ADPOSITIONAL CLITICS AND WORD ORDER IN SOMALI

By John Ibrahim Saeed
Trinity College, Dublin
(Received 25 February 1992; Revised 29 July 1992)

1. Introduction

This paper emerges from an ongoing study of verbal derivational affixes in Somali. The discussion will touch on related languages, so, as background, the diagram in Figure 1 shows the position of Somali in East Cushitic, a subgroup of Cushitic, itself one of the Afroasiatic subfamilies.

Somali, in common with other Cushitic languages, has a productive series of affixes which alter the semantic and argument structure of verbs to which they are attached. Thus we find several causative affixes, passive, autogenitive or ‘middle’, and inchoative affixes. These are described in Andrzejewski (1968), and Saeed (1987, in press). This paper grew out of an investigation of a group of

Figure 1. The Omo-Tana Subgroup within Eastern Cushitic
morphemes, associated with the verb, which have been variously called in English: prepositions, prepositional particles and locative particles; and in Italian, *preposizioni* (Panza 1974, Puglielli 1981a). As we will see, this terminology, though mnemonically convenient, is rather misleading. The original question at issue was whether these morphemes belong to the family of verbal derivational affixes or not; and if not, what their grammatical status might be.

These morphemes have been called prepositions because they assign to NPs a range of semantic or thematic roles similar to English prepositions; but in fact their grammatical status is quite different. In examining cross-linguistic correspondences to English sentences like (2), we are perhaps familiar with information about ‘movement from a source’ being marked in certain ways, such as (3a–3d).

(2) We came from Somalia.

(3) (a) included in the semantics of the verb;
(b) marked by case on the noun;
(c) marked by a preposition forming a unit with the nominal;
(d) marked by a postposition forming a unit with the nominal;
(e) marked by a morpheme satellite of the verb.

Somali seems to present us with an additional strategy (3e) as shown in (4–6).

(4) Soomaaliya ayaanu ka nimi
    Somalia FOC+we from came
    ‘we came from SOMALIA’

(5) waxaannu ka nimi Soomaaliya
    what+w we from came Somalia
    ‘it was SOMALIA we came from’

(6) Cali baa Amina u keenay Shukri
    Ali FOC Amina to brought Shukri
    ‘ALI brought Amina to Shukri’, or less likely, ‘ALI brought Shukri to Amina’

In (4) and (5) we see the morpheme *ka* ‘from’, remaining next to the verb regardless of the position of the NP *Soomaaliya*. In (5) the corresponding morpheme is *u* ‘to’. Constructions with these

morphemes have the three characteristics in (7), as we will see in later examples.

(7) (a) The morpheme remains in pre-verbal position regardless of the position of the NP it semantically ‘governs’;
(b) This governed NP occurs in the basic absolutive form (the default case used for objects, isolation, etc.), whichever morpheme occurs;
(c) When more than one non-subject NP is involved, these structures can give rise to ambiguity, which is resolved by contextual information.

In fact there is only one nominal adposition in Standard Somali: the preposition *ilaal* ‘towards, to’, borrowed from Arabic and which forms PPs with NPs in the usual way as in (8):

(8) [PP | ilaa | [NP Berbera]]  ‘towards Berbera’

In the rest of this paper we ignore this one preposition. The basic meanings of the preverbal morphemes are given in (9), and an idea of the semantic roles they assign to nominals is given by examples (10–18), where the morphemes are in bold type:

(9) *u* ‘to, for’
   
   *ku* ‘in, into, on, at, with (instrument)’

   *ka* ‘from, away from, out of’

   *la* ‘with (comitative)’

(10) bannaanka u saar!
    bannaann+ka  u saar
    out-of-doors + the to release
    ‘release it to the outdoors!’, ‘make it all public!’ (Mumin 1979: 162)

    *u*: role assigned = GOAL

(11) maxaad ii weydiinaysaa?
    maxay+baa  i+u  weydiinaysaa
    what+FOC+you me+for are asking
    ‘what are you asking me for?’, ‘why are you asking me?’

    (Mumin 1979: 62)

    *u*: role assigned = REASON
(12) anigu waxaan ku cayilay waa miyigii baan tegay
   I what+I from grew-fat(is) country + the FOC + I went
   "what I grew fat from is I went to the countryside" (Mumin
   1979: 78)
   ku: role assigned = CAUSE

(13) Baricade iyo walaashii waxay ku naseen dhulkii weyna
    B and sister + his what + they in rested land + the big
    'Baricade and his sister rested in the great country' (Cartan
    Xaange 1988: 21)
    ku: role assigned = LOCATIVE

(14) rajaynimaan ku weynaadey
    orphanhood + FOC + 1 in grew-up
    'I grew up in orphanhood', 'as an orphan' (Cartan Xaange
    1988: 15)
    ku: role assigned = LOCATIVE (metaphoric)

(15) war aynu caday ku rumayanno...
    hey us caday-plant with clean-teeth
    'hey, let's clean our teeth with caday-plant!' (Cartan Xaange
    1988: 55)
    ku: role assigned = INSTRUMENT

(16) haddaad iga muudsatid sunta maska...
    hadda + aad i + ka muudsatid sun + ta mas + ka
    if you me + from suck poison + the snake + the
    'If you suck out the poison of the snake from me . . .' (Cartan
    Xaange 1988: 20)
    ka: role assigned = SOURCE

(17) ka firso hadalladaad leedahay!
    about think talk + the + you say
    'think about the talk you are saying!', 'mind what you are
    saying!' (Mumin 1979: 66)
    ka: role assigned = THEME (or perhaps LOCATIVE in a
    metaphorical sense)

(18) war yaad la hadlaysaa?
    hey who + FOC + you with talking

"hey, who are you talking with?" (Cartan Xaange 1988: 64)
la: role assigned = COMITATIVE

As David Appleyard has reported (Appleyard 1990), close cognates of these preverbal morphemes are found in other languages of the Omo-Tana branch of East Cushitic (shown in Figure 1) and less commonly in other branches of Cushitic where case-endings and postpositions predominate. Preverbal morphemes exactly paralleling these Somali examples have been reported in Boni, Rendille, Elmolo and Dasenah. Bernd Heine in his survey of Eastern Omo-Tana (which he calls the ‘Sam’ languages: Heine 1978) left their status somewhat open:

(19) 'Proto-Sam must have had a set of particles placed immediately before the verb. The function of these particles is intermediate between nominal prepositions and semantic verbal extensions' (Heine 1978: 37–38).

In more recent work by Steve Pillinger on Rendille these morphemes have been identified as verbal derivational affixes: In Pillinger (1989: 244f.) verbs occurring with ka (cognate with Somali ka) and lee (cognate with Somali la) are identified as derived verbs. We can see this in (20) which is Galboran & Pillinger's dictionary entry for ka, labelling it a verbal derivational affix. Indeed the dictionary has separate lexical entries for verbs with and without these morphemes as shown in the selection of entries in (21):

(20) 'kā' (VPX); at, in, from, with, by
    [Locative/ablative/instrumental verb prefix. Also has the
    function of adding an extra object: i.e., makes an intransitive
    verb transitive or a transitive verb ditransitive; cf. riira "shout",
    but kāriira, "shout at (someone)" . . .'] (Galboran & Pillinger
    1990: 43)

(21) abda 'learn' kāabda 'learn from'
    jira 'be, exist' kājira 'be at (a place)'
    (Galboran & Pillinger 1990)

This approach would make these morphemes a Cushitic parallel to the applicative verbal affixes found in Bantu languages, e.g. Swahili (22):
(22) (a) ni-me-lima shamba kwa ajili ya Musa
I-PERF-cultivate plantation for sake of Musa
'I have cultivated the plantation for (the sake
of) Musa'
(b) ni-me-m-lim-ia Musa shamba
I-PERF-him-cultivate-APPL Musa plantation
'I have cultivated the plantation for Musa' (Comrie
1985:317)

In (22b) the applicative affix has made lima 'cultivate' into a so-
called 'prepositional' verb, which, interestingly for our later dis-
cussion, then requires an object agreement affix within the verb.

In this paper I will examine the status of these morphemes in
Somali. I will argue firstly that they are clitics and not affixes, and that
consequently when they occur with verbs we are not dealing with
lexically generated derived verbs. This leaves us with the problem of
terminology. While the term 'preposition' is a useful mnemonic for
speakers of English or Italian, the terms preposition, postposition,
and adposition normally imply constructions where the element
forms a syntactic unit with a nominal. This is clearly not the case with
these Somali preverbal clitics. On comparative and historical
grounds we might want to draw parallels with adpositions in other
Cushitic subgroups, perhaps even positing a historical migration of
adpositions to the verb in Omo-Tana, in which case a term like
adpositional clitic might be defended. Looking at the functions of
these clitics - marking comitative, instrumental, directional,
locative, etc. roles on arguments of the verb - we might prefer a term
like oblique case clitic, where 'case' refers to abstract or deep case in
the spirit of Fillmore's Case Grammar (e.g. Fillmore 1968), rather
than inflectional case. We might note Lyons's (1968:289-304)
observations that similar semantic functions are marked in various
languages by nominal adpositions, inflectional cases, and we might
now add, preverbal clitics. We seem to be dealing with different
grammaticalization strategies for similar semantic functions, and
therefore the choice of terminology will reflect an emphasis on either
the syntactic or semantic behaviour of the items. In this, the problem
presented is reminiscent of the identification of adverbials in many
languages. Since Somali does have nominal inflectional cases for
marking, e.g., subject versus nonsubject, and for genitive (see Saeed
1987:128 ff.), we will not, for the present, adopt the term case clitic,
but use instead the term adpositional clitic.

This first section concludes with a summary of the characteristics
of these adpositional clitics which distinguish them from both words
and affixes. Secondly, I will argue that these adpositional clitics
should be assigned independent lexical status and that the construc-
tions they occur in should be derived by syntactic rules. The final sec-
tion of the paper is concerned with the interaction of these clitics
with word order. While the discussion is of immediate relevance to
the description of Cushitic languages, it is hoped that the con-
cclusions will be of wider, especially typological, interest.

2. The Clitic Status of Adpositional Clitics

Although the definition of clitics is one of the more controversial
areas in syntactic theory, I will assume that for an element to be a
clitic it must satisfy the conditions in (23):

(23) (a) it must be a bound morpheme: for example, not able to
occur in isolation, as an answer to a question, etc.; and,
(b) at least optionally, be phonologically attached to another
element - a host word or other clitic; and,
(c) be distinguished by its grammatical behaviour from
morphological affixes.

These morphemes clearly satisfy (23a): they can only occur with a
verb and, unlike full words whose order is free, as we will see, they
must occur in a fixed position in a clitic cluster preceding the verb.
We can call the verb and its preceding clitic cluster the verbal piece;
and in (24a) the verbal piece, which needs a classifier or other
element to form a clause, has only one permissible order of its
elements, schematized in (24b), as (24c-24e) show:

(24) (a) . . . uu kugu soo noqday
    uu ku + ku soo noqday
    he you + to back came
    ' . . . he came back to you'
Strong evidence for these morphemes satisfying criterion (23b) comes from the collapsing or coalescence rules which apply to these elements causing the host and clitic to merge into a new phonological form. See for example (25) and (26) where adpositional clitics coalesce with absolutive pronoun clitics.

(25) wuu noo qaaday
    waa + uu na + u qaaday (na + u → noo)
    CLASS + he us + for took
    ‘he took it for us’

(26) ii sheeg!
    i + u sheeg (i + u → ii)
    me + to tell
    ‘tell (it) me’!

As described in detail in Andrzejewski (1960), adpositional clitics and pronominal clitics merge into preverbal clitic clusters which are sometimes quite complex, as in (27):

(27) ...
    rarka igala dhiga!
    rar + ja i + ka + la dhiga (i + ka + la → igala)
    loads + the me + from + with take-down
    ‘... take the loads down from (them) with me!’ (Andrzejewski 1960: 107)

To satisfy criterion (23c) we must show the divergence of these elements from morphological affixes, and the most likely contenders here are verbal derivational affixes. The clearest evidence of difference concerns attachment to the verbal root. Derivational affixes attach in a specific order to the root of the verb, relative to other affixes. As shown in (28), they occur closer to the root than inflectional affixes.4

(28) (a) [VERB ROOT + LEX + AGR + TNS]
    qaad + at + t + ay →
    take middle 3fs past
    qaadatay ‘(she) took for herself’

There are a number of phonological arguments for distinguishing adpositional clitics from affixes, and derivational affixes in particular. They all take the following form: lexical affixes in their position in (28) undergo various phonological rules which show them to form a phonological word with the root; these include tone assignment, assimilation, syllable simplification, and other rules. The clitics on the other hand are inert to all of these processes. I will not deal with this form of justification in any great detail here, for two reasons. The first, very practical reason, is that relatively little is yet known of the details of Somali phonological rules. This means that any arguments are likely to be based on shifting sands. The second reason is that it is not a priori clear that inertness is a characteristic which would disqualify a morpheme from affix status. We might remember Dorothy Siegel’s (1979) study of English affixes where she distinguished two classes of affix precisely on the basis that CLASS II affixes were relatively inert to phonological processes with the stem.

Nonetheless I might suggest one or two arguments of this type. One concerns the assignment of tone patterns to verbal roots. As Andrzejewski (1956) has shown, particular paradigms of verbs are marked by specific tone patterns, the choice of pattern cross-cutting with morphological classes. So, for example, in first and second conjugation verbs (to employ the most commonly used terminology, e.g. Bell 1953, Andrzejewski 1956, Saeed 1987 amongst others), the imperative forms have a HIGH-LOW pattern on the last two syllables, as in (29a). If two syllables are not available, we find the following realizations in (29b and c):

(29) (a) polysyllables:
    …CVCV
    H    L
    (where V may be short or long)
(b) long monosyllables:

\[
\text{CVV(C)}
\]
\[
\begin{array}{c}
\text{H} \\
\text{L}
\end{array}
\]

(realized as FALL)

\[
\text{CVC}
\]
\[
\begin{array}{c}
\text{H} \\
\text{L}
\end{array}
\]
i.e. a HIGH

(c) short monosyllables:

Examples of this can be seen in (30)

(30) engőjì! 'dry (it)'
hagāáji! 'straighten (it)'
órōd! 'run!'
jōoji! 'stop (it)'
kēen! 'bring (it)'
qāad! 'take (it)'
tāg! 'go!'
mār! 'pass!'

If grammatical rules add affixes inside this paradigm, the tone shape shifts to reflect the extra syllables, as shown in a single example in (31):

(31) kēen! [base form] 'bring (2sg) it!'
kēena! + [plural affix] 'bring (2pl) it!'
kēenin + [negative affix]:
ha kēenin! 'don’t (2sg) bring it!'
kēenina + [negative and plural affixes]:
ha keenina! 'don’t (2pl) bring it!'

It seems clear that the rules for applying tone patterns to verbal forms treat the affixes as part of the same word. By contrast the preverbal clitic cluster has no effect on this process, as shown in (32):

(32) ú kēen! 'bring (2sg) it to him!'
ingà qāad! 'take (2sg) it from me!'
kā tāg! 'go (2sg) away from it!'
kā bāx! 'get (2sg) out of it!'

It has to be said that without specific phonological analyses, this kind of evidence is not conclusive: one might in the end want to argue that there are different kinds of affixation at work here, only one type of which is ‘visible’, so to speak, to the rules which match tonal shapes to paradigms. Nonetheless, in the absence of such arguments these data do mark a significant difference between affixes proper and our proposed clitics. We might use just one more example of this type of data to show that one of the phonological rules of assimilation – the voicing of intervocalic stops within a word – treats affixes as part of a word with the verbal root (33 below), and adpositional clitics as part of a word with pronominal clitics (34), but each as separate, as the surviving voiceless stops in (35) show:

(33) \[\text{joog} + s + a\theta + \text{aan} \rightarrow \text{joogsadaan} \]
\[\text{+ CAUS + MIDDLE + AGR + INFL} \]
\[\text{‘(they) stop themselves’} \]

(34) \[\text{wuu} \quad \text{iga} \quad \text{qaaday (i + ka \rightarrow iga)} \]
\[\text{CLASS + he me + from took} \]
\[\text{‘he took it from me’} \]

(35) \[\text{wuu} \quad \text{kuu} \quad \text{keenay} \]
\[\text{CLASS + he you + for brought} \]
\[\text{‘he brought it for you’} \]

As mentioned earlier, similar arguments can be made using a variety of phonological processes, all of which do not ‘see’ adpositional clitics as part of the verbal word.

Clearer perhaps than this phonological evidence, and supporting it, are morpho-syntactic arguments based on evidence of continuity, or perhaps better, intrusion. Derivational affixes are always attached next to verbal roots ‘inside’ inflection, and permit no intrusive elements. Adpositional clitics on the other hand may be separated from the verb by a number of elements, most significantly certain sentence-level elements. We see examples of this in (36) to (39), where the intrusive element is in bold type:

(36) \[\text{waan} \quad \text{kuu} \quad \text{soo} \quad \text{celin} \quad \text{doonaa} \quad \text{carrabkaaga} \ldots \]
\[\text{CLASS + I you + to back return will tongue + your} \]
\[\text{‘I will give you back your tongue …’} \quad \text{(Cartan Xaange 1988: 79)} \]
(37) sonkori kuna jirto  
  sugar in + NEG be (NEG)  
  'sugar is not in it', 'there's no sugar in it'

(38) sii uu weysa uga qabto saca ina  
  way he calf + the in + from seized cow + the before  
  ayan uba soo dhowaan  
  she + NEG to + even towards neared (NEG)  
  'in such a way that he seized the calf from  
  the cow before she even got near it' (Cartan-Xaange 1988: 59)

(39) (a) wuu tegay gurigiisa dabadeedna wuu  
  CLASS + he went house + his then + and CLASS + he  
  ka soo noqday  
  from back came  
  'he went to his house and then came back from it'

(b) wuu tegay gurigiisa wuuna ka  
  CLASS + he went house + his CLASS + he + and from  
  soo noqday.  
  back came  
  'he went to his house and he came back from it'

(c) wuu tegay gurigiisa wuu kana  
  CLASS + he went house + his CLASS + he from + and  
  soo noqday  
  back came  
  'he went to his house and he came back from it.'

(d) *wuu tegay gurigiisa wuu ka soona noqday

(e) *wuu tegay gurigiisii wuu ka soo noqdayna

In (36) the adpositional clitic is separated from the verb by the  
deictic element soo 'towards the speaker'; in (37) by the main  
sentence negative word ma 'not', which must be distinguished from  
negative verbal inflection; in (38) by the clausal parenthetic element  
ba 'even'; and most tellingly in (39) by the sentence coordinator -na  
'and'. This links clauses but not NPs, and generally follows the quite  
common cross-linguistic tendency to take up a position after the first  
constituent of the second clause, or 'Wackernagel's position' as it is  
often called. The only variability we find, as shown in (39), concerns  
elements of the verbal piece, where speakers can identify a  
constituent boundary either after the cluster of CLASSIFIER + subject  
pronoun clitic, or after the adpositional clitic. The coordinator  
cannot be placed after the deictic element or the verb, as (39d,e)  
suggests. Needless to say, it cannot be placed between the verb root  
and any of its affixes. The point here of course is that the placement  
of -na seems sensitive to constituent boundaries and the possibility  
of placing it after the adpositional clitic is more evidence that the  
latter is not an affix of the verb.

One final argument for clitic rather than affix status for these  
morphemes derives from an observation made by several linguists,  
for example Zwicky and Pullum (1983), that clitics are less selective  
in their choice of hosts than affixes. Somali adpositional clitics  
contrast with verbal affixes in this regard: we have seen examples of  
them cliticizing onto absolutive pronoun clitics (examples 11, 16, 24  
etc.), other adpositional clitics (27), the negative word ma 'not' (37),  
and the coordinator -na, 'and' (39c). Moreover whereas affixes are  
very rigidly attached to verbal roots, we have adpositional clitics  
cliticizing onto hosts both leftwards (pronoun clitics) and rightwards  
(negative ma and coordinator na).

Perhaps here we can summarize in (40) the differences these  
adpositional clitics exhibit from both words and affixes:

(40) (a) **Un-wordlike behaviour**
  1. they are bound morphemes;
  2. they position in clitic clusters which have a fixed  
     sequence unlike Somali's free word order;
  3. these clusters are attracted to the verb;
  4. these morphemes cliticize onto hosts.

(b) **Un-affixlike behaviour**
  1. phonological rules treat clitic clusters as words, but not  
     sequences of clitic cluster plus verb;
  2. they may be separated from the verb by various  
     elements, some clausal;
  3. they do not exhibit high selectivity in hosts;
  4. their attachment is directionally variable.
On a different level, a point that was recently made to me in Nairobi by Somali speakers, when discussing this topic, was that since the adoption of an orthography in 1972, clitics have consistently been written separately from the verb, and that this is not an area where Somali writers generally show indecision in word division.

We can end this section by noting that compared to clitics cross-linguistically, these adpositional clitics are perhaps unusual in exhibiting high tone, often with accompanying stress, but we might note that this does not prevent coalescence with other elements, both high and low tone, e.g. (41):

(41) (a) ú 'in' + ká 'from' \(\rightarrow\) uga 'in-from'
sidée bi'ir uga baxay?
'In which way did he go out from it?’, ‘how did he get out from it?'

(41) (b) na 'us' + ká 'from' \(\rightarrow\) nagá 'from us'
wwu nagá qaadaay
'He took it from us'

3. The Lexical Status of Adpositional Clitics (ACs)

If we grant then that the evidence seems to support the identification of these morphemes as clitics in preverbal position, there arises the question of their lexical status. While allowing them to be clitics, it might still be argued that the clitic-verb combinations should be listed in the lexicon as forming phrasal semantic units, e.g. as in (42):

(42) qaad 'take'
 u . . qaad 'take for'
 ka . . qaad 'take from'
 la . . qaad 'take with'

I will argue against this, and for adpositional clitics being listed in the lexicon separately from verbs, i.e. that phrases as in (42) should be created by syntactic rules. There are a number of reasons for this. Firstly, through some AC-verb combinations do have idiosyncratic meanings as in (43), for the most part the ACs contribute consistent meanings to the AC-verb combinations, as in (44):

(43) la . . bax 'to elope'
   with leave
   ka . . celi 'to defend'
   from send-back

(44) ka . . tag 'to go from'
   ka . . yimi 'to come from'
   ka . . keen 'to bring from'
   ka . . qaad 'to take from', etc.

Moreover there are a number of constructions involving these clitics which can occur with any verb in the language. One such construction forms adverbial expressions, involving u 'in' assigning a MANNER role to an NP, as in (45).

(45) sidan u samee!
   way + this in do
   'do it this way!', 'do it thus!'

In fact manner adverbials are generally formed by using u to govern an NP, with or without use of si 'way, manner', as shown in (46).

(46) si dirran . . u 'warmly'
   way warm in
   si xun . . u 'badly'
   way bad in
   degdeg . . u 'quickly, hastily'
   haste in
   sahal . . u 'easily'
   ease in

In exactly parallel constructions, adverbials of direction are also regularly formed with u:

(47) sare . . u 'upwards'
   hoos . . u 'downwards'
   hor . . u 'forwards'
   dib . . u 'backwards'
   gees . . u 'sideways'
Since these adverbial constructions can be used with any verb, and since the clitic u can govern an infinite number of NPs, including relative clauses (as in example (48) below sidaad doonto ‘the manner which you want’), it seems reasonable to assume that the clitic has to be listed separately in the lexicon, and not together with either the verb or any listable nominals.

(48) sidaad doonto u samee!
     way + the + you wish in do
     ‘do it in the way you want’!, ‘do it how you want’!

A second, similar argument arises from an examination of comparative constructions, where another adpositional clitic ka is used, as in (49):

(49) agalkan agalkaas wuu ka weyn yahay.
     house + this house + that CLASS + it than big is
     ‘this house is bigger than that house’

Here the NP semantically governed by ka corresponds to the complement of ‘than’ in English comparatives (hence the gloss in (49)). Once again this is a maximally productive construction: any adjective can occur with this clitic in this construction; a few examples are in (50):

(50) ka . . . dheer ‘taller than . . .’
    ka . . . yar ‘smaller than . . .’
    ka . . . badan ‘more than . . .’
    ka . . . qabow ‘colder than . . .’

Once again the regularity of the process and the fact that it can occur with any adjective and any verb in the language suggests that the most economical approach is to list ka independently in the lexicon, and form these constructions by regular rules.

To conclude this brief section, then: adpositional clitics (with some exceptions) combine with verbs and nominals in a semantically predictable way, and furthermore, in adverbials and comparatives they participate in productive phrasal rules, which involve an infinite number of complex constructions, like relative clauses. This suggests that, on a fairly reasonable division of labour between the lexicon and syntactic rules, these clitics are best assigned independent lexical status, and the constructions in which they are involved are to be derived by syntactic rules.

4. Adpositional Clitics and Word Order

To recapitulate the discussion so far: these morphemes assign semantic or thematic roles to NPs but show up as clitics in preverbal clitic clusters. We could deal with their syntactic derivation by constructing verbal pieces (with their preverbal clusters) by a combination of constituent and linearization rules. But what might seem somewhat odd about this is that such an approach leaves us with no syntactic marking of the association between the adpositional clitic and the NP it semantically governs. This is odd, because, of course, in various grammatical theories it is expected that thematic role assignment will have some syntactic, preferably structural, basis. So in Government & Binding (or Principles & Parameters) Theory, for example, theta-roles have classically been assigned under a configurational relation of government (Chomsky 1986a, 1988).

For our Somali adpositional clitics one of the obvious ways to save this expectation might be to assume that an adpositional clitic and an NP do form a structural unit at some level and that the clitic is moved to the verb by a rule of adposition incorporation, in the spirit of rules proposed for other languages by Baker (1988). There might even be a cross-linguistic glimmer of hope for such an approach: Steve Pillinger has told me that the morpheme ka ‘from, with, etc.’ in Rendille which normally occurs as a preverbal particle as in (51a) may also show up as a nominal postposition, as shown in (51b) (both given with Pillinger’s transcription and glosses) which reflect different emphases, the details of which are not yet clear:

Rendille

(51) (a) wül inamá kájahe
     stick boy FOC with-hit
     ‘he hit the boy with a stick’

(b) wül ká inamá jahe
     stick with boy FOC hit
     ‘it was with a stick that he hit the boy’
But note that this option seems not to be true for other Rendille preverbal particles, nor to be true of any of the Somali clitics.

But before examining whether a GB-style movement rule is justified, perhaps by testing whether the postulated rule obeys the expected constraints like the Empty Category Principle or its derivative, the Head Movement Constraint (see Baker 1988, Ouhalla 1989 for discussion of these in relation to incorporation), we might save some time and effort by reviewing in general the role of government in case and theta-role assignment in Somali. As I argued in Saeed (1984) there are good reasons for assigning Somali sentences a flat structure without a VP constituent, i.e. something like (52) for a simple sentence with a transitive verb, where again VPC is the verb and its accompanying template of clitics.

(52) S → NP, NP, VPC

Any linearization rules which might apply to the constituents in (52) will be driven not by grammatical relations, but by discourse-based rules of information structure, or ‘functional-sentence perspective’ as Hans-Jürgen Sasse has shown for the closely related language Boni (Sasse 1981). To take a Somali example paralleling rule (52) and with no NP focus (for now), we find the possibilities in (53) for the sentence ‘the man reached his house’, each belonging to a different conversational context.

(53) ‘the man reached his house’
(a) ninkii áqalkiisi wíiu gaadhay
   man+the house+his CLASS+he reached
   SUBJ OBJ V
(b) áqalkiisií níinkii wíiu gaadhay
   OBJ SUBJ V
(c) níinkii wíiu gaadhay áqalkiisií
   SUBJ V OBJ
(d) áqalkiisií wíiu gaadhay níinkii
   OBJ V SUBJ
(e) wíiu gaadhay níinkii áqalkiisií
   V SUBJ OBJ

(f) wíiu gaadhay áqalkiisií níinkii
   V OBJ SUBJ

If an NP is focused by a focus particle baa or ayaa, then there is one word order constraint which operates, namely, the focused NP must occur to the left of the verb. This is what I called ‘Focus Fronting’ in my study of focus and topic (Saeed 1984), and is one of the reasons that Somali has been called a verb-final language. But the relative order of NPs is still quite free and there is no evidence for a VP. If we, for example, focus the object in (53) we can form the contextually distinct sentence in (54):

(54) (a) níinkii áqalkiisií bííu gaadhay
   man+the house+the FOC+he reached
   SUBJ OBJ V
   ‘the man reached HIS HOUSE’
(b) áqalkiisií bííu gaadhay níinkii
   OBJ SUBJ V

Note that the grammatical relations of subject and object are marked by the NP’s morphological shape (in this case tonally marked) and by agreement with the verb, but not by their syntactic configuration.

So the problem of explaining how case and thematic roles are assigned to NPs by a discontinuous adpositional clitic is actually the same task as explaining how case and thematic roles are assigned to subject and object NPs when they stand in no fixed configurational relationship to the verb. The fact is that in Somali and closely related languages, as far as the syntax is concerned, NPs ‘float’ free around the verbal complex, their order being determined by conversational principles. This means that the governing of NPs from within the verbal complex by an adpositional clitic or verb appears to have no structural conditions; in other words, it seems that governors like V and what would be P in other languages do not form constituents with their governed NPs.

One conclusion from this might be that such dependencies have to be represented at a separate level. Whether this level might best be seen as a separate syntactic level of functional structure as in
Lexical-Functional Grammar (Bresnan 1982), or a semantic representation which would recompose the units, as Hale suggested for Warlpiri in Hale (1983), is a question that I shall not try to settle here.

Instead I would like to sketch in a somewhat different analysis which may go some way to explaining the freedom of NP ordering. This is that the free floating of NPs around the verbal piece is licensed by the verbal piece containing in clitic form the argument structure needed to satisfy the verb.

Firstly note that the verbal piece with a classifier can be a whole sentence, if all the NPs are known or deictic, as in (55):

(55) way naga kaa qaadeen
     CLASS+ they us +from you took
     'they took you away from us'

Now, tracing whether clitics are always there to satisfy the verb’s argument structure is complicated by the fact, long recognized in Somali grammars, that the absolutive pronoun clitics include a gap in their paradigm. As shown in (56) the 3rd person absolutive form of these clitic pronouns is realised as zero; i.e. the omission of an absolutive pronoun clitic is interpreted as a third person argument, the choice being contextually determined. In other words, any unfilled position in the argument structure of a verb or adpositional clitic will be interpreted as this empty pronoun. In (56) the lack of an overt object argument for the verb *arag* ‘see’, will result in one of the interpretations given as glosses:

(56) waan arkay
     CLASS+I (3pro) saw
     'I saw him', 'I saw her', 'I saw it', 'I saw them'

If we have a sentence lacking an overt argument for a case clitic, the result will be as in (57):

(57) waan u siiyey lacagii
     CLASS+I (3pro) to gave money + the
     'I gave the money to him/her/them'

This means that, if we are to see whether the floating NPs always have a clitic copy in the verbal complex, we will need to use examples with non-3rd person arguments governed by verbs and case clitics. The tests will presumably have to show:

(a) whether doubling of arguments can occur, i.e. whether there can be both a clitic in the verbal complex and a coreferential NP ‘floating free’, so to speak; and,

(b) which of the two satisfies the argument structure requirements of the governor, say, for now, the verb; that is, which of the two *has* to occur.

Examples (58)-(60) use the independent pronouns, which are syntactically and morphologically full nouns, as the second non-clitic argument. So, sentences (58a–d) use a second person singular verbal object, (59a–d) a second person plural, and (60a–d) a first person plural. The (a) and (b) examples show part of the full NPs’ ‘floating’ behaviour around the verbal complex (with other permutations omitted for brevity). The (c) examples show that just the clitic pronoun satisfies the verb’s argument structure, while the (d) examples show that just the full NP cannot.

(58) (a) askartu adiga way ku raadinayeen
     police + the you(sg) CLASS+ they you(sg) were-seeking
     'the police were looking for you (sg)'

     (b) askartu way ku raadinayeen adiga
     police + the CLASS+ they you(sg) were-seeking you(sg)
     'the police were looking for you (sg)'

     (c) askartu way ku raadinayeen
     police + the CLASS+ they you(sg) were-seeking
     'the police were looking for you (sg)'

     (d) *askartu adiga way raadinayeen
     police + the you(sg) CLASS+ they were-seeking
     'the police were looking for you (sg)'

(59) (a) askartu idinka way idin raadinayeen
     police + the you(pl) CLASS+ they you(pl) were-seeking
     'the police were looking for you (pl)'

     (b) askartu way idin raadinayeen idinka

     (c) askartu way idin raadinayeen

     (d) *askartu idinka way raadinayeen
(60) (a) askartu annaga way na raadinayaan
    police + the us CLASS + they us were-seeking
    'the police were looking for us'
(b) askartu way na raadinayaan annaga
(c) askartu way na raadinayaan
(d) *askartu annaga way raadinayaan

The (d) examples above are ungrammatical because of the lack of an
object pronoun clitic inside the verbal piece; we can speculate that
the gap in the verbal piece causes the default 3rd person inter-
pretation, which then clashes with the non-3rd person independent
NP. If you like, the (d) examples are interpreted as having one too
many arguments for the verb's required argument structure, some-
thing we could parallel in English by (61):

(61) *the police were looking for him you

Exactly parallel examples occur with arguments governed by adpo-
sitional clitics:

(62) (a) askartu adiga way kaa qaadeen
    askar + tu adiga waa + ay ku + ka qaadeen
    police + the you (sg) CLASS + they you + from took
    'the police took it away from you (sg)'
(b) *askartu adiga way ka qaadeen
    'the police took it away from you (sg)'

In (62b) the adpositional clitic *ka, 'from', is not interpreted as
governing the independent pronoun *adiga, 'you'; the adpositional
clitic has to have a clitic pronoun in the clitic cluster. In other words
it seems that it is the clitic within the verbal piece and not the NP
outside it which satisfies the argument structure requirements of
verbs and adpositional clitics.

The one exception to this is focused NPs. These cannot occur in
clitic form in the verbal cluster as (63) and (64) show:

(63) (a) Cali baa ku raadinaya
    Ali  FOC you seek(pres)
    'ALI is looking for you', 'it's ALI who's looking for you'

(b) *Cali buu ku raadinaya
    Ali  FOC + he you (sg) seek(pres)
    'ALI is looking for you (sg)'

(64) (a) adigay raadinayaan
    adiga + baa + ay raadinayaan
    you (sg) + FOC + they seek(pres)
    'they are looking for YOU', 'it is YOU they are looking for'
(b) *adigay ku raadinayaan
    adiga + baa + ay ku raadinayaan
    you (sg) + FOC + they you seek(pres)
    'they are looking for YOU'

These examples show that focused NPs cannot occur with corefer-
ential clitics in double-argument constructions. Sentence (63) shows
this with a focused subject, and (64) with a focused object. This
seems pragmatically plausible given that clitic pronouns are lexically
underspecified and can participate in an argument structure only
because of the recoverability of their referents. This is shown by their
association with omissible, known NPs and deictic NPs. Focused
NPs, on the other hand, must obligatorily be overtly present in the
predication, as we would expect: NPs are given focus to be made
prominent in the discourse, signalling amongst other things, new
information and contrastive emphasis. They are thus maximally
resistant to omission, deletion, or representation by clitics. We can
describe the facts of (63) and (64) simply by stating that since
focused NPs have to be present as part of the predication, there is no
available position in argument structure for corresponding clitic
pronouns. Thus this behaviour of focused NPs indirectly supports
our analysis of doubled NP arguments as being external to the main
predication, while clitic pronouns are arguments within it.

This requirement for focused NPs to be overtly present and not be
represented by clitics perhaps presents us with a natural explana-
tion of why it is just here that we find a word order constraint on NPs, i.e.
that focused NPs must occur to the left of the verb, whereas there are
no constraints on the order of NPs represented by clitics in the
verbal piece.
If this analysis is correct then we can say that the governing
categories of verb and adpositional clitic in Somali have their
valency requirements satisfied, either: (a) by clitics in the verbal
piece; or (b) by a focused NP. In both cases there are, contrary to our
first impression, structural constraints at the level of syntax which
make government subject to locality constraints: a very strict one
with adpositional clitics, a looser one with focused NPs. In both
instances, the governed elements have to occur to the left of their
governor. Other NPs are completely free of structural constraints as
long as they are coreferential with clitics in the verbal piece. Given
the information status of these non-focus NPs, i.e. usually given or
known, it seems reasonable to see a parallel between these structures
and topic structures, i.e. to suggest something like (65) for the
government of a non-focus argument by an adpositional clitic

\[ \text{government} \]

\[ \text{topic relation} \]

\[ (65) \text{[VPC} \ldots \text{PRO} \ldots \text{AC} \ldots \text{V}] \ldots \text{NP} \]

We might note that the coreferentiality shown in (65) between
pronominal clitic and floating NP puts this relationship at one end of
the range for topic structures in Somali. As described in Saeed
(1984) the relationship between the topic and the predication
proper may range from coreferentiality with one of the predication’s
arguments, through set-subset relations, to purely pragmatic
associations.

One piece of support for this analysis comes from agreement. My
analysis basically claims that government by a verb or adpositional
clitic in Somali is, with nonfocus NPs, a local relation onto clitics in
the verbal piece, to which free-floating NPs are in a topic-like
relation. Now Somali shows subject-verb agreement in features of
number and gender. The question thus arises: is this agreement also a
local relation between pronominal clitic and verb, assigned under
government, as might be suggested by this analysis?

Some support for a positive answer comes from a consideration of
focused subjects. As we have seen, these focused NPs cannot be
represented by clitics in the VPC. So if agreement with non-focus

I argued in Saeed (1984) that this reduced agreement is in fact no
agreement. I shall not go into details here, but basically what we find
in reduced agreement is the base or simplest form, 3rd singular
masculine, except for surviving 3ps and 1pl agreements; these two
forms are often triggered semantically in various contexts where the
syntax predicts other masculine singular agreement, for example by
nonlocal NPs outside the clause, by pragmatically understood entities
overriding grammatical agreement, by semantically feminine
nouns in apposition with grammatically masculine nouns with lesser
lexical content, and so on (see Saeed 1984:90f. where this
phenomenon is called fuzzy agreement). Whatever the details it
seems clear that where no clitic subject pronouns occur in the verbal
piece, normal subject-verb agreement breaks down, suggesting that
agreement is essentially a local relation within the verbal piece.

It is interesting to note that the analysis I am outlining here falls in
quite well with suggestions made by Doug Biber about the historical
development of these morphemes (Biber 1984). He suggests that the
synchronic situation in Omo-Tana languages is a result of two
separate historical processes: firstly, the verbal incorporation of
pronouns which also moved combinations of pronoun + postposition towards the verb; and subsequently, the development of zero anaphora, or empty third person pronominal clitics. The present analysis might even explain one of the problems he reports in his survey: why it is that coalesced forms of pronoun and adpositional clitics sometimes mark government of the pronoun by the adpositional clitic but other times, government of an unrelated NP, i.e. the contrast or mismatch between (67) and (68) (adapted from Biber 1984:56):

(67) \textit{wuu kaa yimid}  \\
\textit{waa + uu }  \textit{ku + ka yimid}  \\
\textit{CLASS + he you +from came}  \\
\hline
\text{‘he came from you’}

(68) \textit{maxay arladini kaa keenay?}  \\
\textit{maxay + baa arlo + diini ku + ka keenay}  \\
\textit{what +FOC land +your you +from brought}  \\
\hline
\text{‘what has brought you (here) from your land?’}

My explanation would be that there is no mismatch and that (67) and (68) are in fact parallel structures. I would argue that (68) should in fact be described as in (69), where the adpositional clitic governs an empty pronoun just as it does an overt pronoun in (67). The external NP is \textit{arladini} ‘your land’ which, as (70) shows, enjoys the usual positional freedom of external NPs:

(69) \textit{maxaa arladini kaa keenay?}  \\
\textit{maxay + baa arlo + diini ku + [e] + ka keenay}  \\
\textit{what +FOC land +your you + 3pro + from brought}  \\
\hline
\text{‘what has brought you (here) from your land?’}

(70) (a) \textit{arladini maxaa kaa keenay?}  \\
(b) \textit{maxaa kaa keenay arladini?}

5. Conclusion

In examining the status of these adpositional clitics we have been led to try to account for the freedom of NP order in Somali sentences. Let me conclude by recapitulating the basic features of the analysis I am now proposing.

First, I have tried to show that these adpositional morphemes are in fact clitics and not some kind of verbal affix, and that they participate in preverbal clitic clusters with other types of clitics. Secondly, I have argued that the clitics have to be assigned independent lexical status and that their constructions with verbs and adjectives have to be derived by syntactic rules. Turning to syntax, we have seen that the presence of pronominal and adpositional clitics in the verbal piece seems to license the extreme positional freedom of non-focused NPs. Focused NPs however cannot be represented by pronominal clitics and are more restricted positionally. I have suggested that, for non-focused arguments, it is the pronominal clitics in the verbal piece which satisfy the valency requirements of adpositional clitics and verbs, with optional ‘doubled’ free-floating NPs being in a topic relation with the clitics.

From a comparative point of view, we can see an interesting parallel between this type of verbal piece with its arguments encoded as clitics and the verb in Bantu, for example, where parallel elements – pronouns and an applied prepositional morpheme – occur as affixes in the verb, as shown in example (22b) earlier. It is tempting to speculate whether Somali and related Omo-Tana languages are evolving towards a Bantu-like verb. Nichols (1986) presents a language typology based on morphological marking; i.e. on whether dependency is morphologically marked on the head or on the dependent(s), or both. In her terminology, Somali, with its verbal clitic clusters encoding the whole predicate-argument structure, seems to employ head marking at the clause level. Interestingly, again as she claims for Bantu, Somali has mostly dependent marking at the phrasal level; for example, determiners are marked for gender and number agreement with their head noun, and adjectives agree in number with head nouns. Thus Somali, in this typology, seems to be a split marking language, the category that Nichols (1986) puts Bantu languages into. This raises the interesting question of whether
the head marking of argument structure on the verb in Somali, and its reflexes in the Omo-Tana subgroup of Eastern Cushitic, is a feature of areal interaction.

Centre for Language and Communication Studies, Arts Building, Trinity College, University of Dublin, Dublin 2, Ireland.

Notes

1. I am grateful to Axmed Xasan 'Cunayyee' and Maxamed Gees for discussions on the language material presented in this paper. Earlier versions of the paper were presented at the 22nd Annual Conference on African Linguistics, University of Nairobi, July 15–19, 1991 and to the Philological Society, 15 November 1991. I am grateful to those present at these meetings for their suggestions and comments, in particular, Steve Fillinger, Martin Mous and Ronnie Sim in Nairobi; Dick Hayward and Wyn Chiao in London. I am also grateful to Theodora Bynon and the anonymous Transactions referees for their suggestions. Research for the paper was supported by the University of Dublin's Arts and Social Sciences Benefactions Fund.

Abbreviations: FOC, CLASS, NEG = focus marker, classifier, negative word; 1, 2, 3 = first, second, third persons; sg, pl = singular, plural; m, f = masculine, feminine; ABS, GEN, SUB = absolutive, genitive, subject cases; PERF, APPL, CAUS, MIDDLE = perfective, applicative, causative, middle (autobenefactive) affixes; LEX, AGR, INF, FFL = derivational, agreement, tense/aspect affixes; pro = pronominal.

2. Indeed, Theodora Bynon has made the suggestion to me that a label for these elements might be adverbial clitics.

3. This schema does not include all possible elements in the preverbal clitic cluster. For example, we omit here discussion of the effect on the linear order of a second or third pronominal clitic, as in (55) below.

4. Here, for clarity, we deal only with the majority so-called 'weak' verbs in Somali; ignoring the four 'strong' verbs and the verb 'ababa to be, which have a different morphological structure and generally do not support derivational affixes (see Saed 1987:98–108).

5. Tone is not normally marked in the Somali orthography, and it is only marked here when important to the argument; the usual conventions are followed, i.e.: high tone is marked with an acute accent: á falling tone is marked by a grave accent: ã low tone is marked by the absence of accent: a

Note that tone marks are only placed on the first letter corresponding to long vowel or diphthong, e.g., ãa, or ùu.

6. Indeed in two cases the high tone of these adpositional clitics prevents homonymy with pronominal clitics:

kú in’ vs. ku ‘you sg, [abs. case]’
lá ‘with’ la ‘one, someone’

where the latter is the impersonal subject pronoun.

7. Nichols’s use of head differs from its use in some other current syntactic theories: for example in Nichols (1986), the verb is seen as the head of a clause and all the nominal arguments, including the subject, are dependent on the verb.

8. For NPs of possession, one of Nichols’s criterial cases, Somali has both a dependent-marking and a head-marking construction, as shown in (a) and (b) respectively:

(a) giriga Cali house + the Ali [Gen]
 ‘Ali’s house’ (lit. ‘the house of Ali’)
(b) Cali girgitisa Ali + his house [ABS]
 ‘Ali’s house’ (lit. ‘Ali his house’)

The choice between the structures seems to be dependent on discourse-based notions of topic, or viewpoint. The other doubtful example, from this perspective, is the marking of subject case on nominals, which seems to be a matter of dependent marking at the clause level.

References


Bresnan, Joan and Mchombo, Sam A., 1987. ‘Topic, pronoun and agreement in


Hale, K., 1983, 'Warlpiri and the grammar of non-configurational languages', Natural Language and Linguistic Theory 1, 5–47.


Tables, charts and diagrams (though not simple rules, examples or formulae) must each appear on a separate sheet, and must be labelled 'Fig. 2', 'Table 4' or the like as appropriate. References in the text to the table or diagram should be made by citing this label.

Phonetic or phonemic transcriptions should make use of the symbols and conventions of the International Phonetic Alphabet. Any departures from IPA practice should be made explicit, and justified.

References in the text should be of the following forms: ‘... applies also to nouns (see Matthews 1981:231–233)’; ‘Rigter & Beukema (1985:116) distinguish between...’.

Bibliographical references at the end should conform to the following models. (Mention one given name for each author, if you have the information; this aids readers following up your references in bibliographies, library catalogues, etc.) Titles of periodicals should be given in full, or abbreviated as in Bibliographie Linguistique. Give place of publication and publisher for books.


Copyright
All rights reserved. Apart from fair dealing for the purposes of research or private study, or criticism or review, as permitted under the Copyright, Designs and Patents Act 1988, this publication may not be reproduced, stored or transmitted in any form or by any means without the prior permission in writing of the publisher, or in the case of reprographic reproduction in accordance with the terms of licenses issued by the Copyright Licensing Agency or the Copyright Clearance Center.