

Between zero and nothing

Transitivity and noun incorporation in Somali

Mauro Tosco

Università di Napoli “L’Orientale”

The article analyzes the expression of understood objects in Somali. There is no object pronoun of 3rd person in Somali; this gap is usually interpreted as a “full \emptyset ”, which saturates the valency of a transitive verb and forces a reading with an anaphoric object. The article shows that this is empirically incorrect: in certain configurations, Somali transitive verbs admit either an anaphoric or a generic reading even in the absence of either an object NP or a non-null pronoun. In order to ensure a generic-object reading, Somali has further recourse to noun incorporation. The article explores the productivity of this strategy as a detopicalizing, backgrounding mechanism and argues that the overall generality of the anaphoric reading is the result of the obligatoriness of focus marking.

Introduction¹

As the late Robert Hetzron once wrote (Hetzron 1989), Somali is without doubt a language with many typological “peculiarities”. Some of them are real, while others are probably more the result of specific theoretical choices on the linguists’ part than inherent facts of the language. This article explores what is, in the present writer’s view, an example of the latter, and will advance a few typological considerations on the expression of “understood objects”.

In Somali, as in many other languages, there is no object pronoun of the 3rd person; according to the current analysis (as reflected, among others, in Puglielli [1981], Svolacchia, Mereu and Puglielli [1995], Saeed [1987, 1993, 1994, 1999], etc.), the object pronoun of the 3rd person is a “full \emptyset ” (rather than a mere gap in the paradigm), which can receive only an anaphoric, definite interpretation, while a reading with a generic object or as an intransitive verb is excluded:

- (1) *waan cunay*^{2,3}
 DECL:I eat:PAST:1s
 ‘I ate it’; *‘I ate’, *‘I ate something’

In other words, whenever a transitive verb is used without an object NP or a SAP (Speech Act Participant) object pronoun, a reading with a generic object is claimed to be barred, and a definite 3rd person object is always understood. In order to ensure an intransitive (or generic object) reading of a transitive verb, a generic noun such as *wax* ‘thing’ must be used:

- (2) *wax baan cunay*
 thing FOC:I eat:PAST:1s
 ‘I ate something’, ‘I ate’

This analysis has momentous consequences: first and foremost, the valency of a Somali verb is always fully saturated, and both the subject and the object (as well as, with ditransitive verbs, the indirect object) of the verb are always encoded. Upon the same “reality of \emptyset ” rests at least another tenet of a much-favored (and mainly couched in GB or minimalist terms) analysis of Somali: the non-argumental status of all non-focused NPs and their coreferentiality with a \emptyset pronoun within the sentence. Somali becomes therefore a non-configurational language, or, better, a discourse configurational one (which entails its being **argument** non-configurational; cf. Nordlinger [1998: 44]).

Sadly, it will be shown that this view rests upon empirically incorrect bases and that, in an admittedly limited number of configurations, the bare absence of an object is enough to ensure a generic or intransitive reading of a transitive verb.

It will further be shown that Somali may use another strategy in order to express a generic-object reading of a transitive verb, namely noun incorporation, in which an argument of the verb is attached (phonologically and morpho-syntactically) to the immediate left of a transitive verbal form. Noun incorporation is mostly used in Somali with patients of transitive verbs, yielding the result of filling the object valency of the verb and detransitivizing it. This productive incorporation is in principle (although not always in fact!) distinguished from the use of incorporation as a lexical device, in which noun incorporation does not affect the valency of the verb.

1. The traditional view

In any Somali main non-negative declarative sentence one and only one NP must be marked with a postnominal element (either *baa* or *ayaa*), usually called a “focus marker”. Apart from clefts, the only possibility to avoid the presence of a focus marker is to use the element *waa*, which has been analyzed either as a Verbal Focus Marker (e.g., Puglielli 1981), or as a Declarative Sentence Classifier (DECL; Saeed 1984, 1999). Non-focused NPs (i.e., those which are not followed by either *baa* or *ayaa*) are considered to be topics, and their position within the sentence is mostly pragmatically-conditioned.

The terms “focus” and “topic” are used in Somali linguistics somewhat idiosyncratically, and the focus markers perform actually a wide variety of functions, not all of them corresponding to what is normally called “focus” in linguistics. As expected, they follow the element which is introduced as new information or is given contrastive emphasis. In addition (and this is especially true when one leaves the question-and-answer framework upon which much syntactic work has been based and looks at actual Somali texts), the focus markers may be used in order to signal a switch in the topic of discourse (and in this case they will mark an element which is, pragmatically speaking, given). Conversely, inthetic (presentational) sentences the subject gets focus (although all the elements are, pragmatically speaking, new). Finally, the focus markers are often used after an adverbial NP as “action-enhancing” devices whenever the attention is upon the development of the action. Against all this, “topic” as used in Somali linguistics is to be understood very loosely as covering whatever NP is not salient enough to attract focus: while it is normally a given element, it can also be new (e.g., a non-subject NP in presentational sentences).

All the grammatical information relevant to the interpretation of any given Somali sentence is found in preverbal position, within what has been variously labelled Verbal Piece, Verbal Group or Verbal Complex (VC), a strictly ordered structure made up of various pronominal clitics, adpositions, a few adverbs, and, finally, the verbal form. The general structure of the preverbal part of the VC is shown in Table 1.

The only position from which an NP is excluded is within the VC. As a consequence of the free word order of NPs (even postverbal NPs may occur in what is otherwise a bona fide verb-final language) and of the structure of the VC, discontinuity is another major feature of Somali sentences, because adpositions, occurring as they do within the VC, are not found next to the NP of which they are the heads. This is explained (e.g., Saeed 1994) on the basis of

Table 1. The structure of the preverbal part of the Somali Verbal Complex

Subject clitic pronoun (max. 1)	Object pronoun/1 (max. 1)	Adposition (max. 2)	Adverbial/1 (max. 1)	Adverbial/2 (max. 1)	Object pronoun/2 (max. 1)
-aan 'I'	<i>i</i> 'me'	<i>u</i>	<i>soo</i> 'here'	<i>wada</i>	<i>kay</i> 'me'
-aad 'you' (S)	<i>ku</i> 'you' (S)	<i>ku</i>	<i>sii</i> 'there'	'together'	<i>kaa</i> 'you' (S)
-uu 'he'	∅	<i>ka</i>		<i>kala</i> 'apart'	∅
-ay 'she'	∅	<i>la</i>			∅
-aanmu (-aan) 'we' (EXCL)	<i>na</i> 'us' (EXCL)				<i>kayo</i> 'us' (EXCL)
-aynu 'we' (INCL)	<i>ina</i> 'us' (INCL)				<i>keen</i> 'us' (INCL)
-aynu 'you' (P)	<i>idin</i> 'you' (P)				<i>kiin</i> 'you' (P)
-ay 'they'	∅				∅
<i>la</i> IMP	<i>is</i> REFL				

the non-argumental status of the NPS: being satellites, the NPs must be co-referenced in the VC through pronominal clitics, which fully satisfy the argument structure of the verb and the adpositions. Since there is a 3rd person object pronoun ∅, this is a handy explanation. As the non-focused NPs are co-referenced in the VC, discontinuity of constituents may therefore be regarded as a surface phenomenon; (3) below is given the analysis shown in (3'), something like 'Cali, the newspaper, the girl, he gave it to her' (obviously, the respective order of ∅_i and ∅_k is purely conventional):

- (3) *wargeyskii Cali inantii wuu siiyay*^{4,5}
 newspaper:M:ANPH Cali:SUBJ girl:F:ANPH DECL:he give:PAST:3M
 'Cali gave the newspaper to the girl' (Saeed 1994:60)

- (3') *wargeyskii_i Cali_i inantii_k wuu_i ∅_i ∅_k siiyay*

Finally, the only NP which is considered argumental and cannot be co-referenced in the VC is the NP under focus, as shown by the grammaticality of (4) vs. the unacceptability of (4'):

- (4) *adiga ayay garteen*
 you.S FOC:they recognize:PAST:3P
 'they recognized you' (Saeed 1994:65)

- (4') **adiga ayay ku garteen*
 you.S FOC:they you.S.O recognize:PAST:3P

2. Generic-object uses of transitive verbs

So far for the standard analysis. But is it really true that the valency of a Somali transitive verb is always saturated — i.e., that a 3rd person \emptyset object pronoun is part of the paradigm? Consider the following sentence:

- (5) *haddaad tagto, is jir!*
 if:you.s go:DEP:2S REFL stay:IMPV:S
 ‘if you go, take care!’ (Saeed 1987:243)

tag ‘to go’ is a transitive verb — at least when used with an inanimate noun as goal; cf.:

- (6) *Xamar tag*
 Mogadishu go:IMPV:S
 ‘go to Mogadishu!’

With an animate goal the adposition *u* is used:

- (7) *Axmed u tag*
 Axmed to go:IMPV:S
 ‘go to Axmed!’

Of course, a translation of (5) as ‘if you go there,...’ or a literal ‘if you go to it...’ is possible; this is what is generally done when we replace *tag* ‘to go’ with, e.g., *keen* ‘to take’:

- (8) *haddaad keento, is jir!*
 if:you.s take:DEP:2S REFL stay:IMPV:S
 ‘if you take it, take care!’

Both sentences are analyzed as:

- (5') *haddaad \emptyset tagto, is jir!*



and

- (8') *haddaad \emptyset keento, is jir!*



The problem is that such an analysis is scarcely possible for the following sentence, which may be uttered when taking one's leave, and which generally receives an intransitive reading:

- (9) *waan tagayaa*
 DECL:I GO:PROG:PRES:1S
 'I am going'

Note that in (9) the verb is marked for the progressive aspect by the infix *-ay-*. Confirming the general validity of a "Transitivity scale" along the lines proposed by Hopper and Thompson (1980), we find also in Somali that the telicity or non-telicity of the action influences the transitivity of the verb.

A generic object reading is particularly possible when the verbal form is in the General Present paradigm, which has a habitual reading and is inherently non-punctual; again, a non-punctual activity is lower in transitivity than a punctual one (cf. Hopper and Thompson's [1980] "Transitivity Scale"). This may be exemplified by the following sentence from a literary text, in which the object of the three transitive verbs is generic and left unexpressed:

- (10) *waa shaqadaada in aad dishaa, dhacdaa, kufsataa*⁶
 DECL work:F:your.s that you.s kill:PRES:2s plunder:PRES:2s rape:PRES:2s
 'your job is to kill, plunder and rape' (Faysal 2000:91)

Even for *cun* 'to eat', which in (1) above received a transitive reading, an intransitive reading is the only possible in such "general statements" as:

- (11) *iska jir, eygani waa cunaa/qaniinaa*
 REFL:from stay:IMPV:S dog:M:this:SUBJ DECL eat:PRES:3M/bite:PRES:3M
 'take care: the dog bites!'

(11) could also be expressed as (11'), in which the Subject Clitic (3rd Singular Masculine *-uu*) is attached to and merges with the Declarative Classifier *waa*:⁷

- (11') *iska jir, eygani wuu cunaa /qaniinaa*
 REFL:from stay:IMPV.S dog:M:this:SUBJ DECL:3M eat:PRES:3M/bite:PRES:3M

On the other hand, a generic-object reading is not limited to non-past sentences; consider (12), taken from Saeed's (1983) grammar, in which evidence for a generic-object reading comes from apparent object gapping in coordination:

- (12) *wuu cunay oo cabay*
 DECL:he eat:PAST:3M and drink:PAST:3M
 'he ate and drank' (Saeed 1987:250)

If an anaphoric object reading were strictly enforced upon every transitive verb, (12) would entail not only the identity of the subject of the two coordinated clauses, but also the identity of the objects — which is obviously not the case. (12) is nevertheless fully acceptable, but only with a generic object reading.

Finally, the same opposition between an anaphoric object and a generic object reading may be expressed through the use of noun incorporation, as in the following alternative to (11):

- (13) *iska jir, eygani waa dad-cunaa*
 REFL:from stay:IMPV:S dog:M:this:SUBJ DECL people.eat:PRES:3M
 ‘take care: the dog bites!’

(13) shows noun incorporation in the verb. The next sections will explore the extent and use of this strategy in Somali.

3. Noun incorporation

Incorporation in Somali has received a certain amount of attention as a lexical device: Caney (1984), in his classical study on the modernization of Somali vocabulary, provides many examples of noun-verb and deverbalized noun-noun compounds. What has not been sufficiently investigated is the productivity of noun incorporation and its syntactic consequences. A perusal of any good Somali dictionary cannot reveal the extension of the phenomenon: per se, a dictionary deals with lexemes, and, in the case of compounds, with lexicalizations (cf. Sasse 2002 for the complex problem of what it means to write a dictionary of a massively incorporating language — in his case, Cayuga). Thus one finds in the *Somali-English Dictionary* (Zorc and Osman 1993), and to a lesser extent also in the slightly older *Dizionario Somalo-Italiano* (Agostini, Puglielli *et al.* 1985), a good number of lexicalized noun-verb compounds, especially if the resulting meaning is scarcely or not fully predictable from the meaning of their elements. They range from, e.g., *hantigeli* ‘to invest’ (from *hanti* ‘property’ + *geli* ‘to make enter, insert’) to *deynqab* ‘to borrow’ (from *deyn* ‘debt’ + *qab* ‘to take’). As lexicalization is more common for nouns than for verbs, one finds in vocabularies *cidla* ‘empty area, wilderness’ (from *cid* ‘person’ + *la* ‘to be lacking, miss’) or *madaxla* ‘dead-end’ (from *madax* ‘head’, therefore “without head, headless”), but not the respective verbs. But consider now (14):

- (14) *dawo webi kama ag dhowaato, yaxaasna weli*
 jackal:SUBJ river to:NEG near approach:NEG:3M crocodile:SUBJ:and still
wuu carrab la'yahay
 DECL:he tongue miss:PRES:3M
 'Jackal still does not go near a river, while Crocodile still does not have
 his tongue' (Ciise 1985:29)

In (14) the object noun *carrab* 'tongue' is found immediately to the left of the verb and to the right of the subject clitic *-uu* (in *wuu*), and could be written as one word (*carrabla'yahay*), or also as *carrab-la'yahay*. The sentence is the last one in Ciise's (1985) version of a famous Somali folktale in which the jackal tricks the crocodile out of his tongue. While for all the length of the text the crocodile's tongue had always been referred to as 'the tongue', or, in dialogues, as 'your tongue', or even, simply, as 'the crocodile's tongue', *carrab* occurs here without any determiner — be it an article or a possessive — or modifier. The use of the bare form of the noun is scarcely possible in any alternative rendering of the sentence, such as, limiting the attention to the last clause in (14):

- (14') *yaxaasna weli carrabkii buu la'yahay*
 crocodile:SUBJ:and still tongue:M:ANPH FOC:he miss:PRES:3M

As discussed in Tosco (forth.), (14'), with nominal (object) focus, is grammatically possible but pragmatically odd, as the tongue is topical in the text, and no contrast with another object is implied. Another alternative to (14) is

- (14'') *yaxaasna weli carrabkii wuu la'yahay*
 crocodile:SUBJ:and still tongue:M:ANPH DECL:he miss:PRES:3M

Here, again, *carrab* should be definite and referential, and all the NPs positioned out of (i.e., before) the VC *wuu la'yahay*. (14'') is probably the best translation of English 'Crocodile still does not have a tongue', while the original (14) should rather be rendered as 'the crocodile is still tongueless': in noun incorporation a nominal element is backgrounded, detopicalized, and thereby de-contextualized and made non-referential. The crocodile's tongue is no longer the topic, as it was for all the preceding text; the sentence is rather about the crocodile's condition as a result of the jackal's trick.

In a situation in which any non-focused NP is topical, noun incorporation becomes a tool for detopicalizing an NP, commonly the object. Noun incorporation is therefore strictly intertwined with the expression of focus, which was seen above to play a central role in the syntax of Somali: rather than a simple bipartite

opposition between focused vs. topicalized elements, one therefore finds in Somali a threefold contrast at both the syntactic and the pragmatic levels:

<i>pragmatic value</i>	<i>syntactic marking</i>
new; contrastive	focus-marked (via a focus marker or clefting)
topical; given	not-focus-marked, out of the Verbal Complex
detopicalized, backgrounded; generic	not-focus-marked, incorporated into the verb

It seems therefore clear that in noun incorporation we are not dealing with a mere lexical-building device, but with a productive syntactic, “valency-filling” device.

As regards the wordhood of the noun-verb compounds, one can easily demonstrate that the incorporated noun does form a single word with the following verbal form on the basis of several phonetic and morphophonological points (and many irregularities notwithstanding):

1. the incorporated noun and the verbal form form one accentual unit, as shown by the frequent spelling as one word or with a hyphen separating the nominal and the verbal part;
2. feminine nouns ending in /o/ change it to /a/ when incorporated; e.g.: (*ka*) *talaqaado* ‘to take advice, to consult (with s.o.)’, from *talo* ‘advice, opinion’+ *qaado* ‘to take’. The same happens before nominal affixes; e.g., in *talada* ‘the opinion’ from *talo* + the feminine article *-ta* (which further undergoes intervocalic voicing);
3. the first stem consonant of the verb is often geminated after an incorporated noun, although this is only seldomly reflected in spelling; e.g.: (*ku*) *talaggal* ‘to consider, plan, decide (on s.t.)’ from *talo* + *gal* ‘to enter’;
4. in a few cases cluster simplification occurs; e.g.: *arrad* ‘lack (esp. of clothing)’ + *tir* ‘to cancel, destroy’ → *arrad-tir* or *arratir* ‘to find new clothes’.

Syntactically, once a noun has been incorporated, the resulting verb behaves as a single word and occupies the rightmost position in the VC, preceded by object pronouns, adpositions and/or adverbials:

- (15) *wuu ii il-jebiyay*
 DECL:he me:to eye.break:PAST:3M
 ‘he looked at me endearingly’ (‘eye-breakingly’)

3.1 Object demotion in noun incorporation

So far, noun incorporation was shown to be used in order to detopicalize and background a noun, but we have not provided evidence that it is used in order to give a transitive verb a generic object reading. Consider again (2):

- (2) *wax baan cunay*
 thing FOC:I eat:PAST:1s
 ‘I ate something’, ‘I ate’

In (2) the semantically very empty and generic noun *wax* ‘thing’ was used in order to saturate the object slot of a transitive verb and give it a generic object reading at a minimum cost (i.e., minimally affecting the semantics of the sentence). Due to the nearly obligatory marking of focus in any Somali declarative sentence, *wax* is put under focus. It is true that the same meaning could in principle be expressed also by the non-nominal-focus-marked sentence:

- (2′) [?]*wax waan cunay*
 thing DECL:I eat:PAST:1s
 ‘I ate something’, ‘I ate’

The problem with (2′) is that, as remarked by Gebert (1986) and Tosco (forth.), Somali speakers have problems in building focusless sentences whenever NPs are present, and the difficulty increases with the number of NPs in the sentence; this makes pragmatically good sense: NPs are generally mentioned in discourse when they are first introduced, or when they are salient enough. In both cases, Somali would mark the new or salient element through the focus markers *baa* or *ayaa*. A focus-unmarked NP is therefore possible, but generally only when another NP is focused, and the focus-unmarked NP would still be a topic. It was seen above that noun incorporation is used instead in order to detopicalize an object and make it non-referential, as seen in (13) above, repeated here below:

- (13) *iska jir, eygani waa dad-cunaa*
 REFL:from stay:IMPV:s dog:M:this:SUBJ DECL people.eat:PRES:3M
 ‘take care: the dog bites!’

If the noun is generic to start with and is further incorporated in the verb, the result will be a maximally empty element, whose only function will be to satisfy the valency requirement of a transitive verb, securing a generic object reading. What counts as a *generic* noun depends of course on the semantics of the verb; with *tol* ‘to sew’ it is apparently the noun *dhar* ‘clothes’. Again, the following sentence may only have an external object interpretation:

- (16) *waan tolay*
 DECL:I SEW:PAST:1S
 ‘I sewed it/ them’

while the following incorporated verb has a generic object reading:

- (17) *waan dhartolay*
 DECL:I clothes.sew:PAST:1S
 ‘I sewed’

Noun incorporation of a patient as a valency-saturating device is operative also with causative (syntactically ditransitive) verbs. In this case, incorporation may apparently only affect the primary object of the causative, which is the only object of the original monotransitive verb from which it is derived. (19) shows the by now usual incorporation of the patient of the monotransitive verb *daaq* ‘to graze’ whose non-incorporated use is shown in (18) vs. the incorporated and intransitive (19):

- (18) *geelu cawska bay daaqeen*
 camels:SUBJ grass:M:ART FOC:they graze:PAST:3P
 ‘the camels grazed (the grass)’
- (19) *geelu way caws-daaqeen*
 camels:SUBJ DECL:they grass.graze:PAST:3P
 ‘the camels fed on grass’

(21) further shows the incorporation of the primary object of the causative verb *daaji* ‘to make X graze Y’, whose non-incorporated use is shown in (20) vs. the incorporated and monotransitive (21):

- (20) *geela baan cawska daajiyay*
 camels:ART FOC:I grass:M:ART graze:CAUS:PAST:1S
 ‘I allowed the camels to graze (the grass)’
- (21) *geela baan caws-dajiyay*
 camels:ART FOC:I grass.graze:CAUS:PAST:1S
 ‘I fed the camels on grass’

3.2 The lexicalization of noun incorporation

The semantically emptiest noun *wax* ‘thing’ can also be incorporated, as in *waxtar* ‘to help, be useful’, built with *wax* and the verb *tar* ‘to assist’. Note in (22) here below the spelling, first as one word and immediately afterwards with the noun separated from the verbal form:

- (22) *Jabhada ayya ii baahan in aan waxtaro.*
 front:F:ART FOC me:to need:PRES:3F:FOC that I thing.assist:DEP:1s
*Wadankayga ayaan wax tarayaa*⁸
 country:M:my FOC=I thing assist:PROG:PRES:1s
 ‘the (Liberation) Front needs my help (lit.: ‘needs that I help’). I’ll be
 useful to my country’ (Faysal 2000: 16)

Here we find a further development: while the incorporated verb *waxtar* is used intransitively in the first clause (but a reading ‘the Front needs that I help it’ is equally possible), a new object is found in the second clause. A more conservative phrasing would have been:

- (22′) ... *Wadankayga ayaan wax u tarayaa*
 country:M:my FOC:I thing to assist:PROG:PRES:1s

(22′) would be in keeping with the *rule* given in Zorc’s (1993:413) dictionary, to the effect that ‘the roots must be separated except in the verbal noun’. The following example is provided:

- (23) *wax baan u taray*
 thing FOC:I to assist:PAST:1s
 ‘I assisted him, I was of use to him’ (Zorc 1993:413)

(22) represents therefore a change in progress in the incorporation of *wax tar*, in which *wax* loses even its valency-saturating role, the verb is treated as transitive and an adpositional phrase gets promoted to the role of primary object. In so doing, quite frequently the new object-incorporating verb receives a specialized meaning which is not the simple sum of the meanings of its component elements.

The simple verb *raac* ‘to follow’ and the compound verb *raadraac* ‘to follow traces’ (with the incorporated noun *raad* ‘trace’) may be used in approximately the same contexts. Only an anaphoric object reading is possible for the simple verb, as in

- (24) *waan raacay*
 DECL:I follow:PAST:1s
 ‘I followed him/her/it/them’

raadraac ‘to follow traces’ can instead either be used intransitively or get the meaning ‘to investigate’ and be used as a straightforward transitive verb

- (25) *waan raadraacay*
 DECL:I trace.follow:PAST:1s
 1. 'I followed, went trace-following'; 2. 'I investigated X'

It is therefore no surprise that, as was seen to be the case with non-incorporated verbs, the speakers more easily accept a generic object or intransitive reading when the verb is in the progressive aspect; obviously, only the transitive, derived meaning is instead possible with a non-null object:

- (26) *way i raadraaceen*
 DECL:they me trace.follow:PAST:3P
 'they investigated me'

Lexicalization is of course a matter of degree; e.g., from *madax* 'head' and *daali* 'to tire, make tired', the compound verb *madaxdaali* is created, with the somewhat specialized meaning of 'to give s.o. else problems, to tire somebody (with one own's problems)':

- (27) *arrintan ha igu madaxdaalin*
 question:F:this NEG me:with head.tire:IMPV:NEG.S
 'do not make me tired with this problem!'

Also *biyocabsi* 'to make s.o. drink (water)' can get a quite obvious lexicalization as 'to make s.o. drink too much water, to drown s.o.':

- (28) *haddaadan i dhaafin waan ku*
 if:YOU.S:NEG me leave:IMPV:NEG:S DECL:I YOU.S.O
biyocabsinayaa
 water.drink:CAUS:PROG:PRES:1s
 'if you don't let me (go) I'll make you drown (i.e., 'I'll make you drink water until you drown')

Full lexicalization, in which the incorporated verb gets a new meaning, is quite widespread in Somali (cf. Caney 1984 and Ajello 1995). As is cross-linguistically common (cf. Mithun 1984), terms for body parts are typical targets for incorporation in Somali as well as elsewhere in Cushitic (cf. Mous [1993:257 foll.] for the incorporation of body part terms in Iraqw). A non-exhaustive list of more or less idiosyncratic semantic shifts involving body parts as the first term of a noun-verb compound comprises:

1. *beer* 'liver' + *caddow* 'to become white' → *beercaddow* 'to become exhausted' (said of a child who has cried for a long time);
2. *calool* 'stomach' + *duubo* 'to fold o.s.' → *caloolduubo* 'to give up, forget about';
3. *dabo* 'tail; back' + *qabo* 'to catch' → *dabaqabo* 'to catch, seize by the tail';
4. *dhego* 'ears' + *faras* 'horse' → *dhegafarasoob* 'to hear things, hear imaginary sounds';
5. *far* 'finger' + *marooji* 'to twist' → (*ka*) *farammarooji* 'to snatch, get by force from s.o.';
6. *gacan* 'hand' + *qabo* 'to take for o.s.' → *gacanqabo* 'to help, assist';
7. *il* 'eye' + *geli* 'to make enter' → 'to look endearingly' (cf. (15) above);
- 7bis. *indho* 'eyes' + *duub* 'to fold; wrap' → *indhadduub* 'to blindfold';
8. *jilib* 'knee' + *dhig* 'to put down' → *jilibdhig* 'to dig in to fight; to assume a position to fight';
9. *kubo* 'back' + *dhagax* 'stone' → *kubadhagaxoob* 'to become constipated';
10. *laf* 'bone' + *gur* 'to pack up' → *lafaggur* 'to mend, doctor broken bones';
11. *lug* 'foot, leg' + *geli* 'make enter' → *luggeli* 'intervene, interfere';
12. *madax* 'head' + *furo* 'to free o.s.' → *madaxfuro* 'to redeem one's own or s.o.'s debts';
13. *san* 'nose' + *taag* 'to raise' → *santaag* 'to be proud'.

At least as common is the incorporation of relational nouns. These are nouns which enter into possessive constructions with a noun, e.g. *miiska hoostiisa* 'under the table' ("the table its-underneath"). They frequently occur as incorporated nouns, generally with little or no semantic shift:

1. *ag* 'proximity' → 'near' + *joog* 'to stay' → *agjoog* 'to stand by, close next to' (cf. also *ag dhowaato* in (14) above)
2. *dabo* 'tail; buttocks' → 'behind' + *dheh* 'to say' → (*ka*) *dabadheh* 'to repeat s.t. (after s.o.)'
3. *dib* 'back (n.), rear' → 'back' + *dhig* 'to put down' → *dibdhig* 'to postpone, delay'
4. *dhex* 'center, middle' → 'inside, between' + *mar* 'to pass' → *dhexmar* 'to transpire; pass through'
5. *dul* 'top, surface' → 'over, on top of' + *mar* 'to pass' → *dulmar* 'to go, pass over'
6. *hoos* 'lower part, bottom' → 'under' + *geli* 'to make enter' → 'to speak in an obscure, covered manner';
7. *hor* 'front' → 'in front of' + *tag* 'to go' → *hortag* 'to confront, face, meet'

8. *kor* ‘top; surface’ → ‘above, on top of’ + *meer* ‘to circle, roam about’ → *kormeer* ‘to inspect, check, supervise’

One further finds what I call, for want of a better name, *culture-specific incorporations*, in which a noun incorporates into a verb in order to create a lexicalized compound with a meaning which is culturally relevant in Somali ecology or social and spiritual life, as in *ab* ‘ancestor’ + *tirso* ‘to count for o.s.’ → *abtirso* ‘to trace one’s ancestry, to recite (‘count’) one’s genealogy’, or *guri* ‘house’ + *gal* ‘to enter’ → *guriggal* ‘to enter a new home after wedding’ (many other similar cultural compound verbs are listed by Ajello 1995). On the other hand, items belonging to the cultural vocabulary may also lose any specificity when incorporated, as in *xero* ‘enclosure, corral’ + *geli* ‘to make enter’ → *xerageli* ‘to obtain, acquire’.

In all these cases, one finds lexicalization with a new meaning which is not the mere sum of the meanings of the component parts, and on this basis one could hardly speak of a productive process. It was probably the role of noun incorporation as a powerful lexicon-enriching device that so far prevented Somali scholars to give noun incorporation its due place as a productive, grammatical mechanism.

4. A look beyond Somali

Since at least Fillmore (1986), one distinguishes in English between definite null complements and indefinite null complements: while the former allow recoverability from the context, the latter do not. Examples are, respectively: *he is eating* ___ (vs. **he is eating up* ___), and *Mary found out* ___ (examples from Velasco and Portero Muñoz 2002). In the case of Somali, we are obviously not interested in understood definite objects — all Somali transitive verbs admitting them in the form of a “ \emptyset object pronoun”. The crux of the question is indefinite objects.

As discussed by Velasco and Portero Muñoz (2002), the acceptability of English indefinite null complements is linked to a range of factors, among which imperfectivity (expressed either as present tense, or with continuous verbal forms or temporal expressions): while *she is knitting* is acceptable, **she knitted* is not (at least, unless an imperfective reading is forced upon the verb, e.g., through an appropriate time expression: *she knitted for three hours, then fixed supper*).⁹ In general, an indefinite object is all the more omissible the more it is predictable: So-called “cognate” objects are the most predictable, and verbs

which are restricted to them (as English *sing, dream, dance, &c.*) are the most obvious targets of indefinite object omission. Next, verbs which “take their object from a very limited range of potential candidates”, which, therefore, “can be easily retrieved if omitted” (Velasco and Portero Muñoz 2002: 10). Examples are English *read, write, spill*, but also *eat* and *sew*. Although we lack cross-linguistic studies of understood objects, it is probable that any language allows a pronominal object to be represented by \emptyset in at least a few contexts. And, of course, that an objectless sentence may be expressed, simply and iconically, through the *absence* of any object should come as no surprise to functionally-oriented linguists. The degree to which this is allowed probably varies from language to language. In Somali, the anaphoric-object interpretation remains nevertheless the default one, against which the generic-object reading must eke out a difficult existence, mainly “artificially” saturating the object valency of the verb through noun incorporation.¹⁰

It is difficult to judge the extent to which Somali represents a crosslinguistically rare pattern in this regard — many descriptions simply fail to provide the necessary details — but it is plausible that in Somali the overall generality of the anaphoric-object reading is connected to the pervasiveness of focus marking: Somali independent personal pronouns are, morphologically and syntactically, nouns, and as such they are normally followed by a focus marker whenever another NP is not available within the same sentence. It is a general rule of Somali text building that, whenever an element is introduced for the first time, or a topic is shifted, it is marked by the focus marker. At the same time, only one element at a time can be marked as “focused”. In such a situation, it is difficult if not impossible for Somali to introduce an anaphoric independent pronoun without marking it as focused. It is probable that languages which have \emptyset object pronouns of the 3rd person resort to the use of the full forms more often than languages with non- \emptyset object pronouns. While the use of a full, independent pronoun of the 3rd person (when available) is always to some degree a marked choice, it is probably never as strongly marked as clefting or other devices which are normally used for contrastive focus.

While the syntactic expression of pragmatic relations is a general feature of most Cushitic languages, its obligatory nature is very much restricted to the Central-Northern Somali dialects and the standard, literary language which is based upon them. Even the rather divergent Southern Somali dialects do not have obligatory focus marking: the following sentence is fully acceptable in, e.g., the Tunni dialect (from the Tunni version of the same folktale from which (14) above was taken):

- (29) *yahàas árrabkéy síyi*
 crocodile tongue:M:his give:PAST:3M
 ‘Crocodile gave (him) his tongue’ (Tosco 1997: 157)

Its Standard Somali equivalent would require both a Focus Marker and a Subject Clitic:

- (30) *yaxaas carrabkiisa buu síiyay*
 crocodile:SUBJ tongue:M:his FOC:he give:PAST:3M

While no Cushitic language has so far been made the object of such an in-depth syntactic analysis as (Standard) Somali, it seems safe to say that Somali represents more the exception than the rule in this regard, and that most other Cushitic languages behave rather like Tunni. For instance, in Gawwada, an Eastern Cushitic language of southwest Ethiopia, the sentence ‘I saw him’ has at least the following translations:¹¹

- (31) *an-híʔi*
 I-see:PAST:1s
- (32) *isó-si an-híʔi*
 he-OBJECT I-see:PAST:1s
- (33) *isó-kka an-híʔi*
 he-FOC I-see:PAST:1s

At the present state of our knowledge it appears that (31) is compatible with both a generic-object reading (‘I saw’) and with an anaphoric reading: in this context, (31) will be possible when the patient is given; it will, e.g., be uttered in replying to a question such as ‘did you see him?’; (32) seems to be the unmarked (and most frequent) solution, and will be uttered whenever the object is topic and the action is the new element, something like ‘as for him, I saw him’. For instance, a typical context for (32) could be when, after having introduced the referent of *iso* ‘he’, the speaker wants to further inform the addressee about what happened to him. Finally, (33) marks contrastive emphasis on the object: ‘it is him that I saw’, and could be uttered, e.g., in replying to the question ‘whom did you see?’

Obviously, the difficult point is to differentiate between (31) and (32). This is exactly what Somali does not do, both (31) and (32) being generally represented by (34), while (35) is limited to contrastive focus, topic shift and the like (cf. Tosco forth. for details):

- (34) *waan arkay*
 DECL:I see:PAST:1S
 'I saw him/her/it/them'
- (35) *isaga baan arkay*
 he FOC:I see:PAST:1S
 'I saw *him*'

In short, a Somali independent pronoun is normally in focus. It is true that Somali may also have recourse to clefting, but, as far as I know, there are no indications that a cleft sentence has a different pragmatic value than the corresponding sentence with nominal focus:

- (36) *waxaan arkay isaga*
 thing:I see:PAST:1S he
 'I saw *him*'

5. Conclusions

Polysynthetic languages are languages which obligatorily encode the marking of all (or most of) the participants in the action expressed by the verb within the verbal form itself; in other words, the valency of the verb is always fully saturated. In a recent article, Evans (2002) has brought the attention to the different strategies used in order to encode a generic object in polysynthetic languages, where the object marking of all transitive sentences is mandatory. Among these strategies, Evans (2002: 41–42) mentions the use of the third object pronouns to encode an indefinite reading, as in Bininj Gun-Wik (Australian); the reduction of 'the transitivity of the verb in situations where the object is not definite', e.g. through the use of a middle or antipassive, as in Eskimo or in Cayuga (Iroquoian); the use of 'a special object form dedicated to non-specific objects', as in Aztec; fourth, the possibility 'to have a conventionalized interpretation of some third person affix', as the use of the third plural for indefinite objects in Warray (Australian) or, again, in Cayuga. Finally, there is of course the possibility 'to exploit the choice between incorporation and the use of pronominal affixes, e.g. by reserving nominal incorporation for indefinite or generic nouns.' Obviously, a language may use, and is likely to use, a mix of more strategies.

Somali may be defined as a mildly polysynthetic language: it has obligatory subject and object marking on the VC and, just as in many other languages, polysynthetic and not, a non-local object participant has no marking. A transitive

verbal form may be interpreted either as containing a \emptyset anaphoric pronoun of 3rd person or no definite object at all; in this case, an internal or generic object reading will result. Which is which depends on a number of factors, yet to be fully investigated, but probably reflecting on the whole Hopper and Thompson's (1980) "Transitivity Scale" or some version thereof.

Just like many other polysynthetic languages, Somali has recourse to noun incorporation as a productive detransitivizing device. Following Mithun's (1984) classification of noun incorporation types and stages, this article has shown that Somali noun incorporation goes well beyond Type I, i.e., the stage of lexical compounding, coined for recognizable, "institutionalized" activities in which '[C]ompounds have a lexical status which their syntactic counterparts lack' (Mithun 1984: 848). Somali has Type II noun incorporation, in which the incorporated noun 'loses its syntactic status as a distinct argument of the clause' (Mithun 1984: 856) and case is manipulated.

Somali object of 3rd person has turned out to be, in Mithun's (1986) formulation, a classical example of a "zero who isn't there". Or, possibly, sometimes it is there, and sometimes not.

It is not among the aims of this paper to sketch, let alone to elaborate, a full-fledged "counter-syntax" of Somali; but I venture to predict that to take seriously into consideration the possibility of a generic/intransitive reading of transitive verbs and the productivity of noun incorporation will have fatal consequences for much current syntactic analysis of Somali. The non-argumental status of **all** non-focused NPs could well be the first casualty.

Notes

1. This article has been written during my stay at the Institut für Sprachwissenschaft of the University of Cologne (March-October 2002), which has been made possible by a scholarship from the Alexander von Humboldt Foundation. I thank Cabdulqaadir Salaad Dhoorre for his help and insights in analyzing the Somali material, as well as Hans-Jürgen Sasse for his comments and criticisms on earlier versions of this article. The usual disclaimers apply.

2. Somali examples are in the standard orthography, in use since 1973. Its main peculiarities are: <c> = [ɕ]; <x> = [ħ]; <dh> = [d]. The following abbreviations are used in the glosses:

ANPH	anaphoric article	NEG	negative particle; negative verbal form
ART	determinate article	O	object
CAUS	causative extension	P	plural
DECL	declarative particle	PAST	past tense
DEP	dependent mood	PRES	present tense

F	feminine	PROG	progressive aspect
FOC	focus particle; verbal form used with a focalized subject	REFL	reflexive object pronoun
IMP	impersonal subject pronoun	S	singular
IMPV	imperative	SUBJ	subject case form
M	masculine	VC	verbal complex

3. Although glossed with English personal pronouns, their Somali equivalents (such as *-aan* 'I' in (1) and (2)) are subject clitics, whose presence is obligatory (with certain limitations for the 3rd Singular Masculine *-uu*; cf. below (11) and (11')), and which are generally analyzed as syntactic subjects, rather than mere agreement morphemes (like, e.g., the subject clitics of many Romance varieties). Their syntactic status is in any case largely immaterial to the present discussion. The independent personal pronouns are instead always emphatic and never obligatory (cf. below (35) and (36)).

4. Subject is expressed here (as on most nouns) tonally through absence of a high tone; i.e., Cali (Absolute case form; [ǎli]) vs. Cali (Subject case form; [ǎli]). Tone is not represented in the standard orthography.

5. This sentence is pragmatically odd and obviously elicited out of context, as it contains three NPs and no focus marking (cf. Section 3.1. below).

6. Faysal's spelling is at times idiosyncratic, especially as geminated consonants and clitics are concerned: the Subject Clitic is generally written as one word with the preceding element; in this case *inaad* would be more common.

7. Both (11) and (11') are grammatically fully acceptable, while their pragmatic difference is still a matter of some debate (cf. Gebert 1986 and Tosco forth. for details). In general, (11'), in which the Subject Clitic is present, is less acceptable with a generic-object meaning, while it is preferred for 'this dog bites it/them'.

8. The usual spelling would be: '*Jabhadda...inaan waxtaro...Waddankayga...*'. Cf. fn. 6. above.

9. I owe this point as well as the example to an anonymous referee.

10. Another possible strategy, which will not be analyzed here, in order to encode a generic or intransitive reading involves the use of an intransitive, semantically object-incorporating verb; e.g., *quraaco* 'to have breakfast', *qadee* 'to have lunch', *cashee* 'to have dinner', instead of the transitive general verb *cun* 'to eat'.

11. The Gawwada data were collected by the author in 2000 and 2001. I gratefully acknowledge the help of the Italian Ministry of Education, University and Research (M. I. U. R.) for funding my research, and of the Institute of Ethiopian Studies at Addis Ababa University for granting the permission to carry on fieldwork in Ethiopia.

References

- Agostini, Francesco; Puglielli, Annarita; *et al.* 1985. *Dizionario Somalo-Italiano*. Roma: Ministero degli Affari Esteri, Cooperazione Italiana allo Sviluppo.

- Ajello, Roberto. 1995. "La focalizzazione in somalo". In: Ajello, Roberto; and Sani, Saverio (eds.), *Scritti linguistici e filologici in onore di Tristano Bolelli*, 1–28. Pisa: Pacini.
- Caney, John Ch. 1984. *The modernisation of Somali vocabulary with particular reference to the period from 1972 to the present*. Hamburg: Helmut Buske.
- Ciise Maxamed Siyaad. 1985. *Favole somale*. Roma: Ministero degli Affari Esteri, Dipartimento per la Cooperazione allo Sviluppo.
- Evans, Nicholas. 2002. "The true status of grammatical object affixes: evidence from Bininj Gun-Wok". In: Evans, Nicholas; and Sasse, Hans-Jürgen (eds.), *Problems of Polysynthesis*, 15–50. Berlin: Akademie.
- Faysal Axmed Xasan. 2000. *Maandeeq*. Scarborough, Ontario.
- Fillmore, Charles J. 1986. "Pragmatically controlled zero anaphora". In: Nikiforidou, Vassiliki; VanClay, Mary; Niepokuj, Mary; Feder, Deborah (eds.), *Proceedings of the twelfth annual meeting of the Berkeley Linguistic Society*, 95–107. Berkeley, California: Berkeley Linguistics Society, University of California.
- Gebert, Lucyna. 1986. "Focus and word order in Somali". *Afrikanistische Arbeitspapiere* 5: 43–69.
- Hetzron, Robert. 1989. "Typological peculiarities in Somali". *Folia Linguistica* 23(1/2): 7–26.
- Hopper, Paul J.; and Thompson, Sandra A. 1980. "Transitivity in grammar and discourse". *Language* 56(2): 251–299.
- Mithun, Marianne. 1984. "The evolution of noun incorporation". *Language* 60(4): 847–894.
- Mithun, Marianne. 1986. "When Zero isn't there". In: Nikiforidou, Vassiliki; VanClay, Mary; Niepokuj, Mary; Feder, Deborah (eds.), *Proceedings of the twelfth annual meeting of the Berkeley Linguistic Society*, 195–211. Berkeley, California: Berkeley Linguistics Society, University of California.
- Mous, Maarten. 1993. *A grammar of Iraqw*. Hamburg: Helmut Buske.
- Nordlinger, Rachel. 1998. *Constructive case: Evidence from Australian languages*. Stanford CA: CSLI.
- Puglielli, Annarita (ed.). 1981. *Sintassi della lingua somala*. Roma: Ministero degli Affari Esteri, Dipartimento per la Cooperazione allo Sviluppo.
- Saeed, John I. 1984. *The syntax of focus and topic in Somali*. Hamburg: Helmut Buske.
- Saeed, John I. 1987. *Somali reference grammar*. Wheaton MD: Dunwoody Press.
- Saeed, John I. 1993. "Adpositional clitics and word order in Somali". *Transactions of the Philological Society* 91(1): 63–93.
- Saeed, John I. 1994. "Syntactic typology and 'free word order' in Cushitic". *Teanga* 14: 58–75.
- Saeed, John I. 1999. *Somali*. Amsterdam: John Benjamins.
- Sasse, Hans-Jürgen. 2002. "Lexicological and lexicographic problems of word families in Cayuga". In: Evans, Nicholas; and Sasse, Hans-Jürgen (eds.), *Problems of Polysynthesis*, 203–221. Berlin: Akademie.
- Svolacchia, Marco; Mereu, Lunella; and Puglielli, Annarita. 1995. "Aspects of discourse configurationality in Somali". In: É. Kiss, Katalin É. (ed.), *Discourse configurational languages* 65–98. New York: Oxford University Press.
- Tosco, Mauro. 1997. *Af Tunni*. Köln: Rüdiger Köppe.
- Tosco, Mauro. Forthcoming. "A whole lotta focusin' goin' on: Information packaging in Somali texts". To appear in *Studies in African Linguistics*.

- Velasco, Daniel García; and Muñoz, Carmen Portero. 2002. "Understood objects in functional grammar". *Working Papers in Functional Grammar* 76.
- Zorc, R. David; and Osman, Madina M.. 1993. *Somali-English dictionary with English index*. Kensington MD: Dunwoody Press.

Author's address

Mauro Tosco
Università di Napoli "L'Orientale"
Dip. di Studi e Ricerche su Africa e Paesi Arabi
Piazza S. Domenico Maggiore 12
Palazzo Corigliano
I 80134 Napoli
Italy
mtosco@iuo.it