

Pronominal Null Conjuncts in Arabic

Michael Hahn
University of Tübingen

Classical (CA) and Modern Standard Arabic (MSA, together henceforth simply 'Arabic') allow pronominal first conjuncts to be null. I will show that this phenomenon can be analyzed as an instance of a more general pattern of zero anaphora that may affect pronominal arguments or their first conjuncts. The basic invariance across pronominal arguments and pronominal first conjuncts, which is visible in the agreement of clitics and verbs and in the constraints on the distribution of zero anaphora, is accounted for by a new feature sharing mechanism. Similar phenomena, for which the analysis might also be interesting, are found in modern varieties of Arabic and other Semitic languages.

1 The Data

Subjects. Arabic pronominal subjects can be realized by a set of pronouns, which I will call *case-neutral pronouns* (case glossed as 'Ø'), or be dropped (1a). Case-neutral pronouns are also used as conjuncts, but in postverbal conjoined NPs, which give rise to first-conjunct agreement, pronominal first conjuncts may be null and have to be reconstructed from verbal agreement (1b):

- (1) a. 'atayta ('anta) came.2SG.M (you.Ø) 'you came'
 b. 'atayta [('anta) wa=Zaydun] came.2SG.M (you.Ø) and-Zayd.NOM 'Zayd and you came'

Pro-drop and null conjuncts are allowed by the same class of heads, i.e. by finite verbs, but in most syntactic environments not by adjectives and participles. The functional motivation for this constraint seems to be that nonfinite verb forms do not have person agreement:

- (2) a. dhaahib-un *('anta) going-SG (you.Ø) 'you are going'
 b. dhaahib-aani [*('anta) wa=Zaydun] going-DUAL (you.Ø) and-Zayd.NOM 'Zayd and you are going'

Complements. Pronominal complements can be realized by a free pronoun or a pronominal clitic. Besides case-neutral pronouns, Arabic also has free accusative pronouns (3, first row). An argument can also be realized by a clitic and a free pronoun at the same time (3-4, third column). As (3 d) and (4 b) show, a case-neutral pronoun in an oblique position has to be accompanied by a clitic:

- | (3) with clitic only | with free pronoun only | with clitic and free pronoun |
|---|---------------------------------------|--|
| a ra'aytu= ka
saw.1SG-you.ACC | b ra'aytu 'iyyaaka
saw.1SG you.ACC | c ra'aytu= ka 'iyyaaka
saw.1SG-you.ACC you.ACC |
| | d *ra'aytu 'anta
saw.1SG you.Ø | e ra'aytu= ka 'anta
saw.1SG-you.ACC you.Ø |

'I saw you' (all examples)

- | (4) with clitic only | with free pronoun only | with clitic and free pronoun |
|-------------------------------------|-------------------------------|---|
| a baytu= ka
house-you.GEN | b *baytu 'anta
house you.Ø | c baytu= ka 'anta
house-you.GEN you.Ø |

'your house' (all examples)

As noninitial conjuncts, pronouns are realized by an appropriate free pronoun. First conjuncts combine the behavior of coordinated subjects and simple complements. They can be realized by a free pronoun, by a pronominal clitic or by two morphemes simulatenously. As the examples show, a case-neutral pronoun again requires a clitic (5 d, 6 b):

- | (5) with clitic only | with free pronoun only | with clitic and free pronoun |
|--|--|---|
| a ra'aytu= ka [wa=Zaydan]
saw.1SG-you.ACC and-Zayd.ACC | b ra'aytu ['iyyaaka wa=Zaydan]
saw.1SG you.ACC and-Zayd.ACC | c ra'aytu= ka ['iyyaaka wa=Zaydan]
saw.1SG-you you.ACC and-Zayd.ACC |
| | d *ra'aytu ['anta wa=Zaydan]
saw.1SG you.Ø and-Zayd.ACC | e ra'aytu= ka ['anta wa=Zaydan]
saw.1SG-you.ACC you.Ø and-Zayd.ACC |

'I saw you and Zayd' (all examples)

- | | | |
|---|--|--|
| (6) with clitic only | with free pronoun only | with clitic and free pronoun |
| a baytu= ka [wa=Zaydin]
house-you.GEN
and-Zayd.GEN | b *baytu ['anta wa=Zaydin]
house you.Ø and-Zayd.GEN | c baytu= ka ['anta wa=Zaydin]
house-you.GEN [you.Ø and-Zayd.GEN] |

'your and Zayd's house' (all examples)

(5a) and (6a) correspond to null conjuncts in subject NPs. While the index features of the conjunct are realized by agreement in the case of subject NPs, they are realized by a clitic in the case of complement NPs. I will focus on the phenomenon of null conjuncts and how its distribution can be predicted.

Every pronoun is realized by at least one morpheme in Arabic, therefore nonnominative pronouns which are realized neither by a free pronoun nor by a clitic are ruled out:

- | | |
|--|--|
| (7) a. ra'aytu=*(ka)
saw.1SG-you.ACC
'I saw you' | b. ra'aytu=*(ka) [wa=Zaydan]
saw.1SG-you.ACC and-Zayd.ACC
'I saw you and Zayd' |
|--|--|

2 The Underlying Structure

It can be noted that a clitic and a corresponding free pronoun need not be adjacent as in the previous examples:

- (8) yahtiku=haa nnaasu [hiya wa=saa'ira 'ahlihaa]
shame-she.ACC the.people [she.Ø and-rest.ACC of.her.family]
'people shame her and the rest of her family'

On the other hand, a free pronoun as a first conjunct is always adjacent to the second conjunct. This suggests that the conjoined NPs in (1-8) form a constituent, from which a clitic realizing the first conjunct is excluded, as marked by the bracketing in (1-2) and (5-6).

Furthermore, it can be noted that there is a parallelism between null conjuncts in subject NPs and *pro*-drop observed with simple NPs. In both cases, the dropped element is the subject pronominal which is used for verbal agreement, and null realization is possible only with finite verbs. This suggests that subject null conjuncts are *pro*-elements:

- | | |
|-----------------------------|-----------------------------------|
| (9) 'You came' | 'Zayd and you came' |
| a. 'atayta ['anta] | 'atayta ['anta wa=Zaydun] |
| b. 'atayta [<i>pro</i>] | 'atayta [<i>pro</i> wa=Zaydun] |
| you.came you | you.came you and.Zayd |

It is straightforward to assume the same status for nonnominative null conjuncts. This entails that oblique pronominals can be *pro* if they are accompanied by a clitic. Further exploiting the analogy between simple pronominals and first conjuncts, I claim that pronominal arguments which are realized by a clitic without being a conjunct should also be analyzed as *pro* elements. The parallelism between the postulated *pro* element and overt pronouns in simple and conjoined NPs is illustrated by (10):

- | | | | |
|----------------------------------|---|---------------------------|------------------------------------|
| (10) 'I saw you' | 'I saw you and Zayd' | 'Your house' | 'Your and Zayd's house' |
| ra'aytu(=ka) ['iyvaaka] | ra'aytu(=ka) ['iyvaaka wa=Zaydan] | baytu=ka ['anta] | baytu=ka ['anta wa=Zaydin] |
| ra'aytu=ka ['anta] | ra'aytu=ka ['anta wa=Zaydan] | baytu=ka [<i>pro</i>] | baytu=ka [<i>pro</i> wa=Zaydin] |
| ra'aytu=ka [<i>pro</i>] | ra'aytu=ka [<i>pro</i> wa=Zaydan] | | |

One may now ask what the syntactic status of clitics, free pronouns and conjoined NPs, respectively, is and how they are related to the underlying argument they realize. It seems that free pronouns, conjoined NPs and the assumed (abstract) *pro* element are standard realizations of arguments, while clitics have a special status. Firstly, a conjoined NP represents the entire argument including arbitrary nonpronominal conjuncts, while a clitic can only represent a single set of index features. Secondly, free pronouns can have an appositional modifier which suggests that they can be heads and are syntactically similar to nouns. Thirdly, clitics are bound to the head while conjoined NPs and free pronouns can be separated from it as in (8), which is also true of other argument NPs in Arabic.

3 An HPSG Analysis

3.1 Licensing Null Conjuncts

Following [4] and others, I assume that the realization of arguments is determined by the subtyping of *synsem* into *canonical-ss* and *non-canonical-ss*. *Canonical-ss* objects are realized syntactically by a sign, while *non-canonical-ss*

objects which include gaps and *pro* are not realized and do not occur as the synsem value of *sign* objects. Gaps have a nonempty SLASH value, while *pro* is specified as PRO +, where PRO is a head feature. Marking the head value of *pro* will be useful to determine whether an NP contains a *pro* conjunct. Thus, any type of zero anaphora, including null conjuncts, is 'passively' licensed by allowing synsem objects to be *non-canonical-ss*. I will now show how the distribution of zero anaphora can be constrained.

Following previous HPSG analyses of coordination phenomena such as [2], I assume that coordination phrases have, in addition to their normal DTRS list, a CONJUNCTS list representing the *synsem* objects of the conjuncts. The noninitial conjuncts on CONJUNCTS are required to be marked by the coordination clitic *wa=*, which is enforced via the feature CRD ([1]). *wa=* is analyzed as a marker forming a constituent with the marked conjunct and therefore is not a daughter of the coordination phrase. Only the canonical elements of CONJUNCTS are mapped to DTRS:

(11)

$$\begin{array}{l}
 \text{coord-phrase} \rightarrow \\
 \left[\begin{array}{l}
 \text{CONJUNCTS} \quad \langle \left[\begin{array}{l} \text{LOC|CAT|CRD} \quad - \\ \text{LOC|CAT|CRD} \quad + \end{array} \right] \dots \left[\begin{array}{l} \text{LOC|CAT|CRD} \quad + \end{array} \right] \rangle \\
 \text{DTRS} \quad \langle \text{SYNSEM} \quad \boxed{1}, \dots \text{SYNSEM} \quad \boxed{72} \rangle
 \end{array} \right] \\
 \wedge \langle \text{list}(\text{SS|L|C|H|PRO+}) \oplus \langle \boxed{1}, \dots \boxed{72} \rangle
 \end{array}$$

This constraint applies to all types of coordination phrases, because there is no constraint on the syntactic categories of the conjuncts. Since conjuncts are allowed to be *pro*, pronominal null conjuncts are possible in principle. Noninitial conjuncts have to be marked, but since the only lexical item with the relevant marking is the coordination marker, they necessarily are phrases. This means that their head, the conjunct itself, is not null and only a first conjunct can be null.

This analysis accounts for the possibility of zero anaphora including null conjuncts, but it leaves open how the agreement of clitics and verbs with first conjuncts (whether null or not) can be derived, and how the distribution of zero anaphora can be constrained. The remaining part of the section will address these questions.

3.2 Deriving First Conjunct Agreement

Since the first conjunct of an argument NP influences the index features used for subject agreement and clitics, some features of the first conjunct must be visible for a head selecting the NP, i.e. must be accessible within the NP's synsem object. [8] make the index features of Portuguese conjuncts accessible by a special head feature. On the other hand, one could analyze first-conjunct agreement as resulting from a special pattern of resolution of the concord feature in conjoined NPs. However, the head must also be able to determine whether the pronominal argument requires a clitic, i.e. whether it is *pro* (ex. 7) or *case-neutral* (exx. 5 d, 6 b). This syntactic information must be mediated so that it can be distinguished from the corresponding information about conjoined NPs, since they must be non-*pro* regardless of the status of the first conjunct. Therefore, the use of a feature for mediating all the relevant syntactic information seems more adequate. I will assume that the full head value of the first conjunct is structure-shared with the value of the feature INTERNAL-HEAD (I-HD) of the full NP, which is appropriate for *cat*. Furthermore, I assume that the INTERNAL-HEAD value is identical to the head value in (most) other structures.

This allows a unified account of simple and conjoined arguments, since the head can always use the argument's internal-head value to decide about agreement and clitics. Thus, the index features used by clitics and verbal agreement are always those in INTERNAL-HEAD|CONCORD ([9]), irrespective of the internal structure of the argument NP.

This analysis also accounts for example (12), which shows that a conjoined NP can trigger resolved index features on verbs and clitics. This agreement pattern is possible if the NP contains a case-neutral pronominal conjunct:

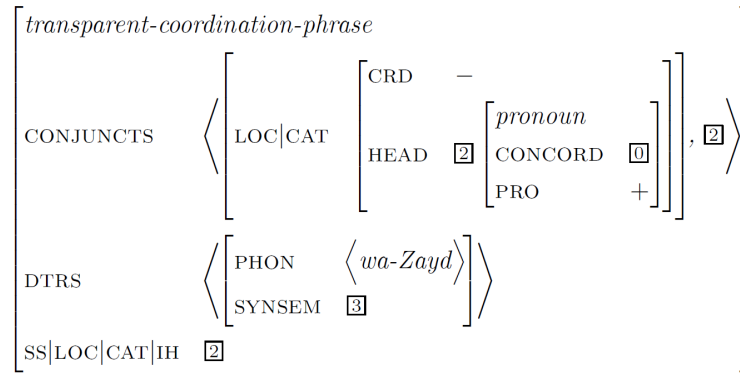
- (12) a. 'alay=naa ['anaa wa='anta] 'an... b. ji'naa ['Abbaas wa='anaa]
 upon-US.GEN I.Ø and-you.Ø that... came.1PL Abbas and-I.Ø
 'it is upon me and you to...' 'Abbas and I came'

I analyze them simply as conjoined NPs where the INTERNAL-HEAD value is shared with the HEAD value of the entire NP, which will have resolved index features, thus providing a uniform agreement mechanism for all NPs: agreement with a verb or a clitic is established by a uniform mechanism operating on the argument NP without any recourse to its internal structure. Technically, the distinction between 'transparent' coordination and the 'opaque' structure in (12) can be implemented by partitioning *coordination-phrase* into *opaque-coordination* and *transparent-coordination*, with *opaque-coordination-phrases* being required to have a case-neutral pronominal conjunct.

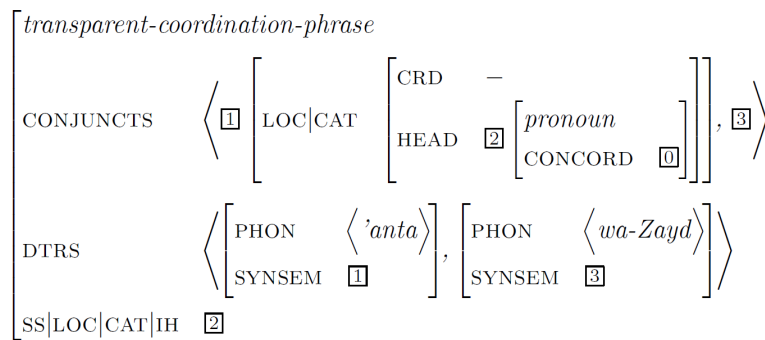
The analysis is illustrated by the phrases *wa=Zayd* "X and Zayd" with null conjunct (13 a), the transparent structure '*anta wa=Zayd* "you.Ø and Zayd" (13 b), its opaque counterpart (13 c) and the pronoun '*anta* "you.Ø" (13 d). The case ending of the name *Zayd* is omitted since the structure is case-independent. The index in SYNSEM|LOC|CAT|

INTERNAL-HEAD|CONCORD, which is used for agreement and clitics, is marked by [0]:

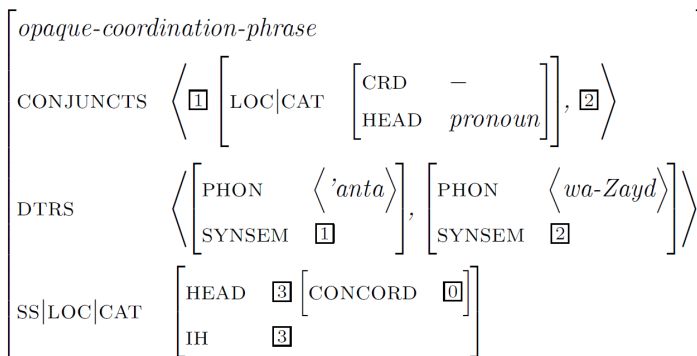
(13) a.



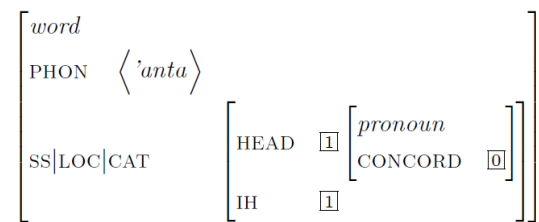
b.



c.



d.



In all four structures, the index used for agreement or a clitic is [0], but its source depends on the internal structure of the NP: In (a), this is the index of the null conjunct, in (b) the index of the overt first conjunct, in (c) the resolved index of the opaque coordination phrase and in (d) the index of the pronoun.

3.3 Constraining the Distribution of Zero Anaphora

It was shown above that the distribution of null arguments and null conjuncts is constrained along two dimensions, both of which have a straightforward functional interpretation. Since both generalizations apply to both pronominal arguments and pronominal first conjuncts of arguments, they can be formalized by constraining the INTERNAL-HEAD value of arguments:

(14) a. A nominative pronoun can be null only if the head is a finite verb (Examples 1-2).

b. A nonnominative pronoun can be null only if there is a corresponding clitic on the head, i.e. the pronoun must be realized by at least one morpheme.

(14 a) can be implemented straightforwardly by constraining the subject arguments of nonfinite predicates to have INTERNAL-HEAD|PRO –. Formalizing (14 b) requires an account of clitical pronouns. It seems that bound pronouns in Arabic, unlike pronominal affixes in some Romance languages [5], are really clitics. While the affix criteria introduced by [10] are not satisfied, there are striking similarities linking bound pronouns and genitive NPs: they are

subject to linearization restrictions operating on the set of bound pronouns *and* genitive dependents of a head, require the same morphological form of the head, do not have wide scope over coordination and are subject to similar restrictions connected to binding theory. Technically, these restrictions can be implemented by retrieving all clitics and genitive arguments in a single phrase of type *head-bound-arg-phrase* with two or three daughters.

Generalization (14b) can then be stated as follows (How the quantificational statement is formalized is immaterial in this context; besides quantifying over elements of ARG-ST, one could also define subtypes of *head-bound-arg-phrase* for different numbers of clitics):

(15) For every element of ARG-ST satisfying

$$\left[\text{LOC|CAT|IH} \left[\begin{array}{ll} \text{CASE} & \boxed{1} \neg nom \\ \text{CONCORD} & \boxed{2} \\ \text{PRO} & + \end{array} \right] \right]$$

there must be a clitic with case $\boxed{1}$ and concord $\boxed{2}$.

These two constraints suffice to predict the ungrammaticality of the starred versions of (2) and (7). The generalization that case-neutral pronouns have to be accompanied by a clitic in nonnominative positions, which explains the ungrammaticality of (3d, 4b, 5d, 6b), can be implemented by similar constraints.

4 Discussion

There are several successful analyses of clitic ordering in various languages using tools of linearization-based HPSG, e.g. [3]. In fact, the existence of apparent null conjuncts after clitics and the multiple, discontinuous realization of one argument could well be explained in a linearization-based setup. One might assume that pronouns may introduce several types of domain objects simultaneously, including those representing free and bound pronominal forms, respectively. Null Conjuncts following clitics could be analyzed as pronouns introducing only a clitical domain object. However, assuming that clitics are introduced by pronouns, while agreement is obviously a morphological process internal to the verb, leaves unclear how the apparent analogy between subject agreement and clitics with regard to first conjuncts can be captured. While this issue is only of a conceptual nature, it seems that a more basic property of this idea is problematic in the case of Arabic, namely the idea that there are no null conjuncts on the level of constituent structure.

In his grammar of Classical Arabic, the eighth-century scholar Sibawayh gives examples with divergent case-marking on what appear to be two conjuncts [7]:

- (16) a. darbu [Zayd-in wa='Amr-in] b. darbu Zayd-in wa='Amr-an
 beating [Zayd-GEN and-Amr-GEN] beating Zayd-GEN and-Amr-ACC
 'beating Zayd and Amr'
- (17) a. maa 'ataa=nii ghayr-u [Zayd-in wa='Amr-in] b. maa 'ataa=nii ghayr-u Zayd-in wa='Amr-un
 not came-me other-NOM [Zayd-GEN and-Amr-GEN] not came-me other-NOM Zayd-GEN and-Amr-NOM
 'Nobody came to me but Zayd and Amr'

In the (b) examples, the two conjuncts have different case-marking, while the (a) examples, which apparently had the same interpretation, are regular. It can be noticed that in both examples, there is a verbal head that could itself assign the case of the second conjunct, and that there is a correlation between the position in argument structure and the case of the second conjunct. In (16), the NP is part of a complement and Arabic verbal nouns are in general allowed to realize at most one complement as genitive NP, while the others receive accusative. In (17), the NP is part of the subject, and Arabic verbs indeed allow more than one nominative dependent, as shown by clauses with nominative expletives [6].

This observation suggests that what is operative here is not some kind of optionality of case agreement in coordination structures, but the normal mechanism of case assignment by which the two conjuncts receive different case-marking from their head. This will follow from the standard HPSG case assignment mechanism if we assume that the first conjunct and the remainder (i.e., *wa=* and the second conjunct) occupy separate positions in the valence lists of the head and thus do not form a constituent. (16-17) thus provide evidence that Classical Arabic had coordination constituents that lack an overt subconstituent corresponding to the first conjunct, i.e. one may assume that the phenomenon of null conjuncts existed at the level of constituent structure. Although we have to leave open how (15-16) can precisely be analyzed, since almost no additional data is available and since there is no evidence for divergent case-marking in MSA, the examples seem to show that an analysis of null conjuncts at the level of constituent structure is required in an 'ideal' grammar of Classical Arabic, thus, there seems to be little motivation for designing an additional linearization-based mechanism to capture the same phenomenon in the case of Arabic pronouns.

References

- [1] John Beavers and Ivan A. Sag. Coordinate Ellipsis and Apparent Non-Constituent Coordination. In Stefan Müller, editor, *The Proceedings of the 11th International Conference on Head-Driven Phrase Structure Grammar*, pages 48-69, Stanford, 2004. CSLI Publications.
- [2] Rui Chaves and Denis Paperno. On the Russian hybrid coordination construction. In Stefan Müller, editor, *The Proceedings of the 14th International Conference on Head-Driven Phrase Structure Grammar*, pages 46–64, Stanford, 2007. CSLI Publications.
- [3] Berthold Crysmann. Clitics and coordination in linear structure. In Birgit Gerlach and Janet Grijzenhout, editors, *Clitics in Phonology, Morphology, and Syntax*, pages 121–159. John Benjamins, 2000.
- [4] Christopher D. Manning and Ivan A. Sag. Argument structure, valence, and binding. *Nordic Journal of Linguistics*, 21, 1998.
- [5] Philip H. Miller and Ivan A. Sag. French clitic movement without clitics or movement. *Natural Language and Linguistic Theory*, 15(3):573–639, 1997.
- [6] Hermann Reckendorf. *Arabische Syntax*. Heidelberg, 1921.
- [7] Sibawayh. *Le Livre de Sibawayhi, Traité de Grammaire Arabe*. Edité par H. Derenbourg. Paris, 1881.
- [8] Aline Villavicencio, Louisa Sadler, and Doug Arnold. An HPSG account of closest conjunct agreement in NP coordination in Portuguese. In Stefan Müller, editor, *The Proceedings of the 12th International Conference on Head-Driven Phrase Structure Grammar*, pages 427–447, Stanford, 2005. CSLI Publications.
- [9] Steven Wechsler and Larisa Zlatic. *The Many Faces of Agreement*. Stanford, 2001.
- [10] Arnold M. Zwicky and Geoffrey K. Pullum. Clitization vs. Inflection: English n't. *Language*, (59):502–513, 1983.