

## **Clause Structure in Palestinian Arabic: The VP-internal Subject Hypothesis and the Derivation of Verb-initial Word Order**

Most studies of Palestinian Arabic have assumed the VP-internal subject hypothesis in their clause structure analysis. In this paper, I present evidence for this widely accepted hypothesis from Palestinian Arabic, setting the grounds for future research of this language.

### **1 Introduction**

The VP-internal subject hypothesis is by now assumed in many syntactic analyses. This hypothesis, first suggested by Kuroda (1988), holds that the subject is base-generated within the VP, where it receives its theta role, and moves into the IP domain in order to check features with the Infl(ectio)nal projection and be assigned Case. In this paper I provide additional support for the VP-internal subject hypothesis, by presenting data from Palestinian Arabic (PA, henceforth), in which attested verb-initial as well as SVO word order must be captured. In section 2, I show that there are two word order possibilities in PA, namely SVO and VSO. In section 3, I consider two hypotheses regarding the underlying structure that generates these two possibilities. In section 4, I give support for the hypothesis holding that the subject is base generated within the VP, by considering matrix questions, and embedded and relative clauses. In section 5 I briefly discuss whether the VP-internal subject is consistent with the Theta Criterion and Case Filter. Section 6 presents a conclusion, and lays out additional issues in the syntactic analysis of this language that should be addressed and pursued next.

### **2 Background on the Word Order of Palestinian Arabic**

PA allows two word order possibilities in sentences with transitive verbs: SVO and VSO, as illustrated in the sentences in (1). Further examination of the two configurations

reveals that the SVO variation is the neutral one, whereas the VSO variation is interpreted either as a question or a marked response to a question. Other word order possibilities are not permitted, even if the subject could in principle be identified from the subject-verb agreement in PA.

- |       |                        |                        |                        |                         |
|-------|------------------------|------------------------|------------------------|-------------------------|
| 1. a. | shaf                   | aḥmad                  | maryam                 | <b>VSO</b>              |
|       | saw.3 <sup>rd</sup> SM | Aḥmad                  | Mary                   |                         |
|       | 'Ahmad saw Mary.'      |                        |                        |                         |
| b.    | *shaf                  | maryam                 | aḥmad                  | <b>*VOS<sup>1</sup></b> |
|       | saw.3 <sup>rd</sup> SM | Mary                   | Aḥmad                  |                         |
|       | 'Ahmad saw Mary.'      |                        |                        |                         |
| c.    | aḥmad                  | shaf                   | maryam                 | <b>SVO</b>              |
|       | Ahmad                  | saw.1 <sup>st</sup> SM | Mary                   |                         |
|       | 'Ahmad saw Mary.'      |                        |                        |                         |
| d.    | *aḥmad                 | maryam                 | shaf                   | <b>*SOV</b>             |
|       | Ahmad                  | Mary                   | saw.3 <sup>rd</sup> SM |                         |
|       | 'Ahmad saw Mary.'      |                        |                        |                         |
| e.    | *maryam                | shaf                   | aḥmad                  | <b>*OVS</b>             |
|       | Mary                   | saw.3 <sup>rd</sup> SM | Ahmad                  |                         |
|       | 'Ahmad saw Mary.'      |                        |                        |                         |
| f.    | *maryam                | aḥmad                  | shaf                   | <b>*OSV</b>             |
|       | Mary                   | Ahmad                  | saw.3 <sup>rd</sup> SM |                         |
|       | 'Ahmad saw Mary.'      |                        |                        |                         |

In sentences with ditransitive verbs, the two word order configurations are Subject-Verb-Indirect Object-Direct Object (SVI<sub>o</sub>Do) or Verb-Subject-Indirect Object -Direct Object (VSI<sub>o</sub>Do), the latter being the marked variation. This is illustrated in the sentences in (2).

---

<sup>1</sup> According to Mohammed (2000) and from previous grammaticality judgments given by my informant, this word order is argued to be grammatical in PA. However, contradictory judgments as well as the sentences given in (2) lead me to doubt that such configuration is as commonly acceptable as SVO and VSO. I attribute this previous claim to the relatively free word order in Standard Arabic (SA) that is used in the media, thereby affecting the syntax of other dialects in Arabic. I will leave further discussion of this configuration to subsequent papers.

2. a. muna a'tat ali el-kitab **SVIoDo**  
 Mona gave.3<sup>rd</sup>SF Ali the-book  
 'Mona gave Ali the book.'
- b. a'tat muna ali el-kitab **VSIoDo**  
 gave.3<sup>rd</sup>SF Mona Ali the-book  
 'Mona gave Ali the book.'

The order of the two objects can be reversed by cliticizing a preposition probably subcategorized for by the verb onto the indirect object. In this case, the preposition *li* ('to') is cliticized onto the indirect object, the NP *ali*, as shown in the sentences in (3).

3. a. muna a'tat el-kitab li-ali **SVDOIo**  
 Mona gave.3<sup>rd</sup>SF the-book to-Ali  
 'Mona gave the book to Ali.'
- b. a'tat muna el-kitab li-ali **VSDoIo**  
 gave.3<sup>rd</sup>SF Mona the-book to-Ali  
 'Mona gave the book to Ali.'

Some ungrammatical word order variations are given in (4). Note that all the possible configurations were tested, with or without the preposition *li* appearing:

4. a. \*a'tat ali muna el-kitab **\*VIoSDo**  
 gave.3<sup>rd</sup>SF Ali Mona the-book  
 'Mona gave Ali the book.'
- b. \*a'tat ali el-kitab muna **\*VIoDoS**  
 gave.3<sup>rd</sup>SF Ali the-book Mona  
 'Mona gave Ali the book.'

Another structure that will be considered, but not discussed thoroughly in this paper, is the cliticization of a pronominal direct object onto the verb, as in the sentences in (5-6). The surface structures given in (5-6a) seem to suggest that there is another word order variation: VOS.

5. a. shafat-hu Maryam **VOS?**  
 saw.3<sup>rd</sup>SF-him Mary  
 'Mary saw him.'

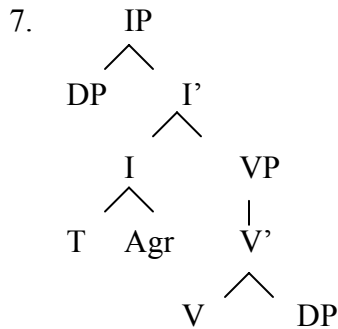
- |       |   |                |
|-------|---|----------------|
| b.    | Maryam shafat-hu<br>Mary saw.3 <sup>rd</sup> SF-him<br>'Mary Saw him.'  | <b>SVO</b>     |
| 6. a. | a'tat-hu                    muna    li-ali<br>gave.3 <sup>rd</sup> SF-it.SM Mona    to-Ali<br>'Mona gave it to Ali.'  | <b>VDoSIo</b>  |
| b.    | muna    a'tat-hu                    li-ali<br>Mona    gave.3 <sup>rd</sup> SF-it.SM to-Ali<br>'Mona gave it to Ali.'  | <b>SVDoIo</b>  |
| c.    | *a'tat-hu                    li-ali    Muna<br>gave.3 <sup>rd</sup> SF-it.SM to-Ali    Mona<br>'Mona gave it to Ali.' | <b>*VDoIoS</b> |
| d.    | *muna    li-ali    a'tat-hu<br>Mona    to-Ali    gave.3 <sup>rd</sup> SF-it.SM<br>'Mona gave it to Ali.'              | <b>*SIoVDo</b> |

Since VOS is ungrammatical in the same structures where the direct object is realized with proper nouns, I will assume that the cliticization is a result of movement of the direct object pronoun to a projection shared with the verb, either by S-structure or PF (Phonological Form). A structure showing this will be presented when I discuss PA matrix questions, in section 4.1.

### 3 VP-internal Subject vs. Subject in Spec,IP

The two D-structure hypotheses that will be considered here contrast in their definition of a thematic domain—the maximal projection in which theta roles are assigned by the predicate to its argument—and whether this domain is distinguished from the inflectional domain, where the arguments check their features and are assigned case.

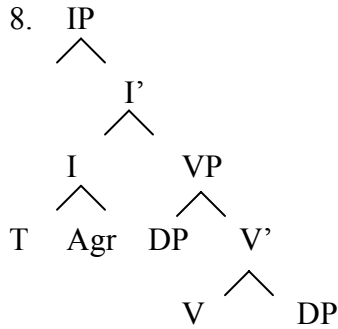
The first proposed underlying structure, depicted in (7), is a structure where the subject is base generated in Spec,IP, where it is assigned Nominative Case by Spec-head agreement, and a theta-role from the predicate heading the XP complement of Infl.



Only the internal argument (usually objects and subjects of passive and unaccusative verbs) is assigned a theta-role within the VP. Thus, the IP is both the thematic and functional layer. This structure also assumes V-to-I raising in order for the verb to check feature such as tense and agreement at Infl<sup>2</sup>. This is the case in many other languages, Semitic languages in particular. This hypothesis (**Hypothesis A** throughout) predicts that SVO is the underlying configuration, which also manifests itself by the unmarkedness of the sentences in (1c) and (2a). To achieve VSO word order, the verb undergoes I-to-C movement.

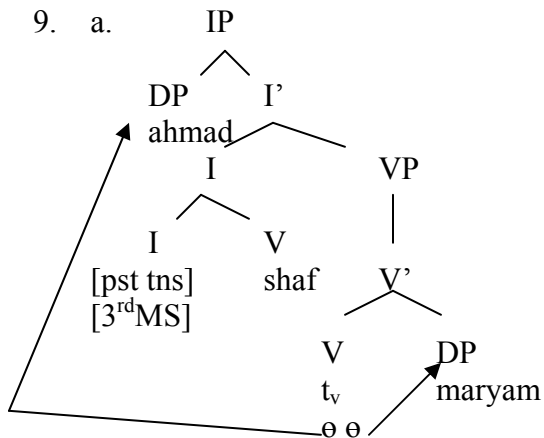
The second proposed underlying structure is a structure where the subject is base generated in Spec,VP, where it is assigned an external theta-role. This proposal, referred to as **Hypothesis B** throughout, considers the VP as the thematic layer, where theta-roles are assigned.

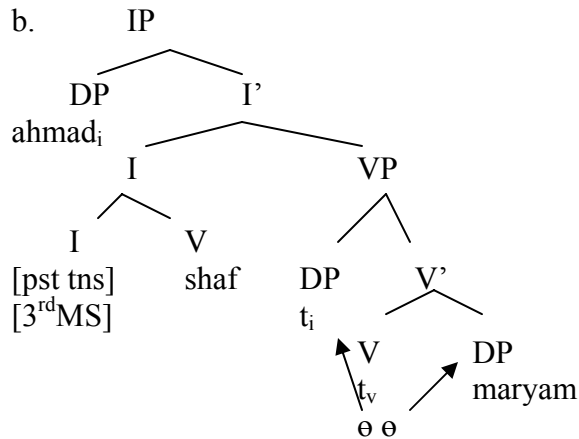
<sup>2</sup> I assume a rather simple Infl structure for the purpose of this paper. However, the split-infl hypothesis (Pollock, 1988) should be evaluated in the future in light of data from PA.



The IP, in this perspective, is a functional layer, where features such as tense and agreement are being checked. Thus, verb raising is also mandatory here for checking features such as tense and subject-verb agreement. In order to achieve the neutral SVO word order, the subject has to undergo A-movement to Spec,IP, in order to be assigned Nominative Case. The subject's staying in situ, at Spec,VP results in a marked VSO word order. The D-structure according to Hypothesis B is depicted in (8).

Structures (9a) and (9b) show the S-structure of the SVO configuration of the sentence in (1c) according to Hypothesis A and Hypothesis B respectively:





#### 4 Support from Matrix Questions and Relative and Embedded Clauses

Now the two hypotheses will be considered in light of their ability to account for the empirical facts of structures of PA matrix questions and relative and embedded clauses

##### 4.1 Matrix Questions in Palestinian Arabic

First, we consider *wh*-questions where both the Spec,CP and C positions are filled with pronounced elements. In such questions, the *wh*-word *aish* ('what') and the complementizer *ili* ('that') occupy Spec,CP and C respectively. As shown in the sentences in (10-11 a), the surface structure seems to be VOS, reminiscent of the sentences in (5a-6a).

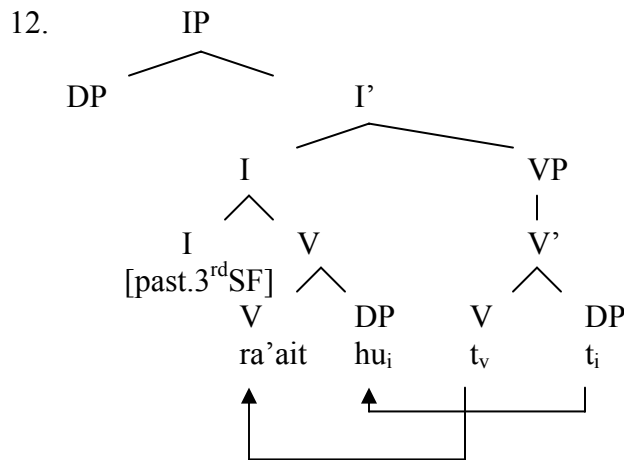
10. a. aish ili hatait-hu maryam ala el-tawila  
 what that put.3<sup>rd</sup>FS-it.SM Mary on the-table  
 'What did Maryam put on the table?'

b. aish ili maryam hatait-hu ala el-tawila  
 what that Mary put.3<sup>rd</sup>FS-it.SM on the-table  
 'What did Mary put on the table?'

11. a. aish ili ra'at-hu maryam  
 what that saw.3<sup>rd</sup>FS-it.SM Mary  
 'What did Mary see?'

- b. aish ili maryam ra'at-hu  
 what that Mary saw.3<sup>rd</sup>FS-it.SM  
 'What did Mary see?'

Whether we consider Hypothesis A or B, we can assume that the cliticization of the direct object pronoun to the verb occurs within the Infl domain. See internal structure of the IP layer in (12), where the identity of the element occupying Spec,IP is yet to be determined. For further discussion of Semitic Clitics, see Shlonsky (1997:177-188).



From the prediction given in Hypothesis A, the verb has to undergo two stages of movement—from V-to-I and subsequently from I-to-C—in order to achieve a verb-initial word order. In the sentences in (10-11), the C position is already filled with complementizer *ili* ('that'), thereby prohibiting I-to-C movement of the verb, and predicting a verb-initial word order to be ungrammatical, as shown in (10a), given here again:

- 10. a. [<sub>C</sub>aish<sub>i</sub> ili [<sub>IP</sub>maryam hatait<sub>v</sub>-hu<sub>i</sub> [<sub>VP</sub>t<sub>v</sub> t<sub>i</sub> ala el-tawila]]  
 what that put.3<sup>rd</sup>FS-it.SM Mary on the-table  
 'What<sub>i</sub> did Maryam put t<sub>i</sub> on the table?'  
 ↑ I-to-C blocked | ↑ V-to-I

The underlying structure proposed in Hypothesis B does not rule out such constructions as in (10-11), as the different possible configurations in PA can be achieved



by head- and phrasal- movements within the IP, leaving the C head available to be filled by the overt *ili* complementizer.

#### 4.2 Relative Clauses in Palestinian Arabic

Relative clauses in PA are externally headed. The relative clause (RC) may be marked by the complementizer *ili*, or not marked at all. Sentence (13a) shows an RC marked by this complementizer. Sentence (13b) has an RC with no marking.

13. a. ra'aitu rajul ili takra-hu maryam  
 Saw.1<sup>st</sup>S man that hate.3<sup>rd</sup>SF-him Mary  
 'I saw a man that Mary hates.'
- b. ra'aitu rajul takra-hu maryam  
 saw-1<sup>st</sup>per-sing man hate.3<sup>rd</sup>SF-him Mary  
 'I saw a man Mary hates.'

The internal structure of RCs seems to be similar to that of matrix sentences. Thus all of the sentences in (14) are grammatical, save (14e), which is with a full DP rather than a resumptive pronoun as direct object. Sentence (14a) shows that cliticization of the direct object pronoun onto the verb occurs in RCs, too. Therefore, the underlying structure of sentence (14a) seems to be SVO. Here, *hu* ('him') refers to the head noun *rajul* ('man'), but was base generated as the DP complement of the verb *takra* (hate.3<sup>rd</sup>SF).

14. a. ra'aitu rajul (ili) takra-hu maryam **VOS (surface)**  
 Saw.1<sup>st</sup>S man (that) hate.3<sup>rd</sup>SF-him Mary  
 'I saw a man that Mary hates.'
- b. ra'aitu rajul (ili) maryam takra-hu **SVO**  
 saw.1<sup>st</sup>S man (that) Mary hate.3<sup>rd</sup>SF-him  
 'I saw a man that Mary hates.'
- c. ra'aitu el-tawila(ili) hatat maryam li-ktab 'alay-ha **VSDoIo**  
 saw-1<sup>st</sup>S the-table(that) put.3<sup>rd</sup>SF Mary the-bookon-it.SF  
 'I saw the table that Mary put the book on.'

- d. ra'aitu el-tawila(ili) maryam hatat li-ktab 'alay-ha **SVDoIo**  
 saw-1<sup>st</sup>S the-table(that) Mary put-3<sup>rd</sup>SF the-bookon-it.SF  
 'I saw the table that Mary put the book on.'
- e. \*ra'aitu el-tawila(ili) hatat li-ktab maryam 'alay-ha **\*VDoSIo**  
 saw-1<sup>st</sup>S the-table(that) put-3<sup>rd</sup>SF the-book Mary on-it.SF  
 'I saw the table that Mary put the book on.'

If we assume a CP adjoining to a domain NP in restrictive RCs, C will be filled with the complementizer *ili* ('that'). If we examine again the structure proposed in Hypothesis A, we see that this structure prohibits the verb from undergoing I-to-C movement in order to achieve a verb-initial word order, as shown in sentence (14c), given here again:

14. b. ra'aitu rajul<sub>i</sub> [CP(ili) [IPmaryam takra<sub>v</sub>-hu<sub>i</sub> [VP t<sub>v</sub> t<sub>i</sub>]]]  
 saw.1<sup>st</sup>S man (that) Mary hate.3<sup>rd</sup>SF-him  
 'I saw a man that Mary hates.'
- 

Thus, the structure in Hypothesis A rules out VSO word order in RCs, while sentences (14a-d) exhibit perfectly grammatical sentences of that same structure. Yet again, the VP-internal Subject proposal in Hypothesis B does not rule out sentences such as those in (14), as they do not involve movement of the V outside of the IP.

So far, the two proposed word order variations were examined in light of possible constructions of *wh*-questions and restrictive RCs. It was shown that Hypothesis A rules out structures that are inarguably grammatical. Now embedded sentences in PA will be considered.

### 4.3 Embedded Clauses in Palestinian Arabic

The sentences in (15) show that the verb *'al* ('say') subcategorizes for a CP, headed by the complementizer *innu* ('that'). The subcategorization frame is given in (16):

15. a. ‘al ‘aḥmad innu ‘aga l-walad  
Said.3<sup>rd</sup>SM Ahmad that came-3<sup>rd</sup>MascSing the-boy  
‘Ahmad said that the boy came.’
- b. ‘aḥmad ‘al innu ‘aga l-walad  
Ahmad Said.3<sup>rd</sup>SM that came-3<sup>rd</sup>MascSing the-boy  
‘Ahmad said that the boy came.’
- c. \*‘al ‘aḥmad ‘aga l-walad  
Said.3<sup>rd</sup>SM Ahmad came-3<sup>rd</sup>MascSing the-boy  
‘Ahmad said that the boy came.’
- d. \*‘aḥmad ‘al ‘aga l-walad  
Ahmad Said.3<sup>rd</sup>SM came-3<sup>rd</sup>MascSing the-boy  
‘Ahmad said that the boy came.’

16. ‘al, V, [ \_\_ CP<sub>innu</sub>]

The sentences in (17) show that the only grammatical word order variations in clauses embedded under verbs such as ‘*al* are SVO, as in (17a,d) and VSO, as in (17b,e), but not VOS, as in (17c,f). This is true regardless of whether the matrix clause is subject-initial, as in (17a-c), or verb-initial, as in (17d-f). The noun *tefuha* (‘apple’) is a feminine noun, and thus different in features than the noun *walad* (‘boy’). This choice of noun was done in order to avoid two possible readings to the sentence, as the verb has to agree in person, gender and number with the subject.

17. a. ‘al ‘aḥmad innu el-walad akal et-tefuha  
Said.3<sup>rd</sup>SM Ahmad that the-boy ate.3<sup>rd</sup>MS the-apple  
‘Ahmad said that the boy ate the apple.’
- b. ‘al ‘aḥmad innu akal el-walad et-tefuha  
Said.3<sup>rd</sup>SM Ahmad that ate.3<sup>rd</sup>MS the-boy the-apple  
‘Ahmad said that the boy ate the apple.’
- c. \*‘al ‘aḥmad innu akal et-tefuha el-walad  
Said.3<sup>rd</sup>SM Ahmad that ate.3<sup>rd</sup>MS the-apple the-boy  
‘Ahmad said that the boy ate the apple.’

- d. ‘ahmad ‘al innu el-walad akal et-tefuha  
 Ahmad Said.3<sup>rd</sup>SM that the-boy ate.3<sup>rd</sup>MS the-apple  
 ‘Ahmad said that the boy ate the apple.’
- e. ‘ahmad ‘al innu akal el-walad et-tefuha  
 Ahmad Said.3<sup>rd</sup>SM that ate.3<sup>rd</sup>MS the-boy the-apple  
 ‘Ahmad said that the boy ate the apple.’
- f. \*‘ahmad ‘al innu akal et-tafuha el-walad  
 Ahmad Said.3<sup>rd</sup>SM that ate.3<sup>rd</sup>MS the-apple the-boy  
 ‘Ahmad said that the boy ate the apple.’

As with *wh*-questions and RCs, Hypothesis A predicts only the word order variation SVO in the embedded clauses. VSO and VOS are ruled out, as the C position occupied by the complementizer *innu* prohibits the verb from undergoing I-to-C movement. Therefore, Hypothesis A predicts VSO word order to be ungrammatical, and thus being too restrictive.

However, the VP-internal Subject proposal in Hypothesis B once again allows for SVO and VSO configurations to occur, since the verb can lie just within the Infl domain and still precede the subject. The ungrammatical structure in (17c) is ruled out by Hypothesis B, as VOS configuration must occur outside of the embedded IP: the verb must undergo I-to-C movement—thereby competing with the complementizer that can optionally lie in C—and its object must move to some projection within the IP. This ungrammaticality is therefore accounted for by Hypothesis B.

## 5 Theta Role Assignment and Case Filter

### 5.1 Theta Role Assignment

There is also a theory-internal advantage for the VP-internal Subject Hypothesis (Hypothesis B): theta role assignment. A predicate has a theta grid—a set of theta roles to

be assigned to different arguments. The theta criterion maintains that each argument must be assigned one and only one theta role. The internal theta role is assigned by the predicate to its complement by sisterhood. This is defined in Chomsky (1981) as theta-government. The external theta role, it was claimed, was assigned to the subject usually at Spec,IP, outside of the VP, with no apparent structural restriction such as c-command. With the subject at Spec,VP, it could be suggested that the predicate assigns the external theta role to the argument by m-command. The verb m-commands the DP: it does not dominate the subject, and the VP—the maximal projection that dominates the verb—does dominate the subject DP.

## 5.2 The Case Filter and the VP-internal Subject Hypothesis

One drawback of Hypothesis B is that it seems as though it does away with the Nominative Case Filter and the EPP (Extended Projection Principle) for the subject. If we accept Hypothesis B, thus allowing for the subject to remain in situ at Spec,VP in verb-initial orders, we must revise the Case Filter (and the EPP) at least for subjects in PA. A possible revision of both the VP-internal Subject Hypothesis and Case Filter would be to suggest that when the subject does not move to Spec,IP in order to get Case, a null subject is inserted into that same position, thereby satisfying the Case Filter and the EPP. The null subject is being licensed by being associated via a chain with the subject at Spec,VP. This coreference involves agreement feature sharing, and would enable this null subject to check features with the verb that has raised obligatorily to Infl. For further discussion on null subjects in Semitic, see Shlonsky (1997:109-143); for a revision of the original proposal, see also Rizzi (1986).

## 6 Conclusion

PA allows a set of word order variations in matrix clauses. This paper endeavored to find an underlying structure for these different configurations. Two different structures were proposed, one of which—the VP-internal subject hypothesis—is taken as a given in many syntactic analyses. Their prediction was examined in light of the structure of *wh*-questions and RCs, and lead to the conclusion that a D-structure where the subject is base generated at Spec,VP is the superior one. The VP-internal Subject Hypothesis accounts for the different word order variations, and rules out the ungrammatical constructions.

Now that this hypothesis has been established, it is necessary to pursue further issues in the syntax of PA. First, in this paper I assumed obligatory verb raising to Infl in order to check tense and agreement features. Independent evidence for such an assumption should be considered for this language. Second, in order to get VSO word order in PA, I proposed verb raising to Infl and an insertion of a null pronoun at Spec,IP. This null subject would share its features with the overt subject in Spec,VP, thereby enabling agreement feature checking with the verb in Infl and satisfying the Case Filter and EPP. This issue, considered by Shlonsky (1997) and others, should also be explored further, and with respect to other alternatives.

Reference

Keenan, Edward L. (1985) *Relative clauses*. In Timothy Shopen (Ed.), *Language Typology and Syntactic Description*, Vol, II (of 3 total): *Complex Constructions* (pp 141-170). Cambridge, UK: Cambridge University Press.

Kuroda S. Yuki (1988) *Whether We Agree or Not: A Comparative Syntax of English and Japanese*. In W. J. Poser (Ed.), *Papers from the Second International Workshop on Japanese Syntax* (pp 103-143). Stanford, CA: CSLI.

Mohammad, Mohammad A. (2002) *Word Order, Agreement and Pronominalization in Standard and Palestinian Arabic*. Philadelphia, PA, USA: John Benjamins Publishing Company.

Rizzi, Luigi. (1986) Null Objects in Italian and the Theory of Pro. *Linguistic Inquiry*, 17:501-557.

Shlonsky, Ur (1997) *Clause Structure and Word Order in Hebrew and Arabic: An Essay in Comparative Semitic Syntax*. New York, NY, USA: Oxford University Press.