant, the subject is realized overtly, the other covertly through morphological marking. T2 sentences are composed of two constituents: A sentential subject and a predicative phrase. The predicative phrase can be an NP, AdjP, PP, or a VP. T3 sentences feature sentences that allow an indefinite subject. The predicative constituent here is a fronted PP. Early Arab grammarians used the theory of Amil to account for these three types. They were in agreement (more or less) regarding the treatment of these sentences except two (Peled 2008). Examples of these sentences are mentioned in (14).

(13) ʔaxū-ka qāʔım-un
brother-your standing-nom.
Your brother is standing.
(14) ʔa-rradʒul-u fī a-ddar-ı
the-man-nom. in the-house-gen.
The man is in the house.

Sentence (13) features a subject followed by an active participle. The source of their disagreement regarding similar sentences is that early Arab grammarians did not assign a separate category for adjectives. Instead, adjectives are included under the category of nouns. In modern Arabic Linguistics, however, adjectives are distinguished from nouns based on their function in the sentence. The active participle here functions as an adjective and can be treated as such. Therefore, this sentence is included within T2 sentences according to the treatment presented here. As for sentence (14), it features a definite subject followed by a prepositional phrase. The early Arab grammarians who did not assign a third type of sentences were faced with the problem that sentences of the same type can sometimes allow a definite subject and sometimes an indefinite one. However, separating the sentences with indefinite subjects in a separate sentence type overcomes this problem.

3. Sign-Based Construction Grammar

Sign-Based Construction Grammar is one a group of linguistic theories known as constructionist approaches to linguistics in which the construction is the main linguistic unit. A construction builds on Saussure's notion of 'sign' and expands it to include grammatical structures that may not include lexical items in them. A construction is a conventionalized form-meaning pairing which can be used to describe all levels of grammatical description (van Trijp 2013). SBCG is a formalized version of Construction Grammar (Michaelis 2012). It was developed to overcome shortcomings of previous Construction Grammar theories, in particular Berkeley Construction Grammar (BCG) (Fillmore and Kay 1995). At the same time, it was introduced "to expand the empirical coverage of HPSG [Head-driven Phrase Structure Grammar], while at the same time putting BCG on a firmer theoretical footing" (Sag 2012: 70). Both BCG and HPSG had strengths and weaknesses and SBCG was introduced to combine the two theories to make use of their strengths and rid of their weaknesses. SBCG represents constructions similar to BCG because it makes use of MOTHER-DAUGHTER levels of representation. SBCG is, on the other hand, similar to HPSG because it uses feature structures and assigns values to them.

The framework of SBCG consists of two parts. The first part is feature structure descriptions of signs and constructions representing linguistic items. The second part is a signature in which types are organized which shows how they inherit each other. An example of a type is *clause* from which different types of clauses are projected. A type that has no other types projecting from it is called a maximal type.

Signs, in SBCG, are linguistic items that are represented using Attribute Value Matrices (AVMs). An example of a sign is presented in Figure 1.

$$\begin{bmatrix} word & & \\ \text{PHON} & /\text{pæt}/ & \\ \text{FORM} & \langle \text{Pat} \rangle & \\ \text{ARG-ST} & \langle\rangle & \\ \text{SYN} & \begin{bmatrix} syn\text{-}obj & & \\ \text{CAT} & \begin{bmatrix} noun & \\ \text{CASE} & nom \\ \text{SELECT} & none \\ \text{XARG} & none \end{bmatrix} \\ \text{VAL} & \langle\rangle \\ \text{MRKG} & det \end{bmatrix} \\ \text{SEM} & \begin{bmatrix} sem\text{-}obj & \\ \text{IND} & i \\ \text{FRAMES} & \langle\rangle \end{bmatrix} \\ \text{CNTXT} & \begin{bmatrix} context\text{-}obj & \\ \text{BCKGRND} & \left\langle \begin{bmatrix} naming\text{-}fr & \\ \text{LABEL} & l_2 \\ \text{ENTITY} & i \\ \text{NAME} & \langle Pat \rangle \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix}$$

Figure 1 A model of the word *Pat* (Sag 2012: 99).

In Figure 1 the word *Pat* is modelled. The features in capitals on the left-hand side of the AVM are the six main parts. Each of these features is assigned values. To explain what this means, the different features in the sign *Pat* is discussed further. First, the top feature is the feature PHON which is given a *phonological-object (phon-obj)* value represented by phonemic transcription. The next feature is the feature FORM which is given a *morph-obj* value which can list all part of the sign. ARG-ST stands for ARGUMENT STRUCTURE and shows all the arguments this sign can satisfy. The angle brackets indicate that what is inside of them is a list of values. The list is empty in this case because the sign represented here is a lexical item out of context. The next feature, SYN, shows the syntactic representation of the sign and is modelled as a *syn-obj*. The sub-features of this feature structure are (CAT)EGORY, (VAL)ANCE, and MARKING (MRKG). The values of CAT include and are not limited to *noun, verb, prep, comp, adv,* and *adj*, each with its own category features. The next feature, SEM, assigns the semantic values of the

sign. Most research uses Frame Semantics (Fillmore, Lee-Goldman and Rhodes 2012). Frames in Frame Semantics are lists of the knowledge one needs to understand the meaning of a word. The last feature is (CNTXT) which specifies the pragmatic contextual features of the sign.

The word *Pat* as represented in the sign in Figure 1 is of type *proper-noun*. It can unify with a verb construction to form a sentence. For example, the word *left*, as sketched out in Figure 2, accepts a NP as its VAL (valance) value in its SYN feature structure which allows it to unify with the word pat as it is a type of NP. Hence, the sentence as a whole *Pat left* is a combination of the two signs in Figures 1 and 2.

$$
\begin{bmatrix}
\textit{word} & \\
\text{PHON} & /\text{lɛft}/ \\
\text{FORM} & \langle \textit{left} \rangle \\
\text{ARG-ST} & \langle \text{NP}[\textit{nom}]_i \rangle \\
\text{SYN} & \begin{bmatrix}
\textit{syn-obj} & \\
\text{CAT} & \begin{bmatrix}
\textit{verb} & \\
\text{VF} & \textit{fin} \\
\text{SELECT} & \textit{none} \\
\text{XARG} & \text{NP}[\textit{nom}]_i \\
\text{LID} & \langle \begin{bmatrix} \textit{leaving-fr} & \\ \text{LABEL} & l_3 \\ \text{SIT} & s \\ \text{S-SRCE} & i \end{bmatrix} \rangle
\end{bmatrix} \\
\text{MRKG} & \textit{unmk} \\
\text{VAL} & \langle \text{NP}[\textit{nom}]_i \rangle
\end{bmatrix} \\
\text{SEM} & \begin{bmatrix}
\textit{sem-obj} & \\
\text{IND} & s \\
\text{LTOP} & l_0 \\
\text{FRAMES} & \langle \begin{bmatrix} \textit{leaving-fr} & \\ \text{LABEL} & l_3 \\ \text{SIT} & s \\ \text{S-SRCE} & i \end{bmatrix}, \begin{bmatrix} \textit{past-fr} & \\ \text{LABEL} & l_2 \\ \text{ARG} & s \end{bmatrix} \rangle
\end{bmatrix}
\end{bmatrix}
$$

Figure 2 A model of the word left (Alhedayani 2016: 200).

This combination between the two signs is licenced by the Subject-Predicate construction in (15). "This construction says that two signs can combine as long as the second is a finite (and hence verbal) sign that selects the first via the VAL feature" (Sag 2012: 146). It allows the finite form *left* to unify with the nominal *Pat* to form a grammatical sentence. It, also, specifies that the head daughter is the

second daughter which works for English. Arabic sentences, however, differ in their assignment of the head daughter because in Arabic the notions of topic and focus are very important. This is further discussed in the following section.

(15) Subject-Predicate construction (subj-head-cxt) (Sag 2012: 146).

$$\textit{subj-pred-cl} \Rightarrow \left[\begin{array}{ll} \text{MTR} & \left[\text{SYN}\ Y\ !\ \left[\text{VAL}\ \langle\rangle\right]\right] \\ \text{DTRS} & \left\langle X, Z : \left[\text{SYN}\quad Y : \left[\begin{array}{ll} \text{CAT} & \left[\begin{array}{ll} \text{VF} & \textit{fin} \\ \text{INV} & - \\ \text{AUX} & - \end{array}\right] \\ \text{MRKG} & \textit{unmk} \\ \text{VAL} & \langle X \rangle \end{array}\right]\right]\right\rangle \\ \text{HD-DTR} & Z \end{array}\right]$$

The semantics feature of the Subject-Predicate construction is not specified because the frames of the two signs unified by this construction combine through the principle of compositionality (Sag 2012).

The following section discusses the novel account that I present for Arabic sentences based on the account presented for English by Sag (2012) and the one presented for Arabic in Alhedayani (2016). However, this account differs from both previous ones in that it adopts a sentence type view not word order view.

4. The Sign-Based Construction Grammar account

This section presents the SBCG treatment of the three sentence types in Arabic. A neighbourhood of constructions are presented in order to account for different idiosyncrasies of Arabic sentences. This treatment is a development on the work presented in Alhedayani (2016), albeit based on a different theory of sentence types rather than word order. The first subsection here presents an account of T1 sentences in which two constructions are adapted from the ones presented by Sag (2012) in order to account for the two variations of this type. The second subsection discusses the T2 sentences. In this section, the Subject-Predicate-Construction (Sag 2012) is modified to account for these sentences. The third subsection introduces a treatment for the T3 sentences.

4.1 Type-1 sentences

This section presents the construction that licenses the T1 sentences in Arabic. There are two sub-types in this section. The first sub-type includes sentences that begin with a verb and have an overt subject. These are the typical T1 sentences similar to the sentence in (16) repeated from (2) above.

(16) ʔakala ʔal-ʔawalād-u ʔa-ttuffāħ-a
ate.S.M the-boys.P.M-nom. the-apples-acc.
The boys ate the apples.

The second sub-type includes sentences that begin with a verb and the subject is

represented morphologically similar to the sentence in (17).

(17) ʔakal-ū ʔa-ttuffaħ-a
ate-P.M the-apples-acc.
They ate the apples.

It is a requirement that the subject of the verb remains represented morphologically and not lexically which prevents the licensing of sentences such as the one in (18) because it is ungrammatical.

(18) *ʔakal-ū ʔal-ʔawalād-u ʔa-ttuffaħ-a
ate-P.M the-boys.P.M-nom. the-apples-acc.
The boys ate the apples.

In sentence (18), the verb carries morphological marking indicating the number and gender of the subject. This is unacceptable in MSA but is allowed in some dialects of Arabic. The saturational-head-complement-construction in (19) accounts for sentences in this category.

The construction in (19) is an adaptation of the construction with the same name that is presented by Sag (2012) for headed constructions like prepositional phrases and verb initial sentences in languages that allow them.

(19) Saturational Head-Complement Construction (↑headed-cxt) (adapted from (Sag 2012))

$$sat\text{-}hd\text{-}comp\text{-}cxt \Rightarrow \begin{bmatrix} \text{MTR} & \left[\text{SYN X } !\left[\text{VAL } \langle\rangle\right]\right] \\ \text{DTRS} & \left\langle \text{Z, Y: } \left[\text{AGR } \left[\text{GENDER } i\right]\right] \right\rangle \oplus \text{L} \\ \text{HD-DTR} & \text{Z: } \begin{bmatrix} word & \\ \text{SYN X: } & \begin{bmatrix} \text{CAT } \begin{bmatrix} verb \\ \text{XARG } none \end{bmatrix} \\ \text{AGR } \begin{bmatrix} \text{NUMBER } sing \\ \text{GENDER } i \end{bmatrix} \\ \text{VAL Y} \oplus \text{L} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

The mother of the construction takes its syntax from the head daughter but differs from it in valence. The daughters are the head verb, its subject, and a possible list. The head daughter is of category verb and has no external argument. The agreement feature of the head daughter determines that its number is always singular but takes its gender from the second daughter (the subject of the verb). The valence feature shows that the verb requires the second daughter to be realized lexically, in addition to a possible list. Sag (2012) explains that '[a]ccording to this "saturating" mode of realization; all the head's valents are realized as sisters of the lexical head' (Sag 2012: 152).

The second variation of T1 sentences includes sentences that begin with a verb but with no overt subject. Sentences like ʔakalu 'They ate' are licensed by a construction that I call saturational-no-subject-construction, in (20).

(20) Saturational no-Subject Construction (↑headed-cxt):

$$sat\text{-}no\text{-}sub\text{-}cxt \Rightarrow \begin{bmatrix} \text{MTR} & \left[\text{SYN } X \text{ !} \left[\text{VAL } \langle\rangle\right]\right] \\ \text{DTRS} & \langle Z \rangle \oplus L\text{: } nelist \\ \text{HD-DTR} & Z\text{:} \begin{bmatrix} word \\ \text{SYN } X\text{:} \begin{bmatrix} \text{CAT} & \begin{bmatrix} verb \\ \text{XARG } none \end{bmatrix} \\ \text{AGR} & \begin{bmatrix} \text{GENDER } i \\ \text{NUMBER } j \end{bmatrix} \\ \text{VAL} & L \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

This construction differs from the previous one in that the valence does not require an overt subject and that the number is not specified to be singular. It is important to note that the daughters which include a non-empty-list, can be a morphological construction either as a suffix attached to the verb as in ʔakalū 'They ate' or built in the verb's morphology as in ʔakala 'He ate'. I, however, do not go into specifications when it comes to morphological constructions.

4.2 Type-2 sentences

Type two sentences consist of two constituents, a subject and a predicate. They are licensed by the subject-predicate construction represented in (21) below, adapted from Sag (2012). However, the subject-predicate construction presented here is different from the one presented by Sag (2012) for English in that the predicate is not a finite form.

(21) Subject-Predicate Construction (↑subj-head-cxt):

$$\begin{bmatrix} \text{MTR} & \left[\text{SYN } X \text{ i}\left[\text{AGR } \boxed{1}\right]\right] \\ \text{DTRS} & \left\langle Y\text{:} \begin{bmatrix} \text{SYN } X & \begin{bmatrix} \text{CAT} & \begin{bmatrix} noun \\ \text{CASE } nom \\ \text{STATE } definite \end{bmatrix} \\ \text{AGR } \boxed{1} & \begin{bmatrix} \text{NUMBER } i \\ \text{GENDER } j \end{bmatrix} \end{bmatrix} \\ \text{label} & l_1 \end{bmatrix}, Z\text{:} \left[\text{SYN } \begin{bmatrix} \text{AGR} & \boxed{1} \\ \text{XARG} & \langle Y \rangle \end{bmatrix}\right] \right\rangle \\ \text{CNTXT} & \left[\text{topic } l_1\right] \\ \text{HD-DTR} & Y \end{bmatrix}$$

This construction licenses the head and its complement. The mother's syntax corresponds to the syntax of the head daughter. It specifies that the agreement features of the mother are the same as the head daughter (X). The daughters in this construction are two because the sentence type consists of two parts. The first daughter (Y) is the subject, which is a nominal with nominative case and definite marking. The second daughter (Z) is the predicate. There is nothing that specifies the category of this daughter because it accepts different types of phrases. However, its syntax feature in the construction specifies that it carries the same agreement features of the head daughter. The context feature shows that the head daughter is

the topic.

The head can be followed by three types of phrasal constructions; NP, AdjP, and PP. For example, when the head is followed by a noun, sentences with the phrasal structure NP NP are licensed; similar to the one in (22). The nominal complement has its own restrictions: 1. the complement is indefinite, 2. they agree in number and gender, 3. they are both nominative not because they agree with each other but because both subjects and predicates are nominative is Arabic.

(22) ʔa-rrıʤāl-u muhandıs-ūn
the-men-nom engineer-P.nom.
The men are engineers.

Type 2 does not allow indefinite heads. However, these sentences are allowed to be inverted. Sentences like the one in (22) are ordered as such because the word muhandıs-ūn 'engineers' is the focus not the head.

(23) muhandıs-ūn ʔa-rrıʤāl-u
engineer-P.nom. the-men-nom
The men are engineers.

Brustard (2000) argues that Arabic is a language that is both topic and subject prominent. Topics (what the sentence predicates about) are heads. Subjects (words that come first in the sentence) are the focus of the sentence. Sometimes the topic and focus are the same word as in (22) above and sometimes they are not as in (23) above. This distinction between topic/focus subject/head can be made clearer by looking at adjectival complements as in (24).

(24) ʔal-ʔawalād-u muʔadab-ūn
the-boys.P.M-nom. polite-P.nom.
The boys are polite.

Similar to the NP NP sentences, the NP AdjP ones posit restrictions on the order not based by headedness, but by focus. The adjective also agrees with the head noun in number and gender bur differs in state. Here, as in the NP NP variant of the construction, the head must be the definite constituent (NP) since the adjective agrees with the noun. The noun cannot take its agreement features from the adjective as they are inherit in the noun. Therefore, the order of constituents does not reflect headedness but reflects focus. The same construction also accounts for NP PP sentences.

A sister construction accounts for the NP VP variation of T2 sentences. To produce such sentences, the saturational-no-subject construction combines with the subject-head construction which produces the one in (25).

(25) Subject-Predicate Construction with the saturational-no-subject construction as the predicate:

$$\left[\begin{array}{ll}
\text{MTR} & \left[\text{SYN } X \left[\text{AGR } \boxed{1}\right]\right] \\
\text{DTRS} & \left\langle Y\text{:} \left[\begin{array}{ll} \text{SYN } X & \left[\begin{array}{l} \text{CAT} \left[\begin{array}{l} noun \\ \text{CASE } nom \\ \text{index } k \end{array}\right] \\ \text{AGR } \boxed{1} \left[\begin{array}{l} \text{NUMBER } i \\ \text{GENDER } j \end{array}\right] \end{array}\right] \\ \text{label} & l_1 \end{array}\right], Z\text{:} \left[\begin{array}{ll} \text{SYN} & \left[\begin{array}{l} \text{CAT} \left[\begin{array}{l} verb \\ \text{XARG } none \end{array}\right] \\ \text{AGR} \quad \boxed{1} \\ \text{VAL} \quad \left[\begin{array}{l} morph \\ \text{index } k \end{array}\right] \oplus \text{L} \end{array}\right] \end{array}\right] \right\rangle \\
\text{CNTXT} & \left[\text{topic } l_1\right] \\
\text{HD-DTR} & Y
\end{array}\right]$$

The construction in (25) licenses sentences like the one in (26). This type of sentences received a lot of attention by linguists investigating the different word orders in Arabic.

(26) ʔal-ʔawlād-u ʔakal-ū
the-boys-nom. ate-they
They ate.

The construction specifies that when the saturational no subject construction combines with the subject predicate construction; the morphological marking of the verb must be co-indexed with the subject of the sentence. Also, the definite feature in the head daughter is removed. This is because the definiteness condition does not apply when the predicate is a verb (Peled 2008).

4.3 Type-3 sentences

Type 3 sentences are similar to type 2 sentences in that they both consist of two parts, but they differ in that type 3 sentences allow indefinite subjects and that the PP predicate is fronted. Sentence (27) is an example of this type.

(27) fī ʔa-ddar-ı radʒul-un
in the-house-gen. man-nom.
A man is in the house.

As sentence (27) shows, the third type of Arabic sentences features no verb in the sentence. They have two constituents: A prepositional phrase, fī ʔa-ddar-ı ‘in the house’, that is followed by an indefinite noun phrase, radʒul-un ‘a man’. The construction in (28) licenses such sentences.

(28) Subject-Predicate Construction (with minor modification) (↑subj-head-cxt):

$$\begin{bmatrix} \text{MTR} & [\text{SYN } X] \\ \text{DTRS} & \left\langle Y\!: \begin{bmatrix} \text{SYN } X & \begin{bmatrix} \text{CAT} & \begin{bmatrix} \textit{noun} \\ \text{CASE } \textit{nom} \\ \text{STATE } \textit{indefinite} \end{bmatrix} \\ \text{AGR } \boxed{1} & \begin{bmatrix} \text{NUMBER } i \\ \text{GENDER } j \end{bmatrix} \end{bmatrix} \\ \text{label} & l_1 \end{bmatrix}, Z\!: \begin{bmatrix} \text{SYN} & [\text{CAT PP}] \\ \text{label} & l_2 \end{bmatrix} \right\rangle \\ \text{CNTXT} & \begin{bmatrix} \text{topic } l_1 \\ \text{focus } l_2 \end{bmatrix} \\ \text{HD-DTR} & Y \end{bmatrix}$$

The construction subject-predicate is modified to account for this type. Here the mother does not specify that the two daughters agree in gender and number as does the sub-pred-cx. Also, the subject is an indefinite noun phrase. The predicate is a PP which is also the focus. This is shown in the context feature in the construction. Word order is determined by focus in Arabic sentences; therefore, the noun head daughter appears last in the sentence.

Determining the head in this construction is not straight forward because agreement is not shown and because there is no XARG. It is not clear whether the sentence predicates about the noun 'a man' or the prepositional phrase 'in the house'. Arab grammarians claim that predication cannot be about an indefinite noun. Hence, the PP must come first in the sentence. However, they still consider the noun to be the subject of the sentence because subjects must be nouns. Accordingly, the noun phrase is the head albeit obligatorily postponed in the sentence.

4.4 Concluding remarks

This section presented a SBCG account of the three types of simple sentences found in MSA. A neighbourhood of constructions are presented to account for the different idiosyncrasies of each type. For Type 1, two sister constructions account for the sentences that begin with a verb which differ based on whether the subject is lexically or morphologically realized. For Type 2, the subject predicate construction allows for variations within this type based on the phrasal structure of the sentence. For Type 3, the subject predicate construction is modified to allow for subtle characteristics that distinguish this type from the previous one. The pragmatic notions of focus and topic determine the head of the construction and the order of constituents.

5. Conclusion

This paper shows that adopting a sentence type view instead of a word order view allows for a better understanding of Arabic sentences. Word order represents one factor affecting agreement but is not the only one. The account presented in this paper derives from the one presented by Sag (2012) for English. However, even though constructions for Arabic and English are different, this does not mean that

SBCG cannot account for universal phenomena. On the contrary, its feature structures and type hierarchies are the same. The constructs are different because the two languages employ different sentence types with different functions.

Moreover, the account presented in this paper is of high typological value. It is the first account given for a language that exhibits VSO word order. This word order is rare among languages (Hahn and Xu 2022). However, Arabic is not the only language with VSO word order. Languages like Irish, Welsh, Māori, and some Native American languages exhibit this word order, too. The sentence type account discussed here following traditional Arabic linguistics can be of use for accounting for languages that exhibit different word order combinations with varying semantic and pragmatic implications.

Rukayah Alhedayani
Assistant Professor of Linguistics
King Saud University, Riyadh, Saudi Arabia
ORCID Number: 0000-0002-5168-9528
Email: ralhedayani@ksu.edu.sa

Endnotes

[1] A third sentence type is aljumlah aDarfiyyah consisting of a noun phrase and a prepositional phrase which some linguists include under type 2. For a, thorough discussion of sentence types in both traditional and modern Arabic Linguistics, see Peled (2008).
[2] The English term ‘subject’ can be ambiguous here. I use ‘subject of the sentence’ to refer to the head, in Arabic mubtadaʔ; and ‘subject of the verb to refer to the doer, in Arabic fāʔıl.
[3] Badawi et al. (2004) use the term ‘comment’ for what I call ‘predicate’ in this paper.
[4] This is ungrammatical unless it is an answer to a question like ‘who is at the house?’ However, answers to questions can be fragments.
[5] A discussion of the pragmatic neutrality (having the least contextual restrictions) of VSO sentences in Arabic can be found in Albuhayri (2019).

References

Albuhayri, Salem. (2019). Information structure in standard Arabic verbal sentences. PhD Thesis, University of Wisconsin Milwaukee.

Alhedayani, Rukayah. (2016). Antonymy in modern standard Arabic. PhD Thesis, University of Sussex.

Aoun, Joseph, Elabbas Benmamoun and Lina Choueiri. (2009). *The Syntax of Arabic*. Cambridge: Cambridge University Press.

Aoun, Joseph, Elabbas Benmamoun and Dominique Sportiche. (1994). 'Agreement and conjunction in some varieties of Arabic'. *Linguistic Inquiry*, 25: 195–220.

Arnold, Jennifer, Thomas Wasow, Anthony Losongco and Ryan Ginstorm. (2000). 'Heaviness vs. newness: The effects of structural complexity and discourse status on constituent ordering'. *Language*, 76(1): 28–55.

Arts, Tessy, Yonatan Belinkov, Nizar Habash, Adam Kilgarriff and Vit Suchomel. (2014). 'ArTenTen: Arabic corpus and word sketches'. *Journal of King Saud University-Computer and Information Sciences*, 26(4): 357-371. https://doi.org/10.1016/j.jksuci.2014.06.009

Badawi, El Said, Michael Carter and Adrian Gully. (2004). *Modern Written Arabic: A Comprehensive Grammar*. New York: Routledge.

Bakir, Murtadha J. (1979). *Aspects of Clause Structure in Arabic: A Study of Word Order Variation in Literary Arabic.* Bloomington: Indiana University.

Boas, Hans C. and Ivan A Sag. (2012). *Sign-Based Construction Grammar.* Stanford: CSLI Publications.

Brustard, Karen. (2000). *The Syntax of Spoken Arabic*. Washington, D.C.: Georgetown University Press.

Chomsky, Noam. (1957). *Syntactic Structures*. Berlin: Mouton and Co.

Chomsky, Noam. (2005). *On Phases*. MA: MIT.

Edwards, Malcolm H. (2010). 'Word order in Egyptian Arabic: Form and function'. In Jonathan Owens and Alaa Elgibali, (eds), *Information Structure in Spoken Arabic*, 93-106. London: Routledge.

Fillmore, Charles J. and Paul Kay. (1995). Construction grammar coursebook. Unpublished manuscript.

Fillmore, Charles J., Russell R. Lee-Goldman, and Russell Rhodes. (2012). 'The FrameNet construction'. In Hans C. Boas and Ivan A Sag (eds). *Sign-Based Construction Grammar*, 309-372. Stanford: CSLI Publications.

Greenberg, Joseph H. (1966). *Universals of Language*. 2nd edition. Cambridge: MIT Press.

Hahn, Michael and Yang Xu. (2022). 'Crosslinguistic word order variation reflects evolutionary pressures of dependency and information locality'. *PNAS: Psychological and Cognitive Sciences*, 119(24). https://doi.org/10.1073/pnas.2122604119

Jackendoff, Ray. (1972). *Semantic Interpretation in Generative Grammar*. Cambridge: MIT Press.

Jarrah, Marwan, Mohammad Rayyan, Ekab Al-Shawashreh and Malek J Zuraikat. (2020). 'Subject-verb agreement in Arabic, Spanish, and English: Typological view'. *Dirasat: Human and Social Sciences*, 47(1): 144–155. file:///Users/ruqwee/Downloads/103999-126883-1-PB.pdf

Kinjo, Kunio. (2015). 'A probe-goal approach to impoverished agreement'. *Proceedings of Seoul International Circle of Generative Grammar* 17. Seoul, 225–244.

Larson, Bradley. (2013). 'Arabic conjunct-sensitive agreement and primitive operations'. *Linguistic Inquiry*, 44(4): 611–631. DOI: 10.1162/LING_a_00140

Michaelis, Laura. A. (2012). 'Making the case for construction grammar'. In Hans C. Boas and Ivan A. Sag (eds). *Sign-Based Construction Grammar*, 31–67. Stanford: CSLI publications.

Owens, Jonathan, Robin Dodsworth and Trent Rockwood. (2009). 'Subject-verb order in spoken Arabic: Morpholexical and event-based factors'. *Language Variation and Change*, 21: 39–67. DOI: https://doi.org/10.1017/S0954394509000027

Peled, Yishai. (2008). *Sentence Types and Word Order Patterns in Written Arabic: Medieval and Modern Perspectives*. Amsterdam: Brill.

Sag, Ivan. (2012). 'Sign-Based Construction Grammar: An informal synopsis'. In Hans C. Boas and Ivan A. Sag (eds). *Sign-Based Construction Grammar*, 69–202. Stanford: CSLI Publications.

Samek-Lodovici, Vieri. (2002). 'Agreement impoverishment under subject inversion'. 11:49-82. https://roa.rutgers.edu/files/431-1200/431-1200-SAMEK-LODOVICI-0-0.PDF

Soltan, Usama. (2007). 'On formal feature licensing in minimalism: Aspects of standard Arabic morphosyntax'. PhD Thesis, University of Maryland, College Park. http://hdl.handle.net/1903/7581

Toutanova, Kristina, Dan Klein, Christopher Manning and Yoram Singer. (2003). 'Feature-Rich part-of-speech tagging with a cyclic dependency network'. In *Proceedings of HLT-NAACL 2003*, 252-259. https://nlp.stanford.edu/pubs/tagging.pdf

van Trijp, Remi. (2013). 'A comparison between Fluid Construction Grammar and Sign-Based Construction Grammar'. *Constructions and Frames*, 5(1): 88-116.

Zwicky, Arnold M. (1994). 'Dealing out meaning: Fundamentals of grammatical constructions'. In *Proceedings of the Twentieth Annual Meeting of the Berkeley Linguistics Society*, 20: 611–625.