Mara Frascarelli and Annarita Puglielli

SOMALI IN A TYPOLOGICAL COMPARATIVE PERSPECTIVE: FROM DATA TO THEORY

1. *Introduction*

Working within a framework that assumes universal principles of grammar, the description of a language is only the first step of linguistic investigation. The morpho-syntactic and intonational properties of a language should, in fact, inspire the scholar to reflect and make a wider consideration regarding the structure and functioning of languages in a comparative perspective, in order to reach a deeper understanding of the language-system as a unified whole. For this reason, once the description of data has reached a satisfactory level, linguistic investigation must then proceed towards theory.

Thirty years of studies on Somali in the generative framework can be considered an important contribution in this direction: several works have been produced in which the analysis of specific phenomena has led to the formulation of relevant proposals for the theory of grammar.

The aim of this paper is to offer a short but, hopefully, significant overview of this analytical approach. In particular, section 2 deals with morphological classification, sections 3 and 4 with NP-internal phenomena (relative modifiers and Case marking, respectively), section 5 is concerned with the sentential level, focusing on subordinate clauses and, finally, sections 6 and 7 deal with the syntax-discourse-prosody interface through the analysis of Focus constructions.

2. *Morphological properties: for a typological classification*

As is known, languages of the world can be classified as inflectional, agglutinative, polysynthetic or isolating, according to their morphological properties (cf. Comrie 1981, Croft 1991). According to standard assumptions, every language is said to belong to one specific group. Let us therefore consider Somali in this respect.

At first glance, Somali could be classified as an inflectional language. This is the conclusion that we reach through the observation of its verbal paradigms. Consider, for instance, a verb like *sheeg* ('say') in the following table:

	present tense	present continuous	past tense
1SG	*sheegaa*	*sheegayaa*	*sheegay*
2SG	*sheegtaa*	*sheegaysaa*	*sheegtay*
3SGM	*sheegaa*	*sheegayaa*	*sheegay*
3SF	*sheegtaa*	*sheegaysaa*	*sheegtay*
1PL	*sheegnaa*	*sheegaynaa*	*sheegnay*
2PL	*sheegtaan*	*sheegaysaan*	*sheegteen*
3PL	*sheegaan*	*sheegayaan*	*sheegeen*

TABLE 1

As we can see, in the relevant paradigm the verbal stem (*sheeg-*) and a number of suffixes combine to realize person, number and gender features (the so-called 'phi-features'). These morphemes pattern fairly consistently across tenses. For instance, first singular and third singular masculine are always marked by the same suffix, as well as second singular and third singular feminine. We also notice that present tenses are characterized by a final long central vowel (*-aa*), while past tense mainly shows a final diphthong (*-ay*). Regular patterns are typical of inflectional morphology, thus confirming our initial hypothesis. However, proceeding with a morphological investigation, we notice the presence of other properties that do not belong to the inflectional type and require

some reconsideration. The most evident fact is the presence of a "Verbal Complex" (VC) in the sentence, that is to say, the existence of a clitic cluster in which arguments are realized as pronouns and elements are disposed in the templatic structure shown in Table 2:

impersonal SUBJ CL *la*	one Obj cl	Preps (2 at most)	Obj cl (poss. pron.)	deictics *soo / sii* of place manner	some Verb adverbials

TABLE 2

The VC can be therefore considered as a "microstructure of the whole sentence" (Puglielli 1981: 15), in which all elements of predication are represented. Consider the following sentences:

(1)
Axmed baa guri-gii [**nooga** (= Ø+na+u+ Ø+ka)
Axmed FM house-AN OCL3-OCL1PL-for-OCL3-from
soo qaaday]
take.PST
'Axmed took it from home for us'

(2)
Xaawo baa [**igu** (*i+ku*) **kaa** **aamintay**]
Xaaxo FM OCL1SG-a you (lit.: your) entrust.PST
'Xaawo entrusted me to you'

This morpho-syntactic property led Svolacchia and Puglielli (1999) to propose that Somali is a polysynthetic language (in the sense of Baker 1996), that is to say, a language in which clitics are only visible for q-role assignment through incorporation onto the verbal head. A crucial consequence of this analysis is that full NPs are realized in a 'dislocated' position and can never carry argument role.[1]

The polysynthetic proposal is supported by additional data. For instance, incorporation is not only specific of verbal arguments, but can also be found in the NP. Consider the following:

(3)
Wiil-kaa-gan-u *waa fiican yahay*
boy-POSS2SG-DEM-NOM DECL nice is
'This boy of yours is nice'

Clearly, determiners and possessives do not modify Nouns as independent elements in the NP, but they are cliticized onto the head-noun. Somali thus appears very similar to polysynthetic languages, as described in works by Jelinek (1988), Mithum (1987) and Baker (1988, 1996). Of course, incorporation in Somali is not a pervasive phenomenon as it is in languages like Mohawk, Oneida, or Cayuga, in which full sentences are realized as a single word, as shown in (4), from Evans and Sasse (2002):

(4)
Ę-s-kakhe-hona'tá-yęthw-ahs (CAYUGA)
FUT-ITER-SCL.1SG/OCL.3PL-potatoes-plant-PERF
'I will plant potatoes again for them'

Nevertheless, the presence of incorporational phenomena in (1)-(3) cannot be denied. Indeed, the recognition of a polysynthetic nature in Somali allows for important predictions. In particular, we can notice that not every 'dislocated' NP is resumed by a clitic pronoun in the VC. Consider the following:

(5)
Annag[oo guri-ga fadhina] bay (*na)
we-PRT house-DET stay FOC.CLSG.3PL 1PL
yimadeen
come.PRS-3PL
'They came while we were at home'
[lit.: 'we, who were at home, they came']

As we can see, the NP *annago* ('we') cannot be resumed in the relevant VC, while resumption is obligatory in sentences like (1)-(2) above. Since the VC is a 'microstructure of the sentence' in which all arguments must be realized, this kind of data leads us to hypothesize that it can only contain elements selected from the verb. As a matter of fact, *annaga* in (5) is not

part of the semantic grid of the verb 'to come', which only sub-categorizes a <theme> (projected as the sentential subject).

This kind of analysis has far reaching consequences for cross-linguistic investigation, since it implies that VCs in poly-synthetic languages can be used as valid diagnostics to check the argument structure of verbs, defining whether a constituent is an argument or not (depending on the presence of clitic resumption). As semantic structure is assumed to be universal, this would be of great significance for the theory, since the definition of theta-grids is very often a problematic issue and the source of longstanding discussion between scholars.

Turning back to morphological classification, Somali also shows the possibility of modifying the argument structure of the verb through affixes in the verbal stem, a typical property of agglutinative languages. For instance, the causative infix -i(s)- can be added to a monoargumental stative verb like buux ('to be full') yielding a biargumental transitive (hence, dynamic) one[2]:

(6)
Koob-ku   *waa*   *buuxay*
glass-DET.NOM   DECL   be.full.PST.3SGM
'The glass was full'

(7)
Cali   *koob-ka*   *waa buux-i-yay*
Ali.NOM   glass-DET   DECL   be.full-CAUS-PST.3SGM
'Ali filled the glass' (i.e., 'Ali made the glass full')

Then, if the morpheme -am- is added to *buuxis*, we obtain a verbal form that is traditionally considered as 'passive' (cf. 8). Unlike passives, however, the <agent> cannot be realized with this kind of verb (cf. 9):

(8)
Koob-ku   *waa*   *buuxsamay (buux-is-am-ay)*
glass-DET.NOM DECL   be.full-CAUS-STAT-PST.3SGM
'The glass filled up'

(9)
Koob-ku   *waa*   **ka**   *buuxsamay,*   **Cali**
glass-DET.NOM DECL from be.full-CAUS-STAT-PST.3SGM Cali
'*The glass filled up by Cali'

Cross-linguistic consideration can provide an explanation. Indeed, languages like Italian and English show that the <agent> of passive structures can always be 'rescued' through a (non argument) PP, while this is not the case with stative (unaccusative) verbs:

(10)
a. *La porta è stata aperta dal vento* (PASSIVE)
'The door was opened by the wind'
b. *La porta si è aperta (*dal vento)* (STATIVE)
'The door opened (*by the wind)'

The explanation for this difference rests on the eventive structure of the verb: passive constructions are dynamic events and, as such, they imply an <agent> which – though 'removed' as an argument – can be rescued as a circumstantial constituent. On the other hand stative verbs describe a state/property of the subject (semantically a <theme>); since there is no <agent>, it cannot be introduced as a PP. As a result, –am– infixed verbs should be classified as statives (for further discussion and data concerning the passive-stative dichotomy cf. Puglielli and Frascarelli 2008).

Resuming the typology issue of this section, we can finally conclude that in a language like Somali polysynthesis, inflectional and agglutinative properties coexist and concur in the interpretation of different phenomena. This means that typological classification is not always a clearcut distinction and 'mixed' types must be recognized. As a matter of fact, the interaction of different morphological properties can stimulate linguistic reflection and must be considered an interesting area for research in a comparative perspective.

3. Morpho-syntactic properties of NP modifiers: the case of relative clauses

Relative clauses in Somali are neither introduced by Complementizers (as in Italian, cf. 11a) nor by relative pronouns (as in Russian, cf. 11b) and the NP heading the clause is not resumed by a pronoun within the sentence (as is the case in Persian, cf. 11c), from Comrie 1981). Relative clauses in Somali thus belong to the 'head-deletion' type (cf. 12):

(11)
a. *Il ragazzo **che** hai incontrato è mio fratello*
'The boy (that) you met is my brother'
b. *Devuška [**kotor-uj** videla] maja sestra*
girl.NOM PRO.REL-3SGF.ACC see.PST.SGF POSS.1SG.F sister
'The girl (that) she saw is my sister'
c. *Man zan-i-ra [ke Hasan be u)*
PRO.S.1SG woman-DET-ACC that Hasan to PRO.O.3SG
jije-ra dad] mišenasad
chicken-ACC give.PST.3SG know.PRES.3SG
'I know the woman to whom Hasan gave a chicken'

(12)
a. *Wiil-ka [Maryan la hadlayaa] waa*
boy-DET.M. Maryan with talk.PRES.PROG.**RED.NOM** DECL
walaal-kay
brother-POSS.1SG
'The boy speaking with Maryan is my brother'
b. *Wiil-ka [af talyaani-ga ku hadlayá]*
boy-DET language Italian-DET in speak.PRES.PROG.**RED**
baan jeclahay[3]
FM.SCL.1SG love.PRES.1SG
'I love the boy that is speaking Italian'
c. *Moos-ka [aad cunaysaa] waa cerin*
banana-DET.M. SCL.2SG eat.PRES.PROG.**2SG.NOM** DECL unripe
'The banana that you are eating is unripe'

From a semantic point of view, relative clauses can be divided into restrictives and appositives, depending on their function with respect to the NP they modify. In particular, restrictive relative clauses provide a reference value for the NP-head (and are therefore necessary for interpretation), while appositives only supply additional information (and are therefore considered as 'circumstantial elements').

This semantic distinction is very often not associated with any morphosyntactic difference in the languages of the world and the relevant interpretation only relies on phonological means or pragmatic information. This is the case of Italian, illustrated in the following sentence:

(13)
Il tuo vicino di casa [che incontro spesso] è molto simpatico
'Your neighbour [that/who I often meet] is very nice'

If the relative sentence is produced within a single prosodic group with the NP-head, it receives a restrictive interpretation (in that case, the relevant neighbour is identified, among others, as the one that I often meet); on the other hand, if the relative clause is realized as an independent prosodic phrase (a sort of parenthetical), it is interpreted as appositive with respect to the NP-head (in that case, the relevant neighbour is not part of a set and its reference is independently identified).

Conversely, in a language like Somali appositive clauses are introduced by a specific morpheme, namely *oo*. Consider the following contrast:

(14)
a. *Wiil-kaas [Maryan la hadlayaa]*
boy-DEM.M.NOM Maryan with talk.PRES.PROG.RED.NOM
waa walaal-kay
DECL brother-POSS.1SG
'That boy that is speaking with Maryan is my brother'
b. *Wiilkaas [**oo** Maryan la hadlayaa] waa walaal-kay*
'That boy, who is speaking with Maryan, is my brother'

Given this formal difference, the morpho-syntactic properties of relative clauses in Somali can be studied without any doubt regarding their semantic nature. Consider for instance *stacking* phenomena:

(15)
a. *Wiilka* [*hadlayá*], [*oo aan ku*
   boy.DET talk.PROG.RED PRT SCL1SG OCL2SG
   baray],   *baa Landan ka yimid*
   introduce.PST.1SG FM London from come.PST.3SGM
   'The boy that is talking, whom I introduced you before, camefrom London'
b. *Wiilka [*oo aan ku baray*], [*hadlayá*], *baa Landan ka yimid*

In a sentence like (15a) we can unambiguously distinguish restrictive from appositive and see that appositive clauses cannot precede restrictives. This is evidence that the relation between a restrictive relative clause and the NP-head is somehow 'closer' and, consequently, these two types of relative clauses cannot be analyzed in a uniform way (as in other languages; for a discussion, cf. Frascarelli and Puglielli 1995b). The study of languages like Somali, in which semantic differences have a morpho-syntactic correlation, can be therefore very important for a deeper understanding of structures and their derivation.

4. Morpho-syntactic properties of syntactic functions: case marking

The realization of syntactic functions shows different strategies cross-linguistically. Some languages realize them through independent morphemes, namely *particles* (as in Chinese and a number of Creole languages) or *prepositions* (as in most European languages), while others use bound morphemes (i.e., *affixes*). Among the latter we can distinguish languages marking the NP (like Tzutujil in 16) from languages marking the verbal head (as in Standard Arabic, cf. 17). This difference

determines the distinction between 'dependent-marking' and 'head-marking' languages, according to Nichols' (1986) terminology:

(16)
x-ø-kee-tij   *tzyaq ch'ooyaa?*
ASP-**PRO.S.3SG-PRO.O.3PL**-eat cloths mouse
'The mouse ate the cloths'
[lit.: 'it ate them, the cloths the mouse']

(17)
'asbaha   *'akūh-u*   *mumaththil-an*
become.PERF.3SGM brother.**PRO.O.3SGM-NOM** actor.ACC-INDEF
'His brother became an actor'

As we can see, in Tzutujil subject and object pronouns are incorporated in the VC, while in Arabic the object pronoun (here interpreted as a genitive) and the NOM Case marking (indicating the subject function) are realized on the nominal element. In this respect, Somali can be classified as a dependent-marking language, as is clear from sentences like (18): Case marking is realized as the rightmost suffix in the NP, while the verbal head only presents inflectional morphemes:

(18)
Wiil-kaa-gan-u   *waa yimid*
boy-POSS2SG-DEM-NOM DECL come.PST-3SGM
'This boy of yours has come'

It is therefore interesting to notice that, in the case of complex NPs, some languages mark both the head-noun and its modifiers and, in coordinated structures both NPs are marked. This is the case of Standard Arabic and Turkish, illustrated respectively in (19) and (20) below:

(19)
[*al-lugha al-'arabiyya*]   *sa'abah*
DET-language.NOM DET-Arabic.NOM difficult
'Arabic language is difficult'

(20)

Ahmet [*uskumru-yu te istakoz-u*] *pişir-di*
Ahmet prawn-ACC and lobster-ACC cook-PST
'Ahmet cooked the prawn and the lobster'

In Somali, on the other hand, only the rightmost element within the NP is marked. From a syntactic point of view, this means that Case marking appears on the *most embedded* constituent. Consider coordination:

(21)

[*qalin-ka iyo buugag-gu*]   *miis-ka   way*
pen-DET   and book.PL-DET.NOM table-DET DECL.SCL3PL
saran yihiin
stay.PRES.3PL
'The pen and the books are on the table'

As we can see, NOM Case is only seen on the second member of coordination (*buggag*), while the first NP (*qalin*) shows the unmarked (non-nominative) Case marking.

This morpho-syntactic property allows for a clear and straightforward identification of the right boundary of NPs and – more importantly – it leads to significant considerations about the internal structure of coordination. Indeed, this kind of data shows that coordinated NPs are included in a 'bigger NP' and that, within this structure, the rightmost NP is syntactically the most embedded one. This is perfectly in line with an analysis of coordinated NPs as members of a 'Coordination Phrase', based on the binary-branching X-bar model (as proposed in Kayne 1994).

According to this proposal, coordinated phrases are inserted in the Specifier and Complement positions of a CoordP, which assumes the categorical specification of its internal members. In other words, the CoordP containing two NPs is, in turn, a 'big NP':

(22)
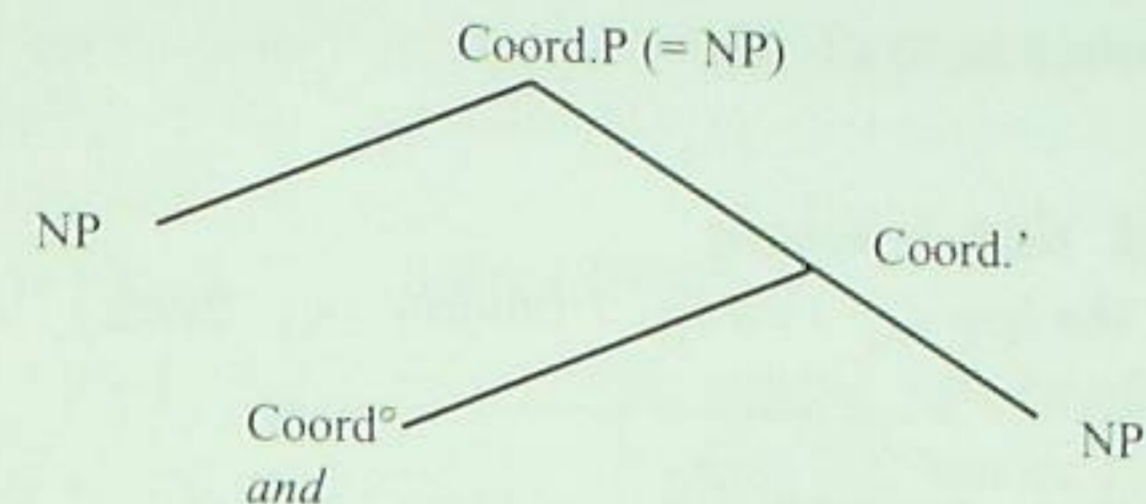
The rightmost constituent in this structure is necessarily the most embedded one, hence – in languages like Somali – the NP carrying morphological markings. No other stipulation is needed to explain this order of facts.[4]

Syntactic embedding is therefore the key to understand rightmost Case marking in any type of complex NP. Interestingly, in a language like Somali this analysis also applies to the relative clause, a type of nominal modifier that is rarely marked for Case in the languages of the world. Consider the following sentence:

(23)

[*wiilka   aan af   Talyaani-ga ku hadlin-i*]   *waa*
boy-DET   NEG language Italian-DET   in talk.NEG-NOM DECL
walaal-kay
brother-POSS1SG
'The boy that cannot speak Italian is my brother'

The fact that the head-noun (*wiilka*) has no Case marking, while NOM Case is present on the relative verb (*hadlin*) shows that these elements are contained within one and the same NP. Hence, Case marking proves that a relative clause is a NP-internal constituent and that the relative verb is its most embedded element.

This piece of evidence strongly supports a *raising* analysis for (restrictive) relative clauses[5], as proposed in Kayne (1994). According to this analysis, the relative clause (a CP) contains the NP-head and is inserted as the complement of the functional D° head. The NP-head then *raises* to scope position

through wh-movement. In other words, a relative clause like (24a) is derived as in (24b):

(24)

a. *The **book** that I bought*

b. [DP [D' *the* [CP [C' *that* [IP *I bought* [NP ***book***]]]]]]

This structure accounts for the argument role of the NP-head (*book*) with respect to the relative clause (namely, <patient>) and for its syntactic properties, after Operator-movement to Spec,CP.

Once again, we have seen that linguistic reflection on the morpho-syntactic properties of a single language can lead to wider considerations about universals and the theory of grammar.

Let us now proceed on this analytical track, focusing on subordination.

5. *Morpho-syntactic properties at the sentential level: subordination*

Subordinate clauses in Somali are constructed like relative clauses. Indeed, they are characterized by the same properties, including the antiagreement phenomenon when the NP-head has a subject role (see note 3). Consider the following:

(25)

a. *gabar-dhii oo [**mar-kaas** guri-ga gashay]*
girl-AN   PRT time-DEM   house-DET enter.PST.RED
baan   arkay
FM.SCL.1SG see.PST.1SG
'I saw the girl *after* she got home'

b. *[**goor-ta**   qorrax-du   dhacdó] imaw*
moment-DET   sun-DET.NOM set.DEP come.IMP
'*When* the sun is setting, come'

c. *[**si-da**   Ahmed uu   doonayó]   erey-ga*
way-DET Ahmed SCL.3SGM want.PROG.DEP word-DET
u   qor
to write.IMP
'Write the word *as* Ahmed wants'

The sentences in brackets, interpreted as adverbial clauses, are introduced in English by elements that can be considered as Complementizers. In Somali, on the other hand, these clauses are dependent on noun: the nominal nature of these elements (*markaas, goorta, sida*) is in fact unquestionable, since they are modified by demonstratives or determiners. This means that the relevant clauses are included within NPs or, in other words, that sentential subordination is realized as *nominal subordination*. This is a crucial conclusion, which deserves further investigation. First of all, let us see whether this property only concerns adverbials or can be extended to all types of subordinate clauses.

At a first glance, completive clauses in Somali seem to fall outside this generalization since they are introduced by an element that might be considered a COMP, that is to say, *in*:

(26)

Waxay   doonaysaa   [in-ay   bisha   dambe
FM-SCL3SGF want.PROG.3SGF **that**-SCL3SGF month-DET next
tagtó]
leave.DEP
'She wants to leave the next month'
[lit.: 'she wants that she leaves the next month']⁶

As we can see, *in* is normally translated as 'that' in English. However, the word *in* is a nominal head in Somali, as is proved by its use in sentences like (27):

(27)

in   lacag ah   i   sii
part/thing money be.PRES.3SG OCL1SG give.IMP
'Give me some money' [lit.: 'give me part of what is money']

This means that completive clauses – like adverbials – must be also considered on a par with relative clauses. Morpho-syntactic evidence in this respect is discussed in Antinucci (1981) and additional support has been provided in recent analyses on intonational properties (Puglielli and Frascarelli 2006).

As a matter of fact, prosodic investigation has shown that *in* is pronounced with a *pitch* and marks the beginning of a *prosodic domain* concluding its (falling) curve at the end of the subordinate clause. This is exactly the prosodic behaviour of NPs heading relative clauses (contrary to COMPs, which are never marked by any prominence). Consider, for instance, the intonational contour of sentence (28), shown in Figure 1 below:

(28)

(waxay   ila   tahay)   Cali in-uu
FM.SCL3SGF OCL1SG-IMPERS seem.3SGF Cali part/thing-SCL3SGM

cajiin qasayaa   la   moodaa
pasta make.3SGM.DEP IMPERS imagine.PRES.3SGM
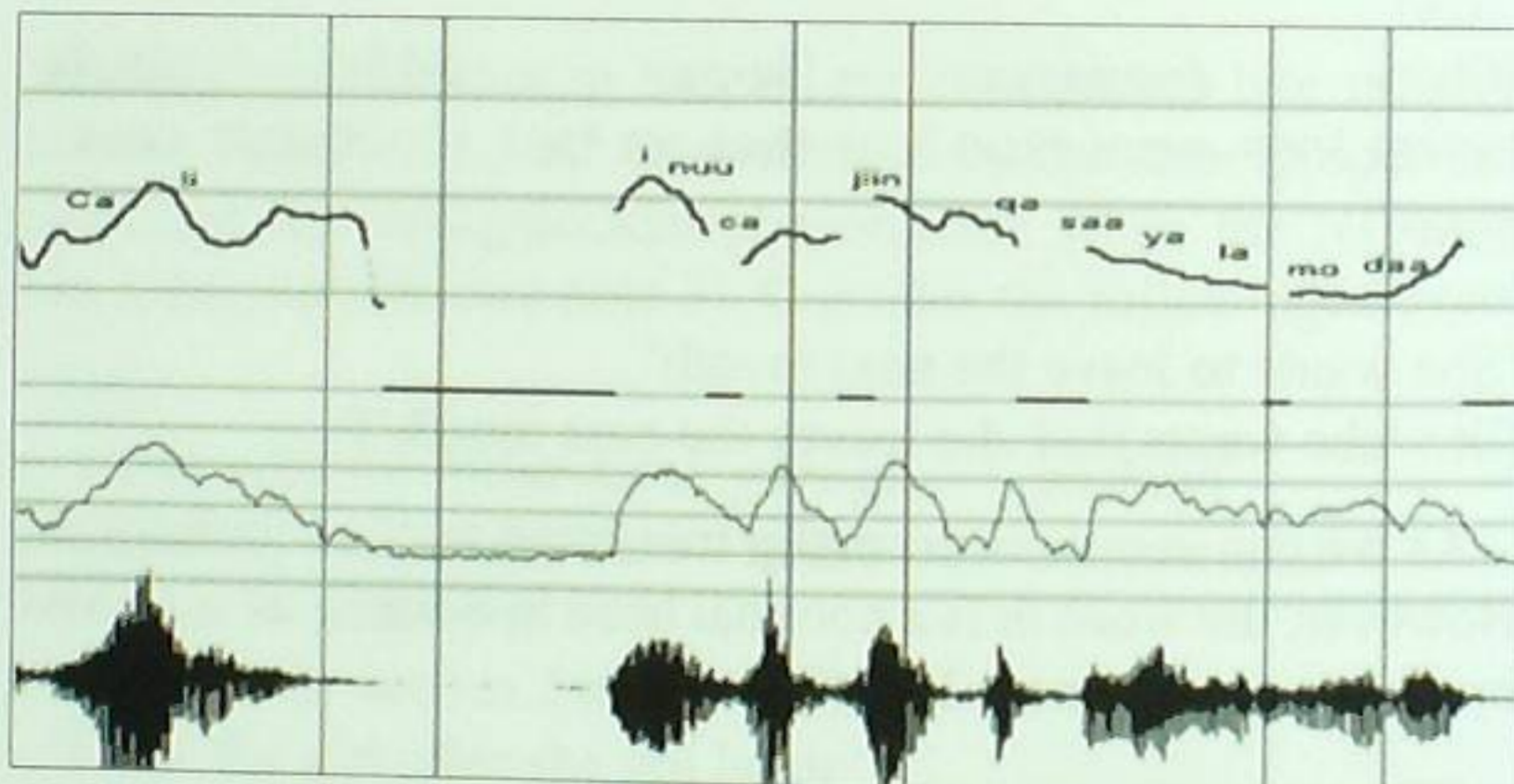'It seems to me that Cali is making pasta, I think'

FIGURE 1

As we can see, *in* represents the highest F0 point in the sentence, while the curve reaches its baseline after the verb *qasaayaa*. We can therefore conclude that the "COMP" *in* is in

fact the head of a relative clause. It is now important to acknowledge that the morpho-syntactic properties discussed for subordinate clauses in Somali do not represent an isolated case cross-linguistically. Indeed, subordinates are realized as nominalized constructions in a number of (typologically diverse) languages. In Turkish, for instance, subordinate verbs show a nominalizer infix and the adverbials are introduced by generic NPs ('time', 'place', 'reason', etc.) located at the end of the relevant clause, consistent with the head-final nature of the language. This is exemplified in (29a-b) (from Kornfilt 1997):

(29)

a. *Müdür   [tatil-e   çik-tiğ-i   **zaman**]*
   manager holyday-DAT go-NOMIN-3SG.POSS time
   ofis   kapa-n-ir
   office   close-RIFL-AOR
   'When the manager is on vacation, the office is closed'
   [lit.: 'The manager, the time he is on vacation, the office is closed']

b. *Hasan   [kitab-i   san-a   ver-diğ-im*
   Hasan   book-ACC PRO.S.2SG-DAT give-NOMIN-1SG.POSS
   ***takdir-de**] çok   kiz-acak*
   case-LOC   very get angry-PST
   'Hasan got very angry because I gave you the book'
   [lit.: 'Hasan, at the instance that I gave you the book, got very angry']

Similarly, adverbial clauses in Maori are introduced by generic NPs and the embedded verbs are not preceded by any tense/aspectual morpheme, while this is always present in matrix clauses. Consider the following (from Bauer 1993):

(30)

a. *[**kia** tae   mai koe]   ka   kai taatou*
   time arrive here PRO.S.2SG T/ASP eat PRO.S.1PL.INCL
   'When you arrive, we will eat'

b. *I   hoki   maatou   ki   te   kaainga*
   T/ASP go back   PRO.S.1PL.EXCL to   DET house
   [*i   te   mea e   ua   ana*]
   from   DET   thing T/ASP rain PST
   'We went back home because it was raining'

Also in these languages – as in Somali – the nominal nature of the elements introducing the adverbial clauses is made clear by the possibility of being modified by determiners and preceded by prepositions (cf. (30b)).

Adverbial clauses are also introduced by nominal heads in Austronesian languages, as is shown below for Tagalog (for additional data and discussion, see Puglielli and Frascarelli 2008):

(31)
[***bago***   *natulog*] *ang mga bata   uminom   ng gatas*
precedent   ATT.sleep TRIG PL   children ATT.drink.PST DIR milk
'Before sleeping, children drank milk'

Given these data, it is feasible to hypothesize that sentential subordination actually originates from nominal subordination in a cross-linguistic perspective. In this line of analysis, wh-constituents introducing embedded clauses in languages like English or Italian should be considered as NPs heading relative clauses; on the other hand, COMPs introducing completives introduce in fact the complement CP of a NP whose head is null (i.e., phonologically unrealized). In other words, subordinate clauses like those in (32a) and (33a) originate from nominal subordinate structures like the ones shown in (32b) and (33b):[7]

(32)
a. [***Quando*** *arriverai*], *saremo tutti molto felici*à
   '***When*** you arrive, we will be all happy'
b. [***Nel momento*** [*in cui arriverai*], *saremo tutti molto felici*
   'In the moment in which you arrive, we will be all happy'

(33)
a. *Ricordo* [***che*** *ti sei comportato molto male*]
   'I remember that you behaved very badly'
b. *Ricordo* [<*il fatto*> [***che*** *ti sei comportato molto male*]]
   'I remember the fact that you behaved very badly'

The nominal nature of embedded clauses is undoubtedly an interesting hypothesis that languages like Somali present to the attention of scholars. Evidently, this kind of analysis has far reaching consequences that need further research and supporting data in a cross-linguistic perspective (see also Testa 2007 for discussion).

6. *The syntax-discourse interface: focus constructions*

The realization of Focus[8] is a much debated issue in the literature and different theories have been proposed. As a matter of fact, languages show different strategies to realize the constituent carrying new information, based on different grammatical means (i.e., prosody, morphology, syntax – or a combination of them).

Despite the great variety of details, however, authors agree in the identification of two main strategies to realize Focus cross-linguistically, namely, the *in situ* and the *extra situm* strategy (cf. Aboh *et al.* 2007). The former is apparently the simplest, since the Focus constituent remains in its basic position (as in Italian, cf. (34)), while the latter requires Focus to move to scope position, left-adjacent to the verbal head (as in Hungarian, cf. (35), from Horvath 1995):

(34)
Ho incontrato LEO *ieri*
'I met LEO yesterday'

(35)
MARI *jött   el   tegnap*
Mari come.PST PV   yesterday
'MARI came yesterday'

Focus is a crucial notion in Somali since it is a 'Focus-prominent' language (cf. Kiss 1995), that is to say, one constituent must be overtly signaled as new information in the sentence. In particular, Focus is identified by means of its syntactic position (obligatorily before the VC) and for the presence of a right-adjacent Focus marker (*baa*):

(36)
a. *Shalay   jamacadda   CALI **baan** (baa + aan)*
   yesterday university-DET Cali FM.SCL1SG

   ku arkay
   to see.PST.1SG
   'Yesterday I saw CALI at the university'
b. **Shalay CALI jamacadda baan ku arkay*
c. **Shalay jamacadda baan CALI ku arkay*

The question is, which Focus strategy does Somali belong to? Given the syntactic position of Focus we might suppose an *extra situm* case. However, deeper consideration of morpho-syntactic data shows that this conclusion is not correct.

First of all, the sentence following the Focus has the morpho-syntactic properties of a relative clause, in contrast to languages like Hungarian. Hence, when the Focus is interpreted as the subject of the sentence, antiagreement effect appears. Consider the following sentence:

(37)
NIMANKÁAS   baa hilibka   cunayá
men-DEM.**NONNOM** FM meat.DET eat.PROG.**RED**
'THOSE MEN are eating the meat'

As we can see, the 'subject' NP 'those men' is not marked with the NOM suffix (i.e., *–u*, cf. section 4) and the verb shows the reduced paradigm. Hence, the relevant NP does not show any of the typical morpho-syntactic properties which characterize a grammatical subject (though it is semantically interpreted as the subject of the sentence). Second, the *extra-situm* strategy does not include Focus markers, whose presence must be explained. In this respect, diachronic research has played a crucial role, shedding light on their *copular origin* in a number of languages (cf. Frascarelli and Puglielli 2005a,b, 2007a,b, Frascarelli 2008).

We must therefore conclude that Somali uses a different strategy for Focus, which includes a relative clause and a copular construction. Let us therefore proceed in the analysis of data, in order to understand this type of construction in detail, starting with the correlation between Focus and relative clauses.

Enlarging the spectrum of our observation, it is important to notice that this correlation has been pointed out by many authors (starting with Schachter 1973) for a number of typologically different languages, like Tigrinya (from Appleyard 1989), Wolof (from Kihm 1999), Byali (from Reineke 2004), Berber (from Ouhalla 1999) and French:

(38)
NKSSU   'Kyyu   waga 'ab kätäma zäxKbbKr (TIGRINYA)
PRO.3SGM COP.3SGM price in town   raise.REL
'HE raises the prices in the town'

(39)
FAS   WI   la   jaaykat bi jënd   (WOLOF)
horse DET COP.3SG merchant DET buy.PST.REL
'The merchant bought A HORSE'

(40)
BíiG   è   ù   n   yi yambK kK   (BYALI)
child.OCL COP SUBJ.3SGF REL PST see.PERF OCL
'She had seen THE CHILD'

(41)
TAMGHARTA ay   yzrin   Mohand   (BERBER)
woman.DEM PRON.INDEF see-PART Mohand
'Mohand saw THAT WOMAN'

(42)
C'est MOI qui est tombé dans l'escalier (FRENCH)
'It's ME who fell from the staircase'

As we can see, in all these languages Focus is followed by a relative clause expressing the presupposed part of the sentence. How can we explain its presence and contribute to the theory of grammar in a cross-linguistic perspective?

In recent analyses in the Generative framework a *cartographic approach* has been proposed to account for the syntax and interpretation of discourse categories (cf. Rizzi 1997). According to this approach, the original CP node has been 'split' into a number of functional projections, each dedicated to the interpretation of a specific discourse category (Focus, Topic, Contrast and so on). The left periphery therefore shows a very complex structure, which can be (minimally) represented as follows:[9]

(43)

[ForceP [TopP [FocP [TopP [FinP [IP   ]]]]]]

For the purposes of the present analysis, this means that Focus moves to the FocP projection in languages with the *extra situm* strategy.

The morpho-syntactic restrictions observed in Somali Focus constructions strongly support this kind of approach. Indeed, the rigid word order observed in sentence (36) above can be easily accommodated in the functional structure given in (44):

(44)

[TopP*shalay* [TopP *jamacadda* [FocP*Cali* [Foc·*baan* [TopP*anigu* [IP*ku arkay*]]]]]]

'Yesterday I saw CALI at the university'

The ban against alternative orders (cf. 36b-c) can be thus considered as a direct consequence of the cartographic order of functional projections in the left periphery. This is a very important issue, since the realization of Focus has been regarded as a mere pragmatic concern for a long time; on the contrary, the crucial role of syntax is made clear by languages like Somali.

However, the cartographic approach by itself cannot explain the presence of a copula, the presupposition embedded in a relative clause and the morpho-syntactic properties of a focused 'subject' (cf. (38)). Theory must be improved to comprehend the realization of Focus in languages like Somali and Tigrinya. Therefore, a third syntactic strategy is needed, which can be defined as the '*cleft-like*' strategy (cf. Frascarelli and Puglielli 2005a, Frascarelli 2008).[10]

Given the presence of a copular construction, this strategy requires the presence of a Small Clause (see Moro 1997), in which Focus is inserted as the *predicate*. Presupposed information, on the other hand, is nominalized and realized as a (free) relative clause (as is the case of all embedded clauses in Somali, cf. Section 5), inserted as the subject of the relevant structure. Finally, the relative clause is headed by a generic null NP (e.g., 'person', 'thing', 'time'), whose value is specified by the Focus after movement to scope position (i.e., SpecFocP). The derivation of sentence (38) can be thus represented as follows:

(45)
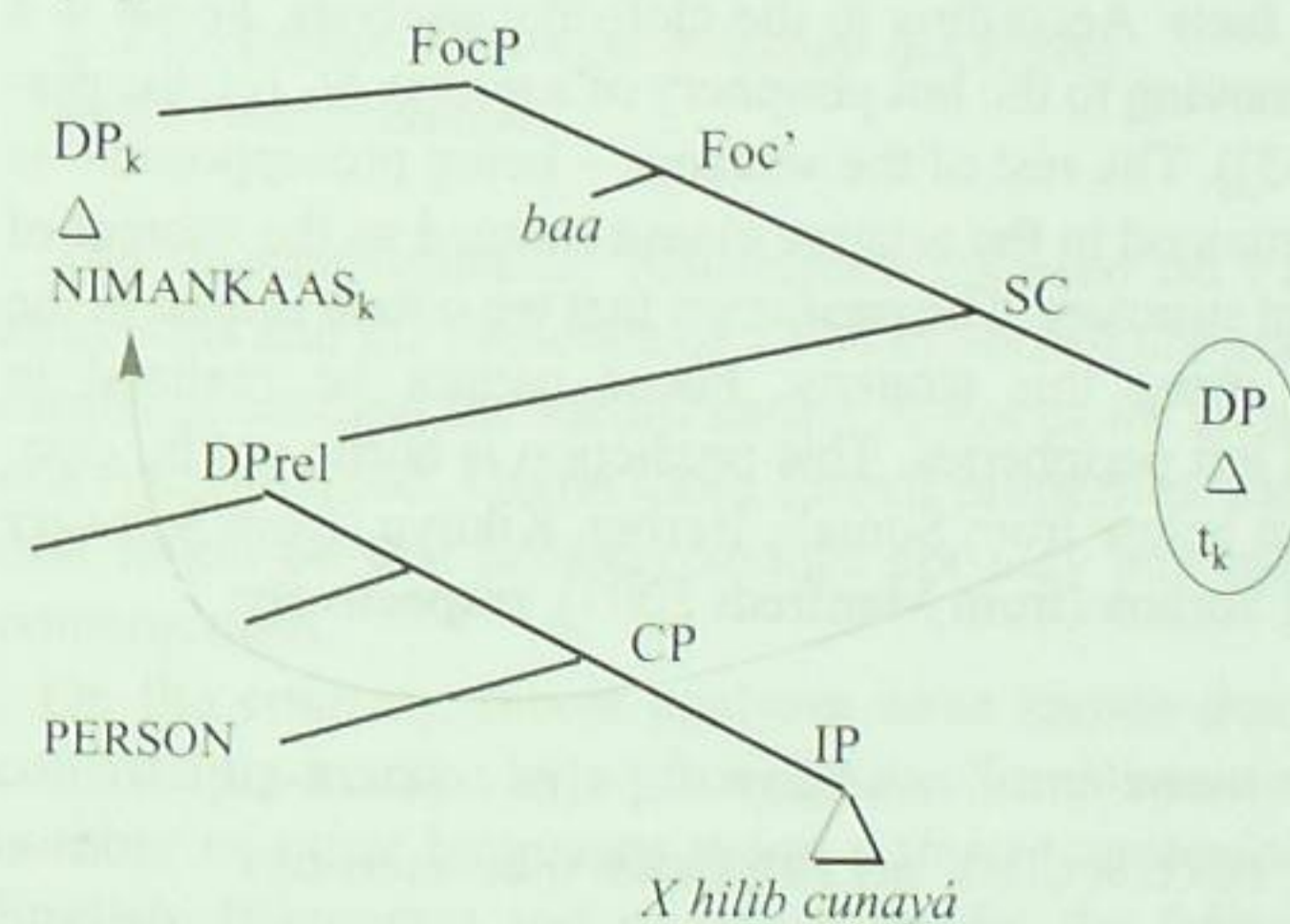
Therefore, this analysis does not exclude the cartographic approach, but integrates Focus movement into a more complex structure in order to account for the morpho-syntactic properties so far observed.

Indeed, the copular structure in (45) accounts for the fact that a Focus cannot have the properties of a grammatical subject since it is not inserted as an argument of the verb (being the predicate itself). It also accounts for the presence of a relative clause: it is the syntactic realization of the set of elements within which the Focus operates, providing a value for an open variable ('X', in 45). In other words, a sentence like (42) implies the recognition of [someone who ate the meat] as the subject of predication and that of a nominal predicate (NIMAN-KAAS) providing a value for the entity having such a property.

The idea of Focus as a predicate has been proposed in different frameworks with different results (cf., among others, Higgins 1973, Szabolcsi 1981, Kiss 1999). Chomsky (1971) also claimed that "the focused constituent is the predicate of a *dominant sentence*" (Chomsky 1971: 72, italics mine). The originality and the typological relevance of the present proposal rests on the assumption that the 'cleft-like' construction qualifies as a *syntactic strategy* to realize Focus in some (non-related) languages.[11]

Finally, the existence of the cleft-like strategy is supported by additional facts. According to the cleft-like analysis, Focus is a predicate moving to the left periphery of a *matrix SC* (cf. the diagram in (45)). The rest of the sentence – being presupposed – is always contained in the relative clause merged as the subject of the relevant structure. The prediction that we obtain is that, in the languages using this strategy, Focus cannot be realized in embedded left peripheries. This prediction is borne out by data, as is shown below from Somali, Berber, Kikuyu (from Schwarz 2004) and Yoruba (from Manfredi 2007), respectively:

(46)
*Hilib-ka waan   sheegay   [in   NIMAN-KA$_i$
meat-DET DECL.SCL1SG say.PST.1SGM that men-DET
baa cunayá]
FM eat.PROG.RED
'I said that THE MEN are eating the meat'

(47)
*Nna-n qa [TAMGHART-A   ay   yzrin   Mohand]
say.PST-1SG that woman-DEM PRO.INDEF see.PST.PART Mohand
'I said that THIS WOMAN saw Mohand'

(48)
*Abdul, [ne IBUKU o-dom-ire]   á-nyu-ire   mae
Abdul   COP book 3SG-read-PST 3SG-drink-PST water
'Abdul, who was reading a BOOK, was drinking water'

(49)
*Won bèèrè pe [boya mo rà FILA ni l'anàa]
3PL ask say if 1SG buy hat COP yesterday
'They ask whether I bought A HAT yesterday'

Given these data we can reach the conclusion that Focus is a discourse feature linked to the left periphery of matrix clauses. This conclusion has important consequences for the theory, and its cross-linguistic extent and validity can be a fruitful subject for future research.

7. *The syntax-discourse-prosody interface: focus interpretation*

The morpho-syntactic restrictions imposed on Focus constructions and the presence of a FM in Somali are grammatical means permitting the identification of Focus for an unambiguous interpretation. Given these formal properties, the expectation might be that prosody should not play any role in such constructions.

On the contrary, recent analyses have shown that Focus is consistently marked by a pitch (H*) in Somali as well, as in a number of other languages using different strategies (Italian, English, Hungarian and so on). Consider the following sentence and its intonational realization in Figure 2 (from Puglielli and Frascarelli 2006):

(50)

dhinaca **waxaan**   (*wax+baa+aan*)   *u jeedaa*   BAKEERIGA
side   FM.SCL1SG   thing-FM-SCL1SG   see.PRES.1SG   glass.DET
MARYAN*uu (= M.+buu) horyalla*
Maryan.FM.SCL3SGM   be in front of.PRES.3SGM
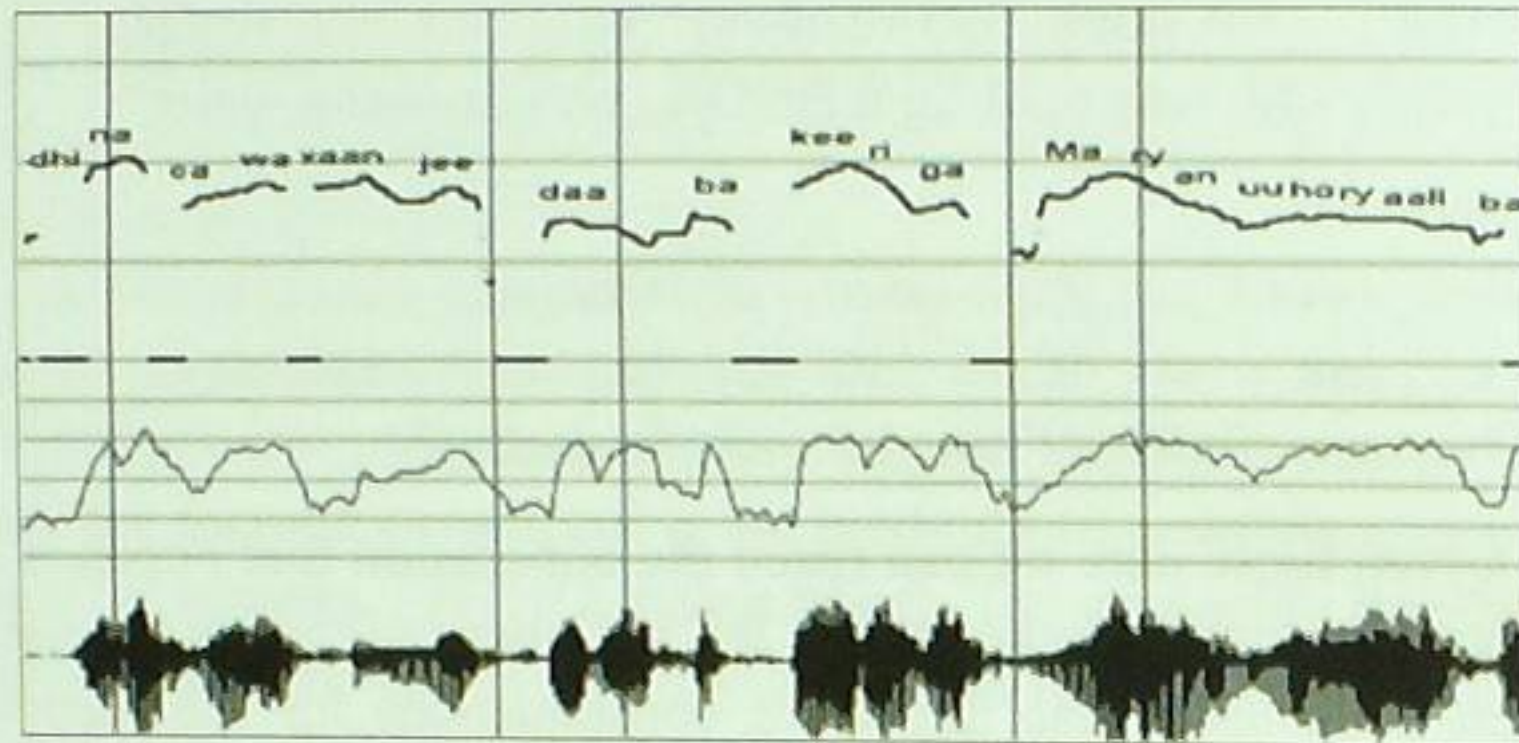'On a side, what I see is a GLASS (and) it is in front of MARYAN'

FIGURE 2

As we can see, both BAKEERIGA ('the glass') and MARYAN are marked by a pitch (H*) and represent the highest point within their intonational phrases. On the other hand, the FMs (*waxaa* and *baa*) are prosodically unstressed.

Both these results are perfectly in line with the cleft-like analysis we propose: Focus moves to Spec,Foc and, as such, is recognized at the syntax-discourse-prosody interface as the element checking the Focus feature. Hence, it is marked by a pitch. The FM, on the other hand, is nothing but a copula, namely, a mere functional head carrying inflectional features and serving as a 'linker' between the presuppositional subject and its nominal predicate. Intonational analysis thus supports the proposal discussed in section 6 and functions as valid diagnostics in verifying the validity of syntactic proposals in cross-linguistic works.

A final suggestion for linguistic reflection concerns the so-called 'division of labour'. As we have seen, despite the richness of morpho-syntactic means to signal Focus, prosody is not

'silent', as might be expected. This means that the economy of system does not work blindly, simply avoiding redundancy of grammatical means, and that a deeper analysis is needed to understand why, especially in the realization of discourse features, syntax, morphology and prosody seem to be jointly required for interface interpretation. This is a crucial issue for the theory of grammar, which can lead to a deeper understanding of data in future research.

NOTES

[1] This means that full NPs are realized in an extrasentential position and associated with a specific discourse value (as Topics or Foci). This is a crucial characteristic of Somali that is not discussed in this paper. The interested reader can refer to Puglielli (1981), Svolacchia et al. (1995), Frascarelli and Puglielli (2005a, 2007a,b, in press).
[2] On the stative-dynamic dichotomy and its morpho-syntatic consequences, cf. Puglielli and Frascarelli (2008).
[3] As we can notice, when the NP-head has a subject role in the relative clause (cf. (12a-b)), verbal inflection appears in a 'reduced' form (the so-called *Antiagreement* phenomenon; see section 6.1). The 'reduced paradigm' has three forms: one for 3SGF, one for 1PL and one for all other persons. It can be therefore considered as a participial-like kind of inflection (cf. Puglielli 1981, Puglielli and Frascarelli 2008).
[4] This implies that non-configurational ('flat') coordinated structures, having all members on the same syntactic level have no explanatory power and must be therefore excluded.
[5] As discussed in section 3, restrictive and appositive relative clauses cannot be given a uniform syntactic analysis.
[6] Evidently, Somali does not allow for infinitive subordinate clauses, a property deriving from its polysynthetic nature (cf. section 2; on the morpho-syntactic consequences of polysynthesis, cf. Baker (1996)).
[7] Null elements are indicated in angle brackets, as in standard use.
[8] As agreed among scholars, the term 'Focus' refers to new information in the sentence. Its extent and semantic value can vary and many distinctions have been proposed (for an overview and some discussion, cf. Rizzi 1997, Frascarelli 2000, Puglielli and Frascarelli 2008). For the purpose of the present work, we only refer to Focus in its *narrow* syntactic realization (i.e., limited to one constituent in the sentence), with a semantic *informative* role (i.e., as the element answering a wh-question).
[9] Indeed, after Rizzi (1997) – from which the structure in (44) is taken – many

authors have provided evidence for the necessity of additional functional projections in the CP area. For the purposes of this paper we concentrate on the original proposal, but refer the interested reader to Puglielli and Frascarelli (2008) for discussion and references.

[10] It is important to insist on the fact that this strategy is used to realize *information* Focus. So, even though this structure is reminiscent of cleft constructions, it is *not* associated with a contrastive interpretation of Focus, as is case of clefted foci in most European languages (like Italian).

[11] For reasons of space we cannot add further details here; the interested reader is referred to Frascarelli (2008).

REFERENCES

Aboh, E.O., Hartmann, K., Zimmermamm, M., eds. (2007), *Focus Strategies in African Languages. The Interaction of Focus and Grammar in Niger-Congo and Afro-Asiatic*, Berlin/New York: Mouton de Gruyter.

Antinucci, F. (1981), "Tipi di frase", in Puglielli A. (a cura di), *Sintassi della lingua somala*, Studi Somali 2, Roma: Ministero Affari Esteri, pp. 219-300.

Appleyard, D. L. (1989), "The Relative verb in Focus constructions: an Ethiopian areal feature", in *Journal of Semitic Studies*, 34 (2), pp. 291-305.

Baker, M. (1988), *Incorporation. A Theory of Grammatical Function Changing*, Chicago: The University of Chicago Press.

Baker, M. (1996), *The Polysynthesis Parameter*, Oxford / New York: Oxford University Press.

Bauer, W.A. (1993), *Maori*, London: Routledge.

Chomsky, N. (1971), "Deep Structure, Surface Structure and Semantic Interpretation", in Chomsky N. (ed.), *Studies on Semantics in Generative Grammar*, The Hague: Mouton, pp. 62-119.

Comrie, B. (1981), *Language Universals and Linguistic Typology: Syntax and Morphology*, Chicago: The University of Chicago Press.

Croft, W. (1991), *Syntactic Categories and Grammatical Relations. The Cognitive Organization of Information*, Chicago/London: University of Chicago Press.

Evans, N. Sasse, H.J., eds. (2002), *Problems of Polysynthesis*, Berlin: Akademie Verlag.

Frascarelli, M. (2000), *The Syntax-Phonology Interface in Focus and Topic Constructions in Italian*, Dordrecht: Kluwer Academic Publishers.

Frascarelli, M. (2008), "Focus, Clefting and Predicate Inversion", ms. (submitted to *Lingua*).

Frascarelli, M., Puglielli, A. (2005a), "A Comparative Analysis of Restrictive and Appositive Relative Clauses in Cushitic Languages", in Brugè, L., Giusti, G., Munaro, N., Schweikert, W., Turano, G. (eds.), *Contribution to the IGG XXX*, Venezia: Cafoscarina, pp. 307-332.

Frascarelli, M., Puglielli, A. (2005b), "The Focus system in Cushitic lan-

guages", in Fronzaroli, P., Marrassini, P. (eds.), *Proceedings of the 10th Meeting of Hamito-Semitic (Afroasiatic) Linguistics*, Firenze: Dipartimento di Linguistica, Università di Firenze, pp. 333-358.

Frascarelli, M., Puglielli, A. (2007a), "Focus in the Force-Fin System. Information Structure in Cushitic Languages", in Aboh, E. O., Hartmann, K., Zimmermamm, M. (eds.), *Focus Strategies in African Languages. The Interaction of Focus and Grammar in Niger-Congo and Afro-Asiatic*, Berlin/New York: Mouton de Gruyter, pp. 161-184.

Frascarelli, M., Puglielli, A. (2007b), "Focus Marking and Universal Grammar", in Azeb, A., Mous, M., Savà, G. (eds.), *Cushitic and Omotic languages: Proceedings of the fourth International Conference on Cushitic and Omotic Languages (Leiden, 10-12 April 2003)*, Köln: Rüdiger Köppe Verlag, pp. 119-134.

Frascarelli, M., Puglielli, A. (2009), "Position, Function and Interpretation of Topics in Somali", in Mereu L. (ed.), *Information Structure and Its Interfaces*, Berlin/New York: Mouton de Gruyter, pp. 325-348.

Higgins, F. R. (1973), *The pseudo-cleft construction in English*, PhD dissertation, MIT.

Horvath, J. (1995), "Structural Focus, Structural Case and the Notion of Feature Assignment", in Kiss K.É. (ed.), *Discourse-Configurational Languages*, Oxford: Oxford University Press, pp. 176-206.

Jelinek, E. (1988), "Voice and Transitivity as Functional Projections in Yaqui", in Butt, M., Geuder, W. (eds.), *Projections from the Lexicon*, Stanford: CSLI Publications, pp. 195-224.

Kayne, R. (1994), *The Antisymmetry of Syntax*, Cambridge, MA: The MIT Press.

Kihm, A. (1999), "Focus in Wolof: A study of what morphology may do to syntax", in Rebuschi, G., Tuller, L. (eds.), *The Grammar of Focus*, Amsterdam / Philadelphia: John Benjamins, pp. 245-274.

Kiss, K.É., ed. (1995), *Discourse-Configurational Languages*, Oxford: Oxford University Press.

Kiss, K. É. (1999), "The English Cleft Construction as a Focus Phrase", in Mereu L. (ed.), *Boundaries of Morphology and Syntax*, Amsterdam / Philadelphia: John Benjamins, pp. 217-230.

Kornfilt, J. (1997), *Turkish*, London: Routledge.

Manfredi, V. (2007), "Nuclear stress in Eastern Benue-Kwa (Niger-Congo)", in Aboh, E. O., Hartmann, K., Zimmermamm, M. (eds.), *Focus Strategies in African Languages. The Interaction of Focus and Grammar in Niger-Congo and Afro-Asiatic*, Berlin/New York: Mouton de Gruyter, pp. 15-45.

Mithum, M. (1987), "Is Basic Word Order Universal?", in Tomlin R. S. (ed.), *Coherence and grounding in discourse*, Amsterdam/Philadelphia: John Benjamins, pp. 281-328.

Moro, A. (1997), *The Raising of Predicates*, Cambridge: Cambridge University Press.

Nichols, J. (1986), "Head-marking and Dependent-marking Grammar", in *Language*, 62, pp. 56-119.

Ouhalla, J. (1999), "Focus in Arabic Clefts", in Rebuschi, G., Tuller, L. (eds.), *The Grammar of Focus*, Amsterdam / Philadelphia: John Benjamins, pp. 335-359.

Puglielli, A., a cura di (1981), *Sintassi della lingua somala*, Studi somali 2, Roma: Ministero Affari Esteri.

Puglielli, A., Frascarelli, M. (2006), "Information Structure in Somali. Evidence from the syntax-prosody interface", paper presented at the Workshop on *Minority Languages. Lexical Aspects and Discourse Grammar*, University of Roma Tre.

Puglielli, A., Frascarelli, M. (2008), *L'Analisi Linguistica. Dai dati alla teoria*, Roma / Cesena: Caissa Italia Editore.

Reineke, B. (2004), "Identificational Operation as a Focus strategy in Byali", paper presented at the Workshop on *Topic and Focus: Information Structure in African Languages*, University of Amsterdam.

Rizzi, L. (1997), "The Fine Structure of the Left Periphery", in Haegeman L. (ed.), *Elements of Grammar. Handbook in Generative Syntax*, Dortrecht: Kluwer Academic Publishers, pp. 281-337.

Schachter, P. (1973), "Focus and Relativization", in *Language*, 49, pp. 19-46.

Schwarz, F. (2004), "Kikuyu Focus Constructions. Syntactic and Semantic Issues", paper presented at the Workshop on *Topic and Focus: Informational Structure in African Languages*, University of Amsterdam.

Svolacchia, M., Mereu, L., Puglielli, A. (1995), "Aspects of Discourse Configurationality in Somali", in Kiss, K. É. (ed.), *Discourse Configurational Languages*, Oxford: Oxford University Press, pp. 65-98.

Svolacchia, M., Puglielli, A. (1999), "Somali as a Polysynthetic Language", in Mereu, L. (ed.), *Boundaries of Morphology and Syntax*, Amsterdam / Philadelphia: John Benjamins, pp. 97-120.

Szabolcsi, A. (1981), "The semantics of Topic-Focus articulation", in Groenendijk, J., Janseen, T., Stokhof, M. (eds.), *Formal methods in the study of language*, Amsterdam: Mathematisch Centrum, pp. 513-541.

Testa, G. (2007), *La sintassi interna ed esterna di alcune frasi avverbiali. Uno studio comparativo*, Tesi di dottorato, Università degli Studi Roma Tre.