Luigi Rizzi
Subjects, topics and the interpretation of *pro*

1 Introduction

In this paper I would like to address the classical issue of the similarities and differences between subject and topic positions. In doing so, I will build on seminal work by Andrea Calabrese, particularly Calabrese (1986) on pronominal interpretation. This paper, as well as other contributions on prosody, syntax and information structure (Calabrese 1982, 1992) anticipated and inspired a view of the syntax-pragmatics interface which turned out to be of crucial importance for cartographic projects.

The theoretical context for the analysis of subject positions is offered by the recent discussion on freezing and labeling, and the "halting problem" for movement. Certain syntactic positions are "halting sites" for syntactic movement, and give rise to freezing effects. These are the criterial positions in the sense of Rizzi (1997), for instance the position hosting the wh-element in questions, but also focus and topic positions. A restrictive approach to labeling (Chomsky 2013, 2015; Rizzi 2015a, b, 2016; see also the papers in Bošković 2016) has been recently shown to offer a comprehensive account of the halting problem, capturing both cases in which movement must continue from an intermediate position, and cases in which movement must stop.

The criterial-labeling approach, originally motivated by properties of A'-syntax, naturally extends to A-syntax. The subject position is the typical "halting site" of A-movement. Under the approach just outlined, it is expected to be a criterial position, expressing a Subject Criterion analogous, *mutatis mutandis*, to the Q Criterion, the Focus Criterion and the Topic Criterion of A'-syntax. One would therefore expect the subject position to manifest the kind of scope-discourse interpretive properties which are typical of criterial positions, i.e., giving rise to an articulation akin to topic – comment, focus – presupposition, operator – scope domain, as in familiar A'-constructions. In Rizzi (2005, 2006, and much related work) I have argued that the subject position occurring in the high part of the IP structure expresses the argument "about which" the event expressed by the predicate is presented. This aboutness property is independent from information structure (see section 4), and has consequences for discourse organization and anaphora resolution. Calabrese (1986) observed that in a null subject language like Italian, *pro* in the following sentence (in certain structural configurations) picks out the aboutness subject: this is what I will refer to henceforth as "the Calabrese effect".

The "aboutness" approach makes the subject position akin to the topic position, which also involves "aboutness" (as well as other properties which distinguish topics from subjects: see section 5): the comment is a statement made about the topic; nevertheless subjects and topics should be distinguished, even in Null Subject Languages in which a full assimilation is at first sight very plausible. In this paper, I will address the different formal and interpretive properties of subjects and topics, taking as a point of departure Calabrese's insight on their role in anaphora resolution. The goal is to arrive at a comprehensive analysis of these positions in the context of the cartographic study of structures and interfaces.

2 The "halting problem" for phrasal movement

Phrasal movement proceeds in successive steps, in accordance with fundamental locality principles (Chomsky 1973). Some intermediate positions require movement to continue (transiting positions), whereas other positions require movement to stop (halting position). Familiar illustrations of these effects are provided by A-bar movement, e.g., wh-movement from the clausal complements of different verbs. A verb like *think* in English requires the complementizer system of its complement clause to function as a transiting position. From an initial representation like (1)a, the wh element must transit through the embedded complementizer system, yielding an intermediate representation like (1)b; but (1)b cannot be the final representation, the embedded C-system is a transiting position, from which the wh-element must move further, yielding the main question (1)c:

(1)   a.   John thinks [C [Bill read [which book]]]
      b.   * John thinks [[which book] [C [Bill read __]]]
      c.   [[Which book] [does [John think [__ C [Bill read __]]]]]?

(that the wh-element transits through the embedded C-system, as predicted by locality, is straightforwardly shown, e.g., in languages with overtly "agreeing complementizers", which morphologically mark the transit of the element: see van Urk 2016 for recent discussion). The reciprocal pattern is offered by the complement of a verb like *wonder*:

(2)   a.   John wonders [C [Bill read [which book]]]
      b.   John wonders [[which book] [C [Bill read __]]]
      c.   * [[which book] [does [John wonder [__ C [Bill read __]]]

movement to the embedded complementizer, as in (2)b yields a well-formed structure here, an embedded question. But this is now a halting position: movement cannot continue as in (2)c, we have a freezing effect in the halting position.

The notions of halting and transiting positions are, of course, descriptive categories: ideally the status of a position wrt movement should be derived from fundamental ingredients of linguistic computations. An approach to the halting problem based on the labeling algorithm is presented in Chomsky (2013, 2015), Rizzi (2015a-b, 2016) (see also the other papers in Bošković 2016).

The crucial structural property here is the "criterial configuration", the fact that in the embedded C-system of (2)b the wh-element and the C agree in criterial feature +Q, the feature designating questions:

(2')   b' John wonders [$_\alpha$ [which$_{+Q}$ book] [+Q [Bill read ___]]]

Here both phrases *[which$_{+Q}$ book]* and *[+Q [Bill read ___]]*, (internally) merged together yielding an XP-YP configuration, carry the criterial feature +Q on their most prominent element. So, both give consistent instructions for the labeling of the mother node α, which can be labeled as Q, an indirect question. The wh-phrase can therefore remain in the embedded C-system, as far as labeling is concerned. Moreover, it must remain there because of the maximality principle of Rizzi (2015a-b, 2016), stating that only maximal objects with a given label can be moved, so that we have the freezing effect illustrated by the ban on (2)c. In a nutshell, after labeling of α has taken place in (2')b, *which book* is not anymore the maximal object carrying the categorial feature Q, as it is immediately dominated by a node labeled as Q, hence it is unmovable in accordance with maximality. Reciprocally, labeling of the embedded clause would not be possible in (1)b (XP and YP do not provide a coherent labeling instruction here), so that further movement of the wh-phrase is required to permit labeling of the embedded clausal node, and its C-system necessarily is a transiting position. The halting or transiting status of a position can thus be derived in a principled way, ultimately from the labeling algorithm and the maximality principle.

If we now consider the A-system, we also find transiting positions, such as the subject position of an untensed raising verb (3)b, and halting positions, such as the subject position of a tensed complement (4)b:

(3)   a.   (it) seems [to have been arrested John]
       b.   * (it) seems [John to have been arrested __]
       c.   John seems [__ to have been arrested __]

(4) a.   (it) seems [could have been arrested John]
    b.   (it) seems [John could have been arrested __]
    c.   * John seems [__ could have been arrested __]

The crucial point here is that the subject position of a tensed clause is the funda-
mental halting position of A-movement, a position in which movement stops, and
which arguably gives rise to freezing effects, as (4)c illustrates for A-movement
(see Rizzi 2006, Rizzi & Shlonsky 2007 and much related work for freezing effects
connected to the subject position with respect to both A- and A'-movement).[1]

If halting positions typically are criterial positions, the halting property of the
high, clause initial subject position (the EPP position in GB syntax) leads us to the
following hypothesis:

(5)   There is a subject criterion

Criterial positions are normally associated to special scope-discourse effects such
as, in the A'-system, the expression of articulations such as topic – comment,
focus – presupposition, operator – scope domain. So, what could be the interpre-
tive property, relevant for scope-discourse interpretation, associated to a subject
criterial position?

Here it is important to distinguish between different positions which can be
referred to as "subject positions": the thematic position of the subject (Spec-vP
in non-unaccusative structures), the low subject position used for subject focal-
ization in (some) Null Subject Languages (Belletti 2004), the high, clause initial,
subject position corresponding to the EPP position of GB syntax (with the possibil-
ity of a further proliferation of non-thematic subject positions in the high IP field:
Cardinaletti 2004). The position which is central to our discussion is the latter:
the canonical subject position, the typical final landing site of A-movement. What
is the contribution of this position to interpretation?

In Rizzi (2005, 2006, and much related work) I have argued that this subject
position expresses the argument which is selected as the starting point in the

---

1 A reviewer raises the question of whether and how the classical distinction between A and
A' movement has an impact on freezing. The analysis developed in the references quoted in the
text gives rise to the expectation that a criterial position (be it A', as the specifier of Q in an
indirect question, or A, as a subject position) always gives rise to freezing effects, and for both
A- and A'-movement. In fact, in the references quoted, that-trace effects are analyzed as cases of
freezing, an analysis which implies that criterial subject positions disallow further A' movement.
If subjects are indeed harder than objects to extract, languages often allow subject extraction in
special configurations (special complementizer forms, etc.), which are analyzed in the referenc-
es quoted as devices to allow a subject wh to avoid the freezing position and remain extractable.
See, in particular, Rizzi & Shlonsky (2007).

description of the event: the subject is the argument "about which" the event expressed by the predicate is presented. The "aboutness" effect is clearly illustrated by active-passive pairs:

(6) a. Piero ha colpito Gianni
        'Piero has hit Gianni'
    b. Gianni è stato colpito da Piero
        'Gianni has been hit by Piero'

The same "hitting" event is presented as being about the agent in (6)a, and about the patient in (6)b. So, passivisation may be seen as a device to shift aboutness from one argument to another in the argument structure of the verb.

3 The Calabrese effect (Calabrese 1986)

The aboutness property seems to be very subtle and elusive, but years ago, Andrea Calabrese discovered a clear test sensitive to aboutness subjects. The null pronominal subject *pro* in Italian picks out the aboutness subject of the immediately preceding clause in certain structural contexts. The effect is particularly clear in the core configuration that Calabrese discusses, i.e., when an adverbial clause precedes the matrix clause (his ex. (1)):

(7) Quando Mario$_i$ ha picchiato Antonio$_k$, *pro*$_{i\,*k}$ era ubriaco
     'When Mario hit Antonio, pro was drunk' (pro = Mario)

Calabrese observes that *pro* can also pick out another salient referent from the previous discourse. For instance, if in the discourse context preceding (7) we have been talking about Francesco, *pro* in (7) could refer to him; e.g., if Francesco is a hanging topic:

(7') (A proposito di) Francesco$_j$, Quando Mario$_i$ ha picchiato Antonio$_k$, *pro*$_j$
      '(As for) Francesco, when Mario hit Antonio, pro was drunk' (pro =
      era ubriaco
      Francesco)

but if we want to search for the antecedent of *pro* in the immediately preceding adverbial clause, *pro* must pick out the subject of predication. In the configuration at issue, if we want to express the fact that Antonio was drunk, we must use the overt subject pronoun *lui*:

(7'') Quando Mario$_i$ ha picchiato Antonio$_k$, *lui*$_{*i\,k}$ era ubriaco
       'When Mario hit Antonio, pro was drunk' (pro = Antonio)

Calabrese also shows that the fact of picking out the aboutness subject is a specificity of *pro* and does not extend to other weak pronominal forms in Italian. A clitic pronoun can refer to both the subject and the object (his ex. (3)):

(8) Quando Mario$_i$ ha picchiato Antonio$_k$, io lo$_{i, k}$ ho visto sanguinare
  'When Mario hit Antonio, I saw him bleed' (ambiguous)

Calabrese treats this effect through two principles (his (6) and (5)):

(9) Use a stressed pronoun only when the occurrence of its referent is not expected

(10) A subject pronoun is expected to have the referent of another subject (in the immediate context)

Principle (9) may be seen as a variant (or perhaps the natural complement) of the Avoid pronoun principle of Chomsky (1981): if there is an alternation between a stressed and a null form, use the null form for expected referents, and the stressed form for unexpected referents. Principle (10) says something specific to subjects, in our sense of aboutness subjects (Calabrese's subject of predication, what he calls "thema"): distinct predicates, in the same local domain, tend to be predicated about the same referent, perhaps a particular case of the functional linguistics' notion of "topic continuity" (Givon 1983), Calabrese argues.[2]

We can now go back to an active-passive pair: in the context identified by Calabrese, *pro* always picks out the aboutness subject, the agent in the active and the patient in the passive:

(11) a. Quando Mario$_i$ ha picchiato Antonio$_k$, *pro*$_{i *k}$ era ubriaco
    'When Mario hit Antonio, pro was drunk' (pro = Mario)
   b. Quando Antonio$_k$ è stato picchiato da Mario$_i$, *pro*$_{*i, k}$ era ubriaco
    'When Antonio was hit by Antonio, pro was drunk' (pro = Antonio)

So, Calabrese's effect offers a reliable way to detect an interpretive property that goes with the aboutness subject. As criterial positions typically involve special interpretive properties, this finding supports the view that the aboutness subject position is a criterial position.

---

2 The basic pattern observed by Calabrese seems to require full native command of null subject properties: in an experimental study, Belletti, Bennati & Sorace (2007) found out that near-native L2 speakers of Italian were significantly more lenient than native speakers in admitting coreference between an overt pronominal and a previous subject.

4 An apparent exception to the Calabrese effect: the relevance of c-command

Calabrese also observes that the system does not seem to work when we have a configuration with a main and a complement clause:

(12)   Francesca ha fatto notare a Maria che *pro* era molto stanca
       'Francesca made Maria realize that pro was very tired'

In fact, the interpretation of *pro* here is fully ambiguous, *pro* could equally well refer to Francesca or Maria.
    I believe that what is going on in this case is that both DP's c-command *pro*, which suggests another principle like

(10')   A subject pronoun is expected to have the referent of a c-commanding DP

That c-command may be relevant here is shown by the following:

(13)   Francesca ha fatto notare alla sorella di Maria che *pro* era molto stanca
       'Francesca made Maria's sister realize that pro was very tired'

Here *pro* can naturally refer to Francesca or to Maria's sister, but not to Maria. To express that interpretation, the overt pronominal subject must be used (even though the exclusion of the null form in the intended interpretation may be less sharp than in the previous cases, the preference pattern seems to me to be clear):

(13')   Francesca ha fatto notare alla sorella di Maria$_i$ che *lei*$_i$ era molto stanca
       'Francesca made Maria's sister realize that pro was very tired'

This pattern follows from the combined action of (9) and (10'): by (10'), both DP's *Francesca* and *La sorella di Maria* are expected to be antecedents of the subject pronoun, which, by (9), can be *pro* in that interpretation. In contrast, *Maria* does not c-command the pronoun, so that it is not an expected antecedent (it is not a subject either, so that (10) is not operative, either); therefore, under (9), the pronoun must be overt in that interpretation.
    As c-command makes every DP an expected antecedent for a pronoun, the Calabrese effect is fully visible only in cases in which c-command (and principle (10')) does not hold, so that only principle (10) is operative in determining what an expected antecedent can be: this happens with preposed adverbials, and also in discourse sequences, as we will see in a moment.
    The relevance of c-command and principle (10') also allows us to address another puzzle in the interpretive properties of *pro*. Consider the coreference

possibilities in "scene setting" environments, such as those discussed by Reinhart (1983):

(14)   In this picture of Mary, she looks sick

In the equivalent in Italian, if we want to express the interpretation in which Mary looks sick, the overt pronominal *lei* must be used:

(15)   In questa foto di Maria$_i$, *pro_i / lei$_i$ sembra malata
        'In this picture of Maria, pro / she looks sick'

Pro is of course possible, but it must refer to another salient referent, not to Maria.[3] Why is it so? Clearly, here neither (10) nor (10') apply, because Maria is not an aboutness subject, nor does it c-command the main subject. So, Maria is not the expected antecedent of the pronominal subject, therefore *pro* cannot be used, and the overt pronominal must be used to express coreference with Maria.

    Back to the main line of argumentation, we can also notice that *pro* naturally picks out a preverbal subject, but not a postverbal subject in a language like Italian, which permits both positions:

(16)   a.   Quando Gianni$_i$ ha telefonato, pro_i era ubriaco
              'When Gianni telephoned, pro was drunk'
        b. * Quando ha telefonato Gianni$_i$, pro_i era ubriaco
              'When telephoned Gianni, pro was drunk'

Coreference in (16)a is fine because Gianni is the subject of predication here, as in the cases considered previously. A postverbal subject is not in the aboutness subject position, hence it is expected that it may not be the antecedent of *pro*, as in (16)b. But there may be an additional reason excluding that dependency. The postverbal subject position is focal (Belletti 2004, 2009), therefore the relevant construal may be excluded by whatever property excludes the construal of a pronoun with a non c-commanding focal element. In fact, as Calabrese (1986) observes, (16)b remains deviant also when the main pronominal subject is overt:

(16)   b'. *? Quando ha telefonato Gianni$_i$, lui$_i$ era ubriaco
               'When telephoned Gianni, he was drunk'

---

3 For instance, if Francesca is a (hanging) topic, a *pro* coreferential to Francesca is possible:
   (i)   (A proposito di) Francesca$_k$, in questa foto di Maria$_i$, pro_k sembra malata
          '(As for) Francesca, in this picture of Maria, she looks sick'

Here, *di Maria* would naturally designate the possessor of the picture, rather than the person reproduced in it.

The persistence of the ill-formedness in (16)b' shows that a more demanding principle is violated here than the requirements of the system in (9)–(10).

Calabrese (1986, fn. 3) reports an observation due to Lidia Lonzi according to which construal with a postverbal subject considerably improves when *pro* is in an adverbial clause following the postverbal subject:

(17)   Ha parlato Carlo$_i$, dopo che *pro$_i$* è arrivato
       'Spoke Carlo (subj), after pro arrived'

Apart from the necessity of a strong pause between the postverbal subject and the adverbial clause, which Calabrese notices, in this case *pro* may be c-commanded by the focal subject: this would come about if the focal subject is locally moved to an IP-internal focus position (as in Belletti 2004), and this position is high enough to c-command clause-final adjunct clauses. This configuration would make *pro* bindable by the focal element, and would make the case analogous to these in which the antecedent c-commands *pro*, such as (12), etc. The c-command option is clearly excluded in cases like (16)b, as the postverbal focal subject is in an embedded adverbial. C-command is also excluded in (16)a, but here *pro* has the other option of picking out a local aboutness subject, an option not available in (16)b.

5 Aboutness is independent from new-given information

In general, the "aboutness" argument tends to be associated with given information, whereas a predicate-internal argument is not. So, there could be a confound between aboutness and givenness: could it be that what *pro* really picks out is given information?

In fact, the link between the aboutness argument and givenness is not necessary. When the subject is not given, for instance in an all-new context, it still functions as the antecedent of *pro*. Here we cannot use the core context discussed by Calabrese, with a preposed adverbial and a main clause, which would often not sound natural in an all-new context: the selective effect on the choice of the antecedent of *pro* is also found, as Calabrese (1986) points out, in the immediately following clause in discourse. Consider for instance a context in which a speaker asks question (18)Q, thus setting up a natural context for a possible all-new sentence (for instance, a speaker may utter (18)Q when he perceives that some action is going on without knowing what kind of event is happening and

who the participants are), and the interlocutor answers with sentence (18)A, and then continues with (18)A':

(18) Q: Che cosa è successo?
          'What happened?'
      A: Un ragazzo$_i$ ha buttato a terra un vecchio$_k$
          'A boy threw an old man to the ground'
      A': ... poi *pro*$_{i, *k}$ ha cominciato a urlare
           '...then pro started to scream'

In (18)A, both a boy and an old man are new information, and still *pro* in the immediately following sentence is restricted to pick out the subject. So, the test is not sensitive to the informational property of givenness, but to the structural position of the antecedent.[4] Again, if the subject of (18)A' is the overt pronoun *lui*, the coreference option shifts, and the natural interpretation is that the old man started screaming.

We have parallel effects when the all-new sentence is passivized: the interlocutor can choose to answer question (19)Q with a passive sentence, as in (19) A; again, in the following sentence (19)A', *pro* picks out the surface subject of predication, in this case the patient of the passive sentence:[5]

(19) Q: Che cosa è successo?
          'What happened?'
      A: Un vecchio$_k$ è stato buttato a terra da un ragazzo$_i$
          'An old man was thrown to the ground by a boy'

---

4 Giuliano Bocci raises the question of whether in this context the effect is restricted to *pro*, or it also extends to other weak pronominal forms. In fact, it seems to me that, here as before, an object clitic can pick out both arguments. For instance two continuations like the following
  (i)   ... poi, lo ho visto sanguinante
        '... then, I saw him bleeding'
  (ii) ... poi, lo ho visto ammanettato
        '...then, I saw him handcuffed'

seem to me to be both possible (pragmatically, (i) is natural with an old man as antecedent, (ii) with a boy as antecedent)

5 Cardinaletti (2004) argued that (at least) two subject positions are to be specified in the high IP field. Building on that, Bianchi & Chesi (2013) argued that the higher and the lower positions correspond to two distinct interpretations, akin to Kuroda's (1973) "categorical" and "thetic" judgments. Given these ideas, the question arises of whether the Calabrese effect involves both subject positions, or is restricted to just one. I will not fully addressed this issue here. Let me just notice that such example as (18), (19) in all-new contexts should instantiate Kuroda's thetic judgments, and they do give rise to the Calabrese effect. If thetic judgments involve the lower subject position, the conclusion seems to be that the Calabrese effect is triggered by this position as well.

A'   ... poi $pro_{k, *i}$ ha cominciato a urlare
'...then pro started to scream'

That the aboutness property is independent from informational structure is also underscored by the reciprocal case wrt (18)–(19). In fact, we have an identifiable effect also in contexts in which both arguments are given, like the following:

(20)   Q: Come mai Gianni e Piero sono così arrabbiati?
             'Why are Gianni and Piero so angry?
       A: Beh, è successo che $Gianni_i$ ha insultato $Piero_k$ davanti a tutti
             'Well, it happened that Gianni insulted Piero in front of everyone'
       A'   ... e subito dopo $pro_{i, *k}$ ha lasciato la riunione
             '...and immediately after pro left the meeting'

One could also answer the questions in (20)Q (=(21)Q) with a passive sentence, as in (21)A; in that case the coreference possibilities are reversed, and *pro* in the following sentence (21)A' must refer to the surface subject of passive, the patient:

(21)   Q: Come mai Gianni e Piero sono così arrabbiati?
             'Why are Gianni and Piero so angry?
       A: Beh, è successo che $Piero_k$ è stato insultato da $Gianni_i$ davanti a tutti
             'Well, it happened that Piero was insulted by Gianni in front of everyone'
       A'   ... e subito dopo $pro_{*i, k}$ ha lasciato la riunione
             '...and immediately after pro left the meeting'

In conclusion, the Calabrese effect is independent from the new or given character of the relevant referent: what the effect is sensitive to is the aboutness property, in accordance with (10), and it holds whether the aboutness subject qualifying as the expected antecedent is new or given information.

6 Subject vs Topic

Reference to "aboutness" stresses the similarity between subjects and topics. In this section and in section 7, I would like to discuss similarities and differences between the two notions, and try to address the question of whether the Calabrese effect singles out subjects, or concerns both.

A classical line of research (e.g., Li & Thompson 1976) addresses similarities and differences between subjects and topics. Let us focus on this issue on the basis of the analysis developed so far. Both subject and topic involve aboutness.

In a subject – predicate configuration, the predicate says something about the subject; in a topic – comment configuration, the comment says something about the topic (Reinhart 1981).

The appropriateness conditions for the use of topics are stricter, though: in "what happened?" contexts, a subject can be felicitously used, as in (18)A, (19)A, but a topic cannot:

(22) Q: Che cosa è successo?
         'What happened?'
      A: # Un vecchio, un ragazzo lo ha buttato a terra
         'An old man, a boy threw him to the ground'

In this context, the topic structure is not felicitous either when the subject is preverbal, as in (22)A, or when it is postverbal, as in the following A':

      A': # Un vecchio, lo ha buttato a terra un ragazzo
          'An old man, threw him to the ground a boy(subj)'

Could the difference between subject and topic be that topics cannot be indefinite? Indeed, various languages put a requirement of definiteness and/or of specificity on topics, and certain cases of topics seem to require definiteness. This is the case, for instance, for hanging topics introduced by prefixes such as *per quanto riguarda X, quanto a X*, etc. (as far as X is concerned, as for X):

(23) Quanto a Maria/questa ragazza/alla ragazza/*una ragazza, parlerò presto con lei
      'As for Maria/this girl/the girl/*a girl, I will talk soon with her'

Nevertheless, in Italian it is not impossible to have an indefinite topic in the clitic left dislocation construction, even a non-specific indefinite, as the following felicitous exchange shows:

(24) Q: Perchè la direttrice è così preoccupata per il dipartimento?
         'Why is the head of department so worries about the dept?'
      A: Una segretaria che sappia tenere la contabilità, non riesce a trovarla
         'A secretary who can(subjunctive) keep the accounting, she cannot find'

The topic is indefinite and non-specific here, as is shown by the subjunctive mood of the verb in the relative clause. Still, the topic structure is felicitous, as the topic directly connects to the context (the accounting of the department).

Examples (22)A, A' can also be made felicitous, e.g., in the following conversational exchange:

(25)   Q: Che cosa è successo? E' vero delle persone anziane sono state portate
            'What happened? Is it true that some elderly people have been taken to
            all'ospedale?
            the hospital?'

         A: Io ho visto solo che un vecchio, un ragazzo lo ha buttato a terra
            'I have only seen that an old man, a boy threw him to the ground'

         A':   Io ho visto solo che un vecchio, lo ha buttato a terra un ragazzo
            'I have only seen that an old man, threw him to the ground a boy(subj)'

Question (25)Q introduces a group of elderly people, and (25)A, A' take up a member of this group (albeit indefinite) as a topic.

In order to have a felicitous topic, some kind of connection with a contextually given set appears to be necessary. An indefinite topic is also possible in Italian, when an unknown individual is linked to a contextually given set (e.g., an old man is connected to the set of elderly people evoked in the immediately previous discourse in (25)). This distinguishes the case in which a topic is possible and impossible, as in (22)A, A'. This link to a contextually given set is clearly reminiscent of D-linking (Pesetsky 1987) (but of relevance may also be Enç's 1984 notion of "partitivity").

This component of connection to the discourse context in the licensing of topics is particularly clear in certain cases (here my presentation is based on Rizzi 2005). Consider a context in which a father is checking on his son's preparation of an exam. If the father says sentence (26), the son could reply with (27) or (27'):

(26)   Father: Oggi non hai fatto niente per preparare l'esame...
                 'Today, you did nothing to prepare the exam...'

(27)   Son: Beh, ho letto un libro...
            'Well, I read a book...?

(27')   Son: Beh, un libro l'ho letto...
            'Well, a book, I read it...'

Both replies are felicitous, but they are not interpretively equivalent. Reply (27') involves a clitic left dislocation of the indefinite object *un libro* (a book): the sentence implies that this book is part of the program for the preparation of the exam: i.e., "a book among those required for preparing the exam, I read it". No such implication is associated with reply (27), in which *a book* remains in object position. The sentence can be very naturally interpreted as "ok, I didn't do anything for the preparation of the exam, but I spent time in another worthwhile activity, reading a book", with the book in question completely disconnected

from the program of the exam. The special interpretation associated with (27')
thus suggests that topics must be licensed by this (explicit or implicit) partitive
connection to a set established in discourse, here the set of documents necessary
for the preparation of the exam.

We may express the interpretive rules triggered by Top at the interface as
follows:

(28) **Top**: a. Interpret the Spec as a D-linked argument about which a comment
is made
b. Interpret the complement as the comment about the Spec.

Beninca' & Poletto (2004), Frascarelli & Hinterhoeltz (2007), Bianchi & Frascarelli
(2011) propose a finer typology of topics: aboutness shift, contrastive, and famil-
iarity topics. I think aboutness is a common feature of all kinds of topics (includ-
ing aboutness shift: "so far we have been talking about topic X, now I want to
shift to a different topic Y"), and D-linking, or the relevant notion of partitivity,
may also be. So, one can think of (28) as the common interpretive core that differ-
ent kinds of topics share.

The aboutness Subject, Calabrese's notion of Thema, shares the aboutness
property of topics, but does not require D-linking. So, a non-D-linked phrase as
un ragazzo, un vecchio in contexts like (18), (19), (22) can be a subject, but cannot
be a topic:

(29) **Subj**: a. Interpret the Spec as the argument which the predicate is about.
b: Interpret the complement as the predicate

Does the Calabrese effect distinguish subject and topic? Calabrese discusses
examples like (30) in which *pro* is unable to pick out the dative topic *a Gianni* of
a preceding sentence, whereas it can take the subject *Carla* as antecedent (his
example (28)):

(30) Poiché a Gianni$_k$ Carla$_i$ gli ha dato un bacio, *pro*$_{i, *k}$ è felice
'Because to Gianni Carla to-him gave a kiss, pro is happy'

I find the judgment clear in the case in which the topic is a dative, as in (30).
When the topic is a direct object, my judgment becomes less sharp. Consider the
following:

(31) a. Poiché Mario, Carlo lo ha severamente criticato, *pro* era piuttosto
'Because Mario, Carlo severely him criticized, pro was rather
imbarazzato
embarrassed'

b. Quando Francesco, Gianni lo ha presentato a Piero, *pro* era molto contento
'When Francesco, Gianni him introduced to Piero, pro vas very happy'

These judgments are difficult, perhaps I still have a preference for binding of *pro* by the previous aboutness subject, but binding by the topic does not seem to me to be excluded. If binding by a topic is disambiguated by gender agreement of the predicate, i.e., with subject and topic mismatching in gender, binding by the topic sounds ok to me:

(32) a. Poiché Mario$_i$, Carla lo ha severamente criticato, *pro*$_i$ era piuttosto
          'Because Mario, Carla severely him criticized, pro was rather
          imbarazzato
          embarrassed'

      b. Quando Francesco$_k$, Giovanna lo ha presentato a Piero, *pro*$_k$ era molto
          'When Francesco, Giovanna him introduced to Piero, pro vas very
          contento
          happy'

Notice that, when the object is not topicalized but *in situ* in object position, gender mismatch does not seem to help:

(33) a. Poiché Carla ha severamente criticato Mario$_i$, *pro*$_{*i}$ era piuttosto
          'Because Carla severely criticized Mario, pro was rather
          imbarazzato
          embarrassed'

      b. Quando Giovanna ha presentato Francesco$_k$ a Piero, *pro*$_{*k}$ era molto contento
          'When Giovanna introduced Francesco to Piero, pro vas very happy'

(*pro* in (33)a–b can obviously refer to another male referent, e.g. Antonio which was mentioned in previous discourse; coreference with the object becomes natural in (33) if the overt pronominal form *lui* is used, as before).

Given the contrast between (32) and (33), topicality of an object seems to at least improve the possibility of coreference with a successive *pro*. If this is so, it would seem that the Calabrese effect is sensitive to aboutness, a property that subjects and topics have in common: we could then restate (10) to the effect that a pronominal subject is expected to have the reference of a +aboutness position (in this view, the degraded character of (30) would be due to some other property, presumably connected to the dative case of that example). Given the less than straightforward nature of the relevant judgments, I will leave the question open of whether (10) is indeed to be revised along these lines, or it selectively picks out subjects, as in Calabrese's original formulation.

7 More evidence that preverbal subjects and topics are distinct positions in Null Subject Languages

In a non-null subject language like French, whether a subject is in subject or topic position is straightforwardly shown by the presence of a resumptive subject clitic in the latter case:

(34) a. Jean a rencontré Marie
          'Jean met Marie'
      b. Jean, il a rencontré Marie
          'Jean, he met Marie'

In (34)a, *Jean* is in subject position, where it is not clitic-resumed, whereas in (34)b it is in topic position, and it is obligatorily clitic resumed in the French clitic left dislocation construction. The obligatoriness of clitic resumption is illustrated by a case like the following:

(35) a.   Jean, ton livre, il le lira demain
           'Jean, your book, he it will read tomorrow'
      b. * Jean, ton livre,＿le lira demain
           'Jean, your book,＿it will read tomorrow'

In (35), Jean necessarily is in a left-peripheral topic position, because it is followed by an object topic; and here clitic resumption is obligatory, as the ill-formedness of (35)b shows.

In a Null Subject Language like Italian, the evidence distinguishing preverbal subjects and subject topics is less straightforward because the resumptive pronoun corresponding the a subject topic is null, so a string like (36) is in principle structurally ambiguous between the two representations of (37):

(36) Gianni ha incontrato Maria
      'Gianni met Maria'

(37) a. [$_{IP}$ Gianni ha incontrato Maria]
          'Gianni has met Maria'
      b. [$_{TopP}$ Gianni Top [$_{IP}$ *pro* ha incontrato Maria]
          'Gianni has met Maria'

One could observe that the two representations of (37) would correspond to two distinct intonational contours, but intonation does not seem to offer completely reliable cues to distinguish between preverbal subjects and topics (Bocci 2013).

This state of affairs has sometimes led to the hypothesis that perhaps Null Subject Languages make a more drastic choice in these cases: perhaps, the preverbal subject position is always filled by *pro* in Null Subject Languages, and "preverbal subjects" are in fact always in the left peripheral topic position. I.e., according to this view, the representation of (36) always is (37)b.

We have already seen one clear piece of evidence against this view. In all new contexts, preverbal subjects are legitimate, but topics are not (see (22)), a pattern which I will now illustrate with another example. In a "what happened" context like (38)Q, SV answers are possible, both in active and passive, as in(38)A, A', but topics are not allowed (as in (38)B, B'). So, (38)A, A' (and previous examples like (18), (19)), illustrate a case of a preverbal subject which is not a topic.

(38) Q   Cosa è successo?
           'What happened?'
     A   Un camion ha tamponato un autobus
           'A truck bumped into a bus'
     A'   Un autobus è stato tamponato da un camion
           'A bus was bumped into by a truck'
     B # Un autobus, un camion lo ha tamponato
           'A bus, a truck it bumped into'
     B' # Un autobus, lo ha tamponato un camion
           'A bus (obj), it bumped into a truck (subj)'

Another kind of evidence leading to the same conclusion, discussed in Rizzi (1985), is provided by the fact that certain quantified expressions, particularly when the quantifier is bare, cannot naturally function as topics:

(39) a.   Non ho incontrato nessuno
             'I met noone'
     b. * Nessuno, lo ho incontrato
             'Noone, I met him'

(40) a.   Ho capito tutto
             'I understood everything'
     b. * Tutto, lo ho capito
             'Everything, I understood it

And still these elements can be found in preverbal subject position:

(41) Nessuno ha aiutato Maria
       'Nobody helped Maria'

(42)   Tutto è successo nella notte
        'Everything happened during the night'

So, it appears that preverbal subjects are not necessarily topics in Null Subject Languages: they can be topics in the left periphery and bind *pro* in IP internal position if their interpretive properties and the discourse conditions are consistent with topicality, but they can also be expressed in the IP internal subject position, much as in non Null Subject Languages.

We can now trace back the deviance of (39)b, (40)b to the interpretive procedure associated to Top in (28). This makes a prediction. If indeed (39)b, (40)b are ruled out because the D-linking element necessarily involved in topicality is missing there, one would expect that by making such elements D-linked the possibility of topicalizing quantified expressions should improve. In fact the following are sharply improved, compared to (39)b, (40)b:

(43)   Nessuno di loro, lo ho incontrato alla festa
        'None of them I him met at the party'

(44)   Tutto questo, lo ho capito
        'All this, I it understood'

The partitive expression *di loro* (of them) and the demonstrative *questo* (this) link the quantified expressions to the discourse context, and make the use of such expressions as topics felicitous. No connection to the context of this sort is needed for quantified expressions in subject position in (41), (42), which follows from the assumed difference in interpretive properties between subject and topic positions.

Conclusions

The preverbal subject position is the typical final landing site of A-movement. Recent approaches to the "halting problem" identify halting positions of phrasal movement chains as criterial positions. If this is correct, we are led to assume a "subject criterion" to be satisfied in the higher part of the IP zone. The natural interpretive property associated to this criterial position is "aboutness": it identifies the argument "about which" the event is presented, with a detectable shift, for instance, from agent to patient aboutness in active-passive pairs. Andrea Calabrese discovered a property which is clearly sensitive to aboutness in this sense: a *pro* subject in Null Subject Languages picks out the aboutness subject of an immediately preceding sentence in a local environment (Calabrese 1986).

The Calabrese effect thus shows that subjecthood is crucial for the structuring of subsequent discourse and anaphora resolution, a scope-discourse property that we may expect to hold for a criterial position.

Topics share with subjects the aboutness property, in the obvious sense that a comment states something about the topic. Nevertheless topics are more demanding than subjects, in that they can be used felicitously only if some kind of connection to the previous discourse (D-linking, or partitivity) can be established. Preverbal subjects are thus possible in "what happened" contexts, but topics are not. Even in Null Subject Languages, in which an identification of preverbal subjects and topics has often been proposed, "what happened" contexts clearly differentiate the two cases. This kind of evidence converges with evidence based on the distributional properties of bare quantified expressions, possible in subject but not in topic positions, a distributional difference which also is ultimately deducible from the different interpretive procedures for subjects and topics.

References

Belletti, Adriana. 2004. Aspects of the low IP area. In Luigi Rizzi (ed.), *The structure of CP and IP: the cartography of syntactic structures*, vol. 2. New York: Oxford University Press.

Belletti, Adriana. 2009. *Structures and strategies*. London and New York: Routledge.

Belletti Adriana, Elisa Bennati, & Antonella Sorace. 2007. Theoretical and developmental issues in the syntax of subjects: evidence from near native Italian. *Natural Language and Linguistic Theory* 25(4). 657–689.

Benincà, Paola and Cecilia Poletto. 2004. Topic, focus and V2: Defining the CP sublayers. In Luigi Rizzi (ed.), *The structure of CP and IP*, 52–75. New York: Oxford University Press.

Bianchi, Valentina & Cristiano Chesi. 2014. Subject islands, reconstruction, and the flow of the computation. *Linguistic Inquiry* 45(4). 525–569.

Bianchi, Valentina & Mara Frascarelli. 2010. Is topic a root phenomenon? *Iberia* 2. 43–48.

Bocci, Giuliano. 2013. *The Syntax–Prosody Interface: a cartographic perspective with evidence from Italian*. Amsterdam: John Benjamins.

Bošković, Želiko (ed.). 2016. *The Linguistic Review* 33 (1). Special Issue on Labeling.

Calabrese, Andrea. 1982. Alcune ipotesi sulla struttura informazionale della frase in italiano e sul suo rapporto con la struttura fonologica. *Rivista di grammatica generative* 7.

Calabrese, Andrea. 1986. Some properties of the Italian pronominal system: An analysis based on the notion of thema as subject of predication. In H. Stammerjohann (ed.), *Tema-Rema in Italiano*, 25–36. Tübingen: Gunter Narr Verlag.

Calabrese, Andrea. 1992. Some remarks on focus and logical structures in Italian. *Harvard Working Papers in Linguistics* I. 19–27.

Cardinaletti, Anna. 2004. Towards a cartography of subject positions. In Luigi Rizzi (ed.), *The structure of CP and IP*, 115–165. New York: Oxford University Press.

Chomsky, Noam. 1973. Conditions on transformations. In S. Anderson & P. Kiparsky (eds.), *A Festschrift for Morris Halle*. New York: Holt Rinehart and Winston.

Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.

Chomsky, Noam. 2013. Problems of projection. *Lingua* 130. Special Issue, Core ideas and results in syntax, 33–49.

Chomsky, Noam. 2015. Problems of projection: extensions. In E. Di Domenico, C. Hamann & S. Matteini (eds.), *Structures, strategies and beyond – Studies in honour of Adriana Belletti*, 3–16. Amsterdam: John Benjamins.

Frascarelli, Mara & Roland Hinterhölzl. 2007. Types of topics in German and Italian. In Susanne Winkler and Kerstin Schwabe (eds.), *On information structure, meaning and form*, 87–116. Amsterdam: John Benjamins.

Givón, Talmy. 1983. *Topic continuity in discourse: a quantitative cross language study*. (Typological Studies in Language 3). Amsterdam: John Benjamins.

Kuroda, Sige-Yuki. 1972. The categorical and thetic judgement: evidence from Japanese syntax. *Foundations of Language* 9. 153–185.

Li, Charles N. and Sandra A. Thompson. 1976. Subject and topic: a new typology of language. In Charles N. Li (ed.), *Subject and topic*, 457–461. New York: Academic Press.

Pesetsky, David. 1987. Wh in situ: movement and unselective binding. In E. Reuland & A ter Meulen (eds.), *The representation of (In)definiteness*, 98–129. Cambridge, MA: MIT Press.

Reinhart, Tanya. 1981. Pragmatics and linguistics: An analysis of sentence topics in pragmatics and philosophy I. *Philosophica anc Studia Philosophica Gandensia Gent* 27. 53–94.

Reinhart, Tanya. 1983. *Anaphora and semantic interpretation*. Chicago: The University of Chicago Press.

Rizzi, Luigi. 1986. On the status of subject clitics in Romance. In Jaeggli & Corvalan (eds.), *Studies in Romance linguistics*, 391–419. Dordrecht: Foris Publications.

Rizzi, Luigi. 1997. The fine structure of the left periphery. In L. Haegeman (ed.), *Elements of grammar*, 281–337. Dordrecht: Kluwer.

Rizzi, Luigi. 2005. On some properties of subjects and topics. In Laura Brugé, Giuliana Giusti, Nicola Munaro, Walter Schweikert & Giuseppina Turano (eds.), *Proceedings of the XXX Incontro di Grammatica Generativa*, 203–224. Venezia: Cafoscarina.

Rizzi, Luigi. 2006. On the form of chains: criterial positions and ECP effects. In L. Cheng, N. Corver (eds.), *On wh movement*, 97–133. Cambridge, MA: MIT Press.

Rizzi, Luigi. 2015a. Cartography, criteria, and labeling. In U. Shlonsky (ed.), *Beyond functional sequence – the cartography of syntactic structures, vol. 10,* 314–338. New York: Oxford University Press.

Rizzi, Luigi. 2015b. Notes on labeling and subject positions. In E. Di Domenico, C. Hamann & S. Matteini (eds.), *Structures, strategies and beyond – Studies in honour of Adriana Belletti,* 17–46. Amsterdam: John Benjamins.

Rizzi, Luigi. 2016. Labeling, maximality, and the head – phrase distinction. *The Linguistic Review* 33(1). 103–127.

Rizzi, Luigi & Shlonsky, Ur. 2007. Strategies of subject extraction. In Hans-Martin Gärtner & Uli Sauerland (eds.), *Interfaces + recursion = language? Chomsky's minimalism and the view from syntax-semantics*, 115–160. Berlin: Mouton de Gruyter.

Van Urk, Coppe. 2015. *A uniform syntax for phrasal movement: A Dinka Bor case study*. Cambridge, MA: MIT dissertation.