Pronouns, agreement and focus in Egyptian Arabic
Malcolm Edwards
m.edwards@bbk.ac.uk

This paper considers the ‘copula pronouns’ of a dialect of Arabic in the light of recent proposals regarding the emergence of functional categories, and in particular agreement marking. It is argued that the sometimes contradictory properties of copular pronouns are compatible with an analysis in which they develop from actual subject pronouns which have undergone a process of reanalysis from elements in Specifier position to heads of functional categories. This process is cyclical, and has implications for the development of functional categories in general and agreement marking in particular.

Introduction
In spoken Egyptian Arabic (EA) present tense equational sentences have the form illustrated in 1-3.

1. il-bint hiyya l-mas’u:la
   the-girl she the-responsible(FemSg)
   the girl is the one responsible

2. il-riga:la humma il-mas’uli:n
   the-men they the-responsible(Pl)
   the men are the ones responsible

3. ‘ana huwwa l-mas’u:l
   I he the-responsible
   I am the one responsible

As these examples show, equational sentences have the schematic structure DP-PRON-DP, where PRON (emboldened in the examples) is apparently a pronoun functioning as a copula element. The grammar of equational sentences in both Arabic and Hebrew has been widely discussed (see for example Eid 1983, Eid 1991, Doron 1986, Ouhalla 1999, among others), with interest centering on the categorical and functional status of PRON. PRON presents problems for analysis, as it displays a mixture of nominal and verbal properties. PRON has been claimed to be either a suppletive form of copula (Eid 1983, 1991), or following Doron’s (1986) agreement marking in INFL. This latter idea has been explored in the context of Arabic, by Eid (1991, 1992) and Ouhalla (1999), among others.

The present paper proposes considers the status of PRON in the light of recent proposals regarding the reanalysis of elements in Spec positions as functional heads (see in particular Simpson and Wu 2002, and Van Gelderen 2004). Specifically it is argued that PRON derives from a cyclical process whereby monomorphemic specifier elements can undergo reanalysis as functional heads. This process in turn can be explained in terms
of an economy condition privileging heads. The analysis also has implications for the development of agreement marking and functional categories.

1. The Distribution of PRON
EA has no present-tense form of the copula verb, with the result that simple predications such as ‘Magdi is an engineer’ are verbless constructions, consisting of a definite Subject, followed by an indefinite predicate:

(4) magdi muhandis
    Magdi engineer
    Magdi is an engineer

Equative sentences in EA are also verbless constructions, and consist of two definite DPs, separated by a nominative pronoun (PRON). The inventory of PRON elements is restricted: all instances of PRON are third-person pronouns, and PRON agrees with the first DP for Gender (M/F) and Number (Sg/Pl).

(5) magdi huwwa il-muhandis
    Magdi he(PRON) the-engineer
    Magdi is the engineer

(5b) il-mulkila hiyya T-Talaba
    the-problem(FemSg) she (PRON) the-students
    the problem is the students

In addition to this morphological restriction, PRON also has a highly restricted distribution, appearing only in the equative construction. PRON cannot occur in sentences expressing simple predication:

(6) *magdi huwwa muhandis/za9la:n/fi l-maktab
    Magdi PRON    engineer/upset/in the-office
    Magdi     is an engineer/upset/in the office

The fact that PRON does not appear in predicational sentences distinguishes it from the copula verb ka:n (‘be’), which appears in both predicational and equative sentences.

(7) magdi ka:n il-muhandis/muhandis/za9la:n/fi l-maktab
    Magdi was the-engineer/an engineer/upset/in the-office

One further property of PRON merits discussion at this point. PRON is negated with the ‘split’ negative marker ma-/. Split negation is the form of negation associated with verbs in EA.
(8) magdi ma-huwwa:∫ il-muhandis
    magdi Neg-Pron(3SM)-Neg the-engineer
    Magdi is not the engineer

(9) faTma ma-hiyya:∫ il-mas’ula
    Fatima Neg-Pron(3SF) the-responsible
    Fatima is not the one responsible

To summarise: PRON has the morphological form of third-person pronouns; it is in apparent complementary distribution with the copular verb ka:n (which has no present-tense form); PRON shows limited agreement with the first DP in the clause; PRON is negated using the form of negation used with inflected verb forms. Superficially, these properties suggest that PRON can be identified with the copula. Several analyses have taken just this view, proposing that PRON is indeed a predicate, and specifically a suppletive form of the copula (eg. Eid 1983), or, more vaguely, that it is a quasi-verbal element (see, for example, Cowell 1966, Brustad 2000, among others). The implications of treating PRON as a predicate are the subject of the next section.

2. Two Views of PRON: Predicate or Agreement?
The notion that PRON in EA is a suppletive copula form is advanced by Eid, 1983. In later work, Eid (1991) attributes a more specific function to PRON – that of an ‘identity predicate’, functioning to signal identity of reference between the Subject and the Predicate. Eid proposes the structure (10) below for equatives with the identity predicate:

(10)

Among the advantages Eid claims for this analysis are that it distinguishes PRON from verbs, by generating PRON as the head of a NP which is itself a predicate. In this way the pronominal nature of PRON, and its apparent predicational function are reconciled, and both receive an explanation. Eid’s analysis presents problems, however. Firstly, in claiming that PRON and the second, ‘predicate’ NP in the construction are a single constituent, the analysis posits a construction in which a pronoun appears as a proper
subconstituent of a NP. Such constructions are otherwise unattested in EA. In fact, there are no constructions in EA in which a predicate element appears inside an NP. The case is further weakened by the fact that PRON is invariably negated with the discontinuous negative *ma-* / as there are no reasons to suppose that a pronoun inside a NP, should such a structure be possible, could be the target of negation independently of its containing phrase. Additional problems for the analysis arise in connection with the agreement features of PRON. Eid claims that the type of agreement found in these constructions is not subject-verb agreement, but agreement within a noun phrase. Here a parallel could be drawn with agreement between demonstratives and heads of NPs, but Eid’s proposal is called into doubt by the fact that the type of agreement found in equative constructions resembles subject-verb agreement, in that it is the first DP in the sentence, rather than the supposedly containing DP, which determines the gender and number of PRON.

(11a) il-mu[ǐ]kila hiyya iT-Talaba
    the-problem(F) she the-students
    the problem is the students

(11b) iT-Talaba humma il-mu[ǐ]kila
    the-students they the-problem
    the students are the problem

These agreement requirements suggest that PRON is a sentential constituent, rather than a subconstituent of a NP. Finally, it can be pointed out that Eid’s analysis assumes a functional role for PRON, specifically that PRON is a predicate. The analysis fails, however, to offer an explanation as to why PRON, in formal terms, a pronoun, should have a functional role. This feature of PRON requires a consideration of the informational characteristics of equative constructions in EA, to which we now turn.

2.2 PRON as Agreement

Ouhalla (1999) explores the relationship between PRON and information structure, and in particular the role of focus. Ouhalla points out that the forms taken by PRON are those of the strong (independent) pronouns in Arabic, and that pronouns of this type generally have either an emphatic or contrastive function. This is the case in EA, a pro-drop language in which strong pronouns (identical with the class of subject pronouns) typically serve functions which are primarily pragmatic, including disambiguation, subject switch, and contrastive focus (cf. Eid 1983, Jelinek 2002). Strong pronouns contrast with both weak (enclitic) pronouns and agreement inflection precisely in the fact that strong pronouns can be the target of focus. Focussing of a weak pronoun, for example, can only be realized by doubling the weak pronoun with a matching strong pronoun. The pattern of phonological stress illustrated in (10), in which the putative target of phonological stress (and hence focus) is capitalised, is not grammatical. In (11) the strong (independent) pronoun is the focused item, and receives phonological stress.

(12) *uf-[ǐ] ṭ-U
    saw(TS)-HIM
Regarding the syntactic locus of PRON, Ouhalla suggests that the Agr/I element of equatives in Arabic includes the focus feature [+f], and following Doron 1986, that PRON is in INFL. As Ouhalla observes, and as has been demonstrated above for EA, the fact that strong pronouns may be assigned a focus feature distinguishes them from clitics, on the one hand, and inflection on the other, and Ouhalla attributes the lack of an inflectional realization of agreement in equatives to this distinction. In support of this argument, Ouhalla cites the fact that in Moroccan Arabic clefts, PRON can carry the focal stress. This is also true of equatives in EA, in which PRON can be focused:

(14) magdi HUWWA il-muhandis (mi sami:r) 
    MAGDI Pron the-engineer (not Samir) 
    (it’s MAGDI who’s the engineer (not Samir)

In this way, Ouhalla simultaneously explains the association of PRON with focus, and the relationship between PRON and INFL, echoing the analysis of PRON in Hebrew proposed by Doron (1986), who provides arguments based on evidence from that language that PRON is an agreement clitic situated in INFL.

Ouhalla’s analysis has much to commend it, but it raises further questions. Firstly, it seems likely that, for EA at least, to view PRON as necessarily the carrier or indicator of focus is erroneous. Equatives with PRON appear to be true equatives, which can undergo inversion with no grammatical consequences, and with no alteration to the propositional content of the sentence. If these constructions are indeed true equatives, we would expect them to be both reversible, and informationally neutral. This appears to be the case. While Magdi huwwa il-muhandis and il-muhandis huwwa magdi can be said to differ in what they are ‘about’, one being about Magdi, and the other about the engineer, this difference is essentially context-dependent. Equally, the informational properties of the sentence magdi huwwa il-muhandis can vary according to context. ‘Magdi’ may be focused when the sentence is used to assert that it is Magdi, and not anyone else, who is the engineer. Alternatively, il-muhandis may be focused if what is being asserted is that Magdi is the engineer, and not, say, the architect. Further evidence that the informational properties of the construction are contingent can be seen in sentences such as (8):

(15) il-mulkila hiyya ‘inn-ak mi 9a:yiz 
    the-problem she(Pron) that-you Neg want 
    the problem is that you don’t want to [do s.t.]

The claim that PRON is located in INFL has been proposed, and adopted, by Doron (1986) (for Hebrew), Eid (1983, 1991), and Benmamoun (1992), among others.
What is asserted by the sentence in (8) appears not so much to be a case of identity between the two phrases, but rather one of specification (cf. Heycock and Kroch, 1999). It is presupposed that there is a 'problem', and the post-Pron constituent specifies what that problem is. It seems likely, therefore, that while PRON may be selectively focused, it is not invariably associated with focus.

We turn now to the proposal that PRON is located in INFL. This proposal, too, poses problems, related to the nature of agreement with PRON, the distribution of negation, and the assignment of focus. With regard to agreement, it has already been noted that the paradigm of PRON is deficient. Although the forms that PRON takes are, formally, those of regular third-person pronouns, the clearest evidence that PRON is not to be identified with the class of pronouns, can be seen in the fact that the elements which constitute PRON do not correspond to the full pronominal paradigm.

PRON displays agreement only for gender (M/F) and number (S/Pl), agreement for person being effectively neutralized. Sentences (19)-(20) illustrate:

(16) ‘ana huwwa il-Hara:mi
I Pron(3SM) the-thief
I’m the thief (‘Am I the thief?’)

(17) ‘inti hiyya il-mas’u:la
you(FS) Pron(3SF) the-responsible
You are the one responsible

The restricted agreement found with PRON is problematic for the claim that PRON is a direct spelling-out of Agr features in I. It is clear that the absence of agreement for person indicates that the type of agreement found with PRON differs from standard subject-verb agreement, as would, for example, be found with the verbal copula.

Negation facts also call into question the claim that PRON is located in INFL. The elements which can take ‘split’ negation are typically of two types – predicates, including verbs, and subject pronouns, including PRON. In the case of verbs, split negation surrounds the verb, complete with any inflectional material and clitics that may be attached to it. It is reasonable to assume that verbs, together with both inflectional affixes, and any enclitics, must raise to Neg, which in turn suggests that the Neg is higher in the sentence than INFL. If this is so, then PRON must also be assumed to raise to Neg.2

Finally, while the idea that INFL is associated with focus, and hence hosts focused PRON, is tempting, it begins to appear somewhat ad hoc when we consider that other elements in INFL are not accessible to focus. As we have seen, inflectional material on verbs, for example, cannot be selectively focused, but must be doubled by a focused pronoun. Similarly, the claims that Agreement is spelt-out in equative constructions, but not, for example, in simple predicational structures, or that the INFL of equatives is uniquely allocated the +f feature do not receive independent support.

In conclusion, the evidence that PRON is situated in INFL is at best ambiguous, and requires a number of supporting assumptions which cannot be independently justified.

---

2 It appears that there is, as suggested by Benmamoun (2000) a close relationship between Neg and the Subject, such that Neg must discharge a +D feature. This idea will be taken up subsequently.
3. PRON and the Subject
To summarise the discussion so far, the sometimes contradictory properties of PRON to be explained are the following:

- PRON occurs only in equative constructions
- PRON is, or appears to be, in complementary distribution with verbs
- PRON is pronominal, but appears to function as a copula
- PRON is negated with the split negative associated with verbs
- PRON shows restricted agreement with the subject

As will be demonstrated, these properties of PRON are compatible with an analysis of PRON which assumes PRON to be what it appears to be – a Subject pronoun, albeit one which has undergone reanalysis.

3.2 Negation and PRON
The interaction of PRON with negation provides an indication as to its likely status.

(18) magdi ma-huwwa:-ʃ il-muhandis
    magdi Neg-Pron(3SM)-Neg the-engineer
    Magdi is not the engineer

(19) faTma ma-hiyya:-ʃ il-mas’u:la
    Fatima Neg-Pron(3SF) the-responsible
    Fatima is not the one responsible

The fact that the split negative is the standard form of negation used with inflected verbs has led some linguists to suggest that PRON is a ‘quasi-verb’ (see, for example, Cowell 1966, Brustad 2000). However, as Benmamoun (2000) has noted, the distribution of split negation in EA is closely linked to the syntax of the subject. Jelinek (1981) observed that split negation in EA attaches to a range of constituents, belonging to a number of distinct categories, all of which have the property that they mark what Jelinek calls ‘person subject’. The essence of this observation is that the split negative appears on constituents encoding person, number and gender features associated with the Subject. Constituents which can be negated in this way include inflected verbs, pronominal elements which realize Subject features, and a set of ‘quasi-verbal’ predicates such as 9and-u (literally ‘at him’) expressing possession. Importantly, elements which do not encode the full set of subject features cannot be negated with the split negative. Thus participles, which generally function as verbal predicates, and can accept the same range of complements as verbs, do not permit this form of negation.

The association of split negation with focus may also be discerned in the phonology of negated PRON. We have already seen examples of negated PRON in equatives, but, as the examples below show, subject pronouns can be negated in non-equatives.

Eid (1991) calls pronouns such as these as ‘negative pronouns’, and notes that the set of negative pronouns is exactly congruent, in terms of featural content, with the set of independent pronouns. The distinguishing property of the negative pronouns is that
attachment of negation triggers a lengthening of the final syllable of the pronoun throughout the paradigm, and this feature appears to be associated with focus: independent (non-negative) pronouns may be phonologically stressed for reasons of focus or contrast, but the stress, in this case, will fall on the first syllable of the pronoun, and never on the second syllable.

Unlike equative PRON, negative pronouns can appear with any predicate, and they represent an alternative negation strategy to that most commonly found in Predicational sentences, whereby the independent negative marker $mi$/$mu$ is placed before the predicate:

(20) huwwa miʃ Hara:mi

he Neg thief

he’s not a thief

(21) huwwa ma-huwwa-ʃ Hara:mi

he Neg-Pron-Neg thief

he’s NOT a thief

The fact that negative pronouns can co-occur with subject pronouns suggests that they represent an emphatic ‘doubling’ of the subject. The nature of this doubling, and its implications for the structure of equatives, is discussed in the next section.

3.3 From Spec to Head

Simpson and Wu (2002) present evidence from a variety of languages that functional categories may evolve from focusing constructions. The process starts when a functional head selects for a (lower) focus projection, which initially has true semantic content. Over time, the focus projection loses semantic content, and decays into an agreement (or concord) element. The development of negation in French offers a clear illustration of the process. Simpson and Wu claim that Neg in French selected for a lower focus position, containing the lexical item $pas$, originally a marker of emphasis, with real semantic content. In Simpson and Wu’s analysis, the association of Neg and the lexical item $pas$ results in a two-part shell structure, in which the two morphemes are associated with two layers of structure, but realize a single (functional) value. The two layers then become related as an instance of agreement, and the original emphatic and focusing role of one, in this case $pas$ is lost. In turn, the higher head, here $ne$, may weaken to become an expletive element, while the lower head becomes the primary realization of the original semantic value of the construction. The next step is for the two-part shell structure to collapse into a single functional projection, headed by the earlier focused

---

3 As the gloss to example (18) suggests, negative pronouns are associated with ‘emphatic negation’ (Salib 1986), and are not simply a neutral alternative to negation with $mi$/$mu$. Jelinek (2002) claims that in EA, the marker of Neg(ation) becomes ‘the focus peak’ of the clause, and the remainder of the sentence is thus the ‘frame’. The details of Jelinek’s claim are not wholly clear, although the association of negation with focus is attested in many languages.
head. Thus, in modern French, *pas*, originally a lexical focusing element, has become the primary expression of negation.

In the latter part of Simpson and Wu’s paper, these cyclical mechanisms are argued to account for the development of subject agreement. A key mechanism in this analysis is ‘Spec-Head reduction’, 4 which is defined by Simpson and Wu as follows: ‘a morphologically simplex element…. in the specifier position of a functional head X0 may over time be reanalysed as instantiating X0 . . . .’ (Simpson and Wu, 2002, 308). In the case of subject agreement, Simpson and Wu propose that subject pronouns, base-generated as subjects in Spec vP, may be grammaticalised as agreement morphemes, and subsequently reanalysed as base-generated lexical elements in the head v0. In a footnote, Simpson and Wu remark on the use of pronouns as copulas in Arabic, suggesting that such pronouns instantiate verb-like ‘semifunctional’ heads.

The relevance of Simpson and Wu’s proposals to the analysis of PRON is as follows. Among the peculiarities of PRON are that it has the form of subject pronouns, that it displays limited agreement, and that it has a functional (or semifunctional) role, as a copula element. To these properties we can add the fact that PRON is found only in equative constructions. These properties of PRON can be explained on the assumption that equatives in EA have the form that they do because they derive from an original left-dislocation construction which has been reanalyzed as a copular construction. Left-dislocated structures are common in EA, and serve a range of informational, and in some cases, grammatical purposes (cf. Eisele 1999).5 All instances of left-dislocation feature a left-displaced definite DP, followed by a string constituting a complete predication, and containing a resumptive pronoun associated with the displaced DP. The initial structure of equatives is that diagrammed in 21. Here the leftmost DP is in topic position. In the case of equatives, these can be seen as a dislocation of a definite DP, with concomitant resumptive pronominalisation, in the form of a strong pronoun. Pursuing this idea, the development of equatives would start with the (slightly simplified) structure in (21a).

---

4 Van Gelderen (2004) proposes that analogous processes of Spec-Head conversion, as well as Head-Head conversion are a widespread feature of language change, and account for the forms of negative and wh-constructions in English and other languages.

5 Cowell (1964) proposes a similar explanation for equatives featuring PRON in Syrian Arabic. Cowell treats equatives as instances of ‘extraposition’, in which the initial DP has been moved leftwards.
Here the initial DP is in Topic position, and is ‘resumed’ by the pronoun *huwwa* in Spec vP. The siting of *huwwa* in Spec vP is in line with a suggestion by Chomsky (1995), and allows the pronoun, as a regular subject, to receive Nominative Case and its $\vartheta$-role.

The next stage is diagrammed in 21b. Here the pronoun has undergone reanalysis, and has become the functional (or semifunctional) head of vP.

If this analysis is along the right lines, it explains the combination of subject-like and semifunctional, predicate-like properties of PRON. It is important to stress, however, that it is not being claimed that PRON is a predicate. If it were a predicate, it would presumably form a constituent with the second (rightmost) DP. As 22 suggests, PRON
and the DP do not form a constituent, because modals such as la:zim ‘must’ cannot intervene between a predicate (or predicator) and its complement.  

(22) il-walad huwwa la:zim il-mas’u:l
   the-boy PRON MUST the-responsible
   the boy must be the one responsible

More generally, the analysis of PRON presented here claims that the properties of PRON reflect its original function as a subject, which has become ‘fossilised’ within a specific, informationally marked, construction type. The formal details of the analysis owe much to Simpson and Wu, but merit further comment. The assumption that PRON, originally generated in (or merged in) Spec vP, has been reanalyzed as the head of vP, is somewhat stipulative, and requires a slight, but essentially harmless, extension of the typical status and function of v-projections. An alternative analysis, in which PRON is generated or merged at Spec TP, and subsequently becomes the head of TP can be questioned on the grounds of the incongruity of having a nominal element occupying the head of TP. Also, when T is realized within an equative construction by the copular verb ka:n/yiku:n, PRON cannot appear. This restriction does not follow from a general ban on the appearance of two realizations of tense within the clause, as tensed verbs can appear with the copula, but reflects a restriction on subject-marking to the effect that the features of the subject can only be marked once – either as inflection, or as PRON.

One final observation concerns the notion of ‘multiple subject constructions’ (MSCs) discussed by Doron and Heycock (1999). Doron and Heycock point out that in Semitic, and other languages including Japanese, there are sentence types which contain more than one nominative noun phrase. Constructions containing PRON are not treated in Doron and Heycock’s paper, but appear, superficially at least, to be strong candidates for membership of this class of sentence. Doron and Heycock present arguments that in the data they review, MSCs are not analyzable as instances of dislocation. There are clear affinities between the notion of multiple subjects and the analysis of PRON presented above, which it is beyond the scope of this paper to explore. For the present, it can be pointed out that in Doron and Heycock’s data, examples of MSCs contain two lexical, and non-coreferential noun phrases, whereas PRON is (necessarily) pronominal, and as such represents a ‘doubling’ of the initial DP.

References
Benmamoun, El-Abbas (2000)

---

6 Adger and Ramchand 2003 present a detailed analysis of pronominal ‘copula’ elements in Scottish Gaelic as predicate elements.


Jelinek, Eloise (2001): Agreement, Clitics and Focus in Egyptian Arabic. In Ouhalla and Shlonsky (eds)


Simpson, Andrew and Zoe Wu (2002): Agreement, Shells and Focus. Language Vol. 78, 2. 287-313