Chapter 12

Rethinking the ECP: Subject–object asymmetries as freezing effects

Luigi Rizzi
Collège de France

The ECP had a major explanatory role in GB syntax. Conceptual and technical difficulties with the principle diverted the focus of theoretical attention from core ECP effects in minimalism. Nevertheless, the empirical motivation for such effects remains robust across languages. In this article, I would like to rethink core ECP effects such as subject–object asymmetries in extraction contexts in terms of a different theoretical apparatus which emerged in recent years in connection with cartographic studies. Criterial positions determine freezing effects. If there is a subject criterion, subjects will undergo such effects, and will be unmovable, unless special devices are used by the language. We observe that-trace effects with subjects but not with objects because there is no general object criterion. This alternative theoretical apparatus can be shown to be empirically advantageous with respect to the ECP approach in connection with a number of phenomena discussed in the classical ECP literature.

1 The classical ECP approach

The empty category principle (ECP) played a major explanatory role in government-and-binding (GB) analyses. First and foremost, it captured different kinds of subject–object asymmetries in extraction contexts: all other things being equal, subjects are harder to extract from embedded domains than objects (or other complements). The classical illustration is the that-trace effect. An object is extractable from an embedded declarative introduced by that, but a subject is not:

(1) a. * Who do you think [ that [ ___ will come ]]?
   b. Who do you think [ that [ Mary will meet ___ ]]?

Judgments gathered with controlled methods have confirmed such asymmetries, while revealing new facets of the phenomenon.\textsuperscript{1} There are factors of empirical complexity, though: certain varieties of English admit (1a) as acceptable, so that in such varieties the asymmetry tends to disappear (Sobin 2002; Schippers 2012); nevertheless, the constraints on extractability are not simply subjected to arbitrary variation: for instance, the asymmetry reappears, also for speakers who accept (1a), in other contexts, such as the extraction from indirect questions (here the contrast is in terms of relative acceptability, as extraction from the weak island is always degraded to some extent):

\begin{enumerate}
\item a. *Who do you wonder if ___ will come?
\item b. ??Who do you wonder if Mary will meet ___?
\end{enumerate}

In other languages, things are even sharper. Subject extraction in (3a) appears to be systematically excluded in French, while object extraction in (3b) is possible (Berthelot 2017):

\begin{enumerate}
\item French
\item a. *Qui penses-tu que ___ va venir?
\quad ‘Who do you think that will come?’
\item b. Qui penses-tu que Marie va rencontrer ___ ?
\quad ‘Who do you think that Marie will meet?’
\end{enumerate}

So, the asymmetries are a real, robustly attested phenomenon. The ECP tried to capture the asymmetries by appealing to independent properties differentiating subjects and complements. According to the classical approach of Chomsky (1981), traces must be lexically governed (or antecedent-governed, an option that I do not discuss here). The object is governed by a lexical element, the verb, while the subject is governed by a functional head, the node Infl, or T, which is not sufficient to satisfy the requirement. So, the asymmetry follows from the nature of the governing element.

This analysis was extremely influential and gave rise to an important literature both on the cross-linguistic scope of the phenomenon, the ways of circumventing it (e.g. via complementizer deletion in English), the exact format of the principle, etc. (see, e.g., Pesetsky 1982; Kayne 1984; Rizzi 1982; 1990, a.o.) In spite of its empirical success and its capacity to generate important syntactic research, the ECP approach was abandoned in Minimalism.

\textsuperscript{1}For instance also in case of object extraction the optimal case is from a clause not introduced by an overt complementizer, but a clear contrast with subject extraction persists: Schippers (2012).
I think the main problem which led to this step is conceptual: Minimalism permits a very limited variety of UG principles: principles operating at the interfaces with sound and meaning, and somehow enforced by the needs of the interface systems (e.g., linearization at the PF side, the theta criterion at the LF side, etc.), and principles of optimal computation, operating on the computing machine (including principles of economy, locality, labeling, etc.). The ECP does not naturally fit into any of these categories, so it has no natural place in the minimalist universe.

There were also technical problems, due to the reliance of the ECP on government, a structural relation not assumed in Minimalism. Personally, I never found such considerations compelling: government is minimal c-command, i.e., c-command constrained by locality, and Minimalism must assume both some form of c-command (perhaps derivatively from the extension condition or no tampering) and some form of minimality, so that the ingredients for government are there, even if a primitive government relation is not postulated. But, even if the technical argument may be unconvincing for these reasons, the conceptual argument remains compelling. So, research on the asymmetry was somehow demoted from center stage in Minimalism.

Nevertheless, the facts are clear, and cross-linguistically robust. True, some languages do not manifest the asymmetry, so that the phenomenon has sometimes been qualified as “language specific”, and, as such, not bearing on UG principles. But this kind of reasoning is highly questionable. On the one hand, systematic exceptions to that-trace effects have turned out to be amenable to independent principled explanations, such as the systematic absence of the asymmetries in null subject languages (Rizzi 1982). On the other hand, the cross-linguistic distribution is clearly constrained: we don’t seem to find clear cases of the “mirror image” of English or French, i.e., a language freely allowing subject extraction across an overt complementizer and banning object extraction.

In classical discussions of such issues, poverty of stimulus considerations were typically invoked to support the necessity of a principled explanation. How does the learner of (the relevant variety of) English, or French, come to know that (1a), (2a), (3a) are excluded? Why don’t all language learners analogically generalize from cases of extraction they hear, (such as 1b), to all cases of extraction, assuming no asymmetries? An anonymous reviewer observes that some qualification is needed here because through statistical learning techniques it may be possible to infer the ungrammaticality of a structure such as (1a) from its non-occurrence. The point is well-taken, even though one should make sure that such techniques can be selective enough, i.e., do not generalize from cases like (1b) to cases like (1a), in the absence of any principled guidance. More importantly, a technical
approach to these problems based on statistical learning would remain too “lo-
cal”: why should the asymmetries be systematically found across languages, and
always in the same direction? Somehow, the systematic higher difficulty with
subject extraction, robustly attested language after language, must come from
some internal pressure and be connected to a principled reason, exactly what
the ECP approach assumed.

These considerations pave the way for the search of a principled alternative
to the ECP to capture the asymmetries.

2 Criterial freezing and the subject criterion

According to the criterial approach to scope-discourse semantics, the initial pe-
riphery of the clause is populated of functional heads such as Q, Top, Foc, etc.,
which attract a phrase with matching feature, creating criterial (Spec-head) con-
figurations, and guide the interpretation of such criterial configurations at the
interfaces with sound and meaning (Rizzi 1997).

One salient property of such criterial configurations is that the attracted ele-
ment is frozen in the criterial position, i.e., it cannot be attracted to a higher posi-
tion. The canonical example is the case of a wh-phrase satisfying the Q-criterion
in an embedded interrogative, selected by a verb like wonder. In such cases, the
wh-element cannot be moved further:

(4) Lasnik & Saito (1992), Bošković (2008)
   a. Bill wonders \[ \textbf{which book} \ Q \ [ \text{she read } ___ ] \]
   b. * \textbf{Which book} \ Q \ does Bill wonder \[ ___ \ Q \ [ \text{she read } ___ ] \]?

While obvious options come to mind to rule out (4b) (one could invoke inter-
face problems with the derived representations, or an “inactivation” analysis à
la Bošković 2008), more complex cases discussed in Rizzi (2006; 2011) and much
subsequent work suggest that the problem is deeper. So, a descriptive principle
like the following seems to hold:

(5) Criterial freezing: An XP meeting a criterion is frozen in place.

In fact the phrase meeting a criterion is not completely frozen: if the phrase
is complex, part of it can be subextracted. E.g., taking Italian (6a) as baseline,
focalization of the PP \textit{di Piero} is possible, with subextraction and clefting:
Rethinking the ECP: Subject–object asymmetries as freezing effects

(6) Italian
   a. Non è chiaro [[ quanti libri di Piero ] Q siano stati censurati ]
      ‘It isn’t clear how many book by Piero Q have been censored’
   b. E’ di Piero che non è chiaro [[ quanti libri ___ ] Q siano stati censurati] (non di Gianni)
      ‘It is by Piero that it is not clear how many books Q have been censored, not by Gianni’

The formulation in (5) should be refined to permit this kind of subextraction. In fact, the element of the specifier which is frozen is the carrier of the criterial feature, the criterial goal, if we assume that the criterial head enters into a probe–goal relation with the attracted phrase (Chomsky 2000). So, (5) should be refined as follows:

(5′) Criterial freezing: In a criterial configuration, the criterial goal is frozen in place.

See Chomsky (2013; 2015), Rizzi (2015a; 2015b; 2016) for attempts to derive the effects of (5′) from the labeling algorithm. I will not address this important point here, and will just assume a descriptive formulation like (5′).

Criterial freezing separates specifier positions targeted by movement into two classes: H A L T I N G positions, and T R A N S I T I N G positions. The criterial positions are halting positions, where movement stops; transiting positions are specifier positions from which movement can (and in fact must) continue, for instance the C-system of a verb like think, which can function as an escape-hatch for a wh-phrase, but not as the final landing site of wh-movement.

If we now turn to the system of A-movement, the typical halting position of A-movement chains is the subject position of finite clauses (as opposed to transiting A-positions, such as the subject positions of raising clauses, participial constructions, etc.). If halting positions are equated to criterial positions, these considerations lead us to assuming a criterial position for A-movement, a subject criterion (Rizzi 2006, and much subsequent work).

Criteria typically go with scope-discourse interpretive effects, such as the topic–comment or focus–presupposition articulations. So, what could be an analogous interpretive effect for subjects? Interpretively, the subject position designates the referent “about which” the event is presented. Active-passive pairs clearly differ in this aboutness property. The following sentences are both appropriate in “all new” contexts, e.g., as answers to questions like “What happened?”, or, with a narrower contextualization, “How did the battle start?”:
Both sentences felicitously depict an attacking event in the given context, but (7a) depicts the event as being about a plane, the agent, and (7b) depicts it as being about a cruiser, the patient. The choice of the aboutness subject has consequences for discourse organization. For instance, as Calabrese (1986) pointed out, the choice affects anaphora resolution in null subject languages, in that a pro subject in the following sentence picks out the aboutness subject. So, if (8) is uttered immediately after (7a), the intended interpretation is that the plane asked for backup; if (8) is uttered after (7b), it’s the cruiser which did (see also Belletti et al. 2007 on this effect):

(8) Italian
    … poi, pro ha chiesto rinforzi
    ‘… then ___ asked for backup’

In much current work initiated in Rizzi 2006, and building on Cardinaletti 2004, I have assumed that a nominal head Subj is an obligatory component of the clausal spine. This head occurs immediately under the lowest head of the complementizer system Fin, hence higher than T, so that we have a partial map of the high part of the IP as follows:

(9)  ... Fin ... Subj[+N] ... T ...

In syntax, Subj[+N] attracts the closest nominal expression to its Spec. At the interface, it triggers an interpretive routine along the following lines: “interpret my Spec as the argument which the predicate is about, and my complement as the predicate”.

3 +N as an attracting feature to the subject position

Why should +N be the relevant feature here? The obvious intuition is that the system needs a nominal expression, capable of referring to an argument, to trigger the appropriate aboutness interpretation. An alternative that comes to mind, perhaps more in line with standard assumptions, is that the attracting feature could be the set of Phi features.
One motivation for assuming +N to be the attractor is provided by the widespread existence of quirky subject constructions, in which a non-nominative nominal expression occupies a subject position (typically with psych-verbs and a few other verbal classes in some languages):

(10) Italian
    A Gianni piacciono queste idee
    ‘To Gianni please these ideas’

In such cases, it is not very plausible that the attracting features would be the Phi set, as the clause initial nominal does not enter into an agreement relation with the inflected verb, whereas if the attracting feature is +N, quirky subjects are expected. In languages like Icelandic, the quirky subject with an inherent case may be a KP, in languages like Italian it could be a KP or a PP, but in any event it plausibly is an “extended projection” the nominal element, in Grimshaw’s (1991) sense, hence accessible to being attracted by a +N attractor.²

The point is relevant in the context of this paper because the hypothesis that the attractor is +N may help explain other subject–object asymmetries originally ascribed to the ECP. One has to do to do with en cliticization in French. As was shown by Ruwet (1972), the clitic en can pronominalize both a PP (in partitive constructions such as [ la première partie [PP de ce roman ] ] ‘the first part of this novel’) and an NP (contained in a larger structure headed by a numeral, such as [ trois [NP romans ]] ‘three novels’):

(11) French
    a. Jean en a publié [ la première partie ___] en 1968 (de ce roman : en = pro-PP)
       ‘Jean of-it published the first part in 1968 (of this novel)’
    b. Jean en a publié [ trois ___ ] en 1968 (romans : en = pro-NP)
       ‘Jean of-them published three in 1968 (novels)’

But if the DP is in subject position, e.g., in the passivized versions of (11), only PP extraction is possible, and NP extraction is barred:

²That the dative experiencer is in subject position, and not a topic, is shown, among other things, by the fact that it does not interfere at all with A-bar extraction, whereas a genuine topic does: Calabrese 1986; Belletti & Rizzi 1988. The special properties of expletives as elements formally satisfying the subject criterion are discussed in Rizzi & Shlonsky (2007).
(12) French (adapted from Ruwet 1972)

a. a [ la première partie ___ ] en a été publiée ___ en 1968
   ‘The first part of-it was published in 1968’

b. * [ Trois ___ ] en ont été publiés ___ en 1968
   ‘Three of-them have been published in 1968’

Why this asymmetry? In Rizzi (1990: 37–38) I proposed an ECP analysis: in object position both traces are lexically governed, by the noun partie in (11a), and by the verb in (11b) (under the definition of government adopted there). In (12a) the trace is still lexically governed by the noun, but in (12b) there is no lexical governor available, hence the structure is excluded.

How can this asymmetry be captured without appealing to the ECP? Under the assumption that the attractor of subject is Subj[+N], the contrast between (12a) and (12b) also follows: in the derivation of (12a), after en has been extracted, the remnant DP still contains a nominal part, and can be attracted; in (12b), the nominal part has been entirely extracted by en cliticization, hence the remnant DP is not extractable any longer (under the copy theory of traces the trace of en is still there, but traces typically are not attractable elements).

It should also be noticed that the asymmetry shown by (12) disappears under A’-movement of the object after en cliticization:

(13) French

a. [ Combien de parties ___ ] il en a publiées ___ en 1968?
   ‘How many parts ___ he of-it published in 1968)?’

b. [ Combien ___ ] il en a publiées ___ en 1968?
   ‘How many he of-them published in 1968?’

Here the +N analysis may have an advantage over the ECP analysis: according to the latter, there is no obvious reason why the lexical government requirement could be lifted in the case of the output of A’-movement, as (13b). By contrast, the alternative involving +N as an attractor captures the contrast between (12b) and (13b): in (13b) the attractor is +Q, and combien clearly carries the Q feature, so the fact that the NP has been extracted is irrelevant, and the remnant can undergo A’-movement.3

3An anonymous reviewer observes that the contrast between (12a) and (12b) is reproduced if the clause is embedded under an “exceptional case marking” verb like laisser (let) in French:

(i) * Il a laissé [ trois ___ ] en être publiées.
   ‘He let three of-them+to+be published’
A somewhat analogous, but also different case of an asymmetry previously connected to the ECP concerns the fact that that deletion cannot affect a moved sentential complement:

(14) a. Bill didn’t say (that) John could win
    b. * (that) John could win, Bill didn’t say ___
    c. * (that) John could win wasn’t said by anyone

Here, contrary to en extraction in French, both A- and A′-movement affect the structure. Stowell (1981) originally observed that the asymmetry in (14) recalls the ECP, and Pesetsky (1995) captured this intuition by assuming that the deleted complementizer is in fact (abstractly) cliticized to the main verb, so that the complementizerless clauses do involve a trace of head movement, arguably in the scope of the ECP.

An alternative to the ECP analysis, still based on the Stowell–Pesetsky insight, could be the following: the clause, in order to move in (14a,b) must be attracted, but its head, the complementizer, has already been attracted and incorporated into the verb; so, if traces are not attractable, the whole clause cannot undergo movement, and must remain in complement position, as in (14a). Notice that this analysis implies that head movement (however it is implemented) is part of narrow syntax, as argued for in Roberts (2010), against the frequently made assumption that head movement is post-syntactic. The difference between that deletion and en cliticization is that in the latter case the head of the construction (the numeral, or possibly a higher abstract determiner) is not affected by cliticization, so that there is no general ban on movement of the whole phrase, but only a selective ban linked to the +N attractor. In case of that deletion, the head of the whole construction has been moved and has become a trace, so that the whole configuration is unmovable.

4 Subject–object asymmetries in extraction contexts

We can now come back to subject–object asymmetries under A′-movement. If criterial configurations are frozen, and there is a subject criterion, nominal elements

(ii) Il a laissé la première partie en être publiée.
    ‘He let [ the first part ___ ] of-it+to+be published’

The reviewer observes that the ECP would not draw the right distinction in this case because the trace of en would be lexically governed by laisser in (i). The contrast follows from the analysis proposed in the text if infinitival clauses of this kind also involve a Subj+[N] head.
which reach Subj will be frozen there. I.e., the attempt of deriving a sentence like (1a) would go through an intermediate representation like (15):

\[(15) \quad \text{You think [ that [ who Subj[+N] will come ___ ]]}
\]

Where who will be frozen and will become inaccessible to further movement. No similar effect arises in the case of object extraction (1b), as there is no object criterion. The asymmetry thus follows from criterial freezing and the subject criterion, which provide an alternative to the classical ECP analysis.

As usual, it is important to look for empirical differences between competing analyses. One class of facts (originally pointed out to me by Paul Hirschbühler) which seems to support the freezing analysis is the following. The wh operator 
*combien* in French can be extracted from an object, or pied-pipe the whole object, as in (16):

\[(16) \quad \text{French}
\]

a. *Combien de personnes veux-tu rencontrer ___?*  
   ‘How many of people do you want to meet?’

b. *Combien veux-tu rencontrer [ ___ de personnes ]?*  
   ‘How many do you want to meet of people?’

Extraction of *combien de NP* from an embedded subject position gives rise to ungrammaticality (as in 17a), but subextraction of *combien* from subject position is only mildly degraded, as in (17b) (Obenauer 1976; Kayne 1984):

\[(17) \quad \text{French}
\]

a. *Combien de personnes veux-tu que [ ___ Subj viennent à ton anniversaire]?*  
   ‘How many people do you want that come to your birthday?’

b. ? *Combien veux-tu que [ [ ___ de personnes ] Subj viennent à ton anniversaire ]?*  
   ‘How many do you want that of people come to your birthday?’

Under the ECP analysis, the ungrammaticality of (17a) is expected, but (17b) would be predicted to be equally ill-formed: if there is no lexical governor for a trace in subject position, a fortiori there should not be a lexical governor for a trace in the specifier of the subject. So, the improvement manifested by (17b) is not expected.

The freezing analysis, by contrast, predicts the ill-formedness of (17a) as a violation of criterial freezing, whereas it makes no claim on (17b), which does not
fall under the scope of formulation (2): only the criterial goal, the nominal part of the DP, is frozen in the criterial configuration with Subj[+N]. The marginality of the example will be linked to other factors constraining extractions from left branches (on such factors, and their interplay with criteria, see Lohndal 2010, Berthelot 2017).

Other cases of special behavior of subjects may be amenable to the same analysis. The complex inversion construction in French (Kayne 1972; Rizzi & Roberts 1989, and subsequent work) involves a wh element (or a null yes/no operator), a subject DP and the inflected verb with an encliticized subject clitic, doubling the subject, as in (18):

(18) French
    Où Jean est-il allé?
    ‘Where John did-he go?’

If the inversion is a reliable cue that I to C (or, in current terms, T to Fin) has occurred, the subject must sit in a special subject position higher than Fin, hence in the left periphery.

Among the many noticeable properties of the construction there is the fact that the left peripheral subject must be distinct from the wh-element, i.e., the following is impossible:

(19) French
    * Qui est-il parti?
    ‘Who did-he leave?’

Rizzi & Roberts (1989), following a suggestion due to Marc-Ariel Friedemann, analyzed (19) as an ECP violation: movement from the left-peripheral subject position to the landing site of wh-movement would violate the head-government requirement of the ECP. How does this analysis translate into the system developed here?

Evidently, in this construction, an extra subject position is licensed in the lower part of the left periphery. One possible way to go is to assume that I–to–C movement can carry along the Subj head to the left periphery, where it remains active to license an A-specifier. If it is so, the subject criterion configuration is reconstituted in the left periphery, yielding a representation like the following:

(20) Où Foc [ Jean est+Subj+Fin [ il ... allé ]]

If this derivational option is taken, and the subject is a wh-element, we would obtain an intermediate representation like:
(21) Foc [ qui est+Subj+Fin [ il ... parti ]]

But here *qui* satisfies the subject criterion, therefore under criterial freezing it cannot move further to the landing site of a wh-element, Foc. The impossibility of (19) can thus be captured, and another case for which the ECP had been evoked can fall under the freezing approach.

5 Conclusions

The ECP had a broad explanatory role in GB syntax, where it offered a coherent account of different constraints on movement across languages. The core case was the asymmetries between subject and object extraction from embedded domain, the former being more severely constrained than the latter, all other things being equal. Starting from the analysis of the core cases, a very large array of phenomena across languages turned out to be amenable to an ECP analysis.

Under minimalist guidelines, the ECP showed problematic features both conceptual and technical: on the one hand, it did not seem to naturally fit the principled typology of principles foreseen by minimalism; on the other hand, its crucial reliance on government was problematic in a framework explicitly attempting to do away with the government relation. So the principle was abandoned, and the vast body of empirical discoveries connected to the ECP fell out of center stage in the minimalist literature.

In this article I have tried to show that certain important effects analyzed in terms of the ECP in previous literature (including my own work) could be advantageously reanalyzed in different terms, relying on cartographic work and on the system of criteria in particular. Criterial configurations are Spec-head configurations which go with special interpretive instructions of the scope-discourse kind. So, criterial heads such as Top, Foc, Q, Rel, etc. attract phrases with matching features to the specifier position, and guide the interpretation of the structure, e.g., as expressing the topic-comment or focus-presupposition articulation, or explicitly marking the scope of operators. One remarkable syntactic property of criterial positions elucidated in the recent literature is the freezing effect: a phrase meeting a criterion (or, more accurately, the criterial goal) is frozen in the criterial configuration and cannot undergo further movement. Criterial positions thus are “halting” sites for syntactic movement. In a number of articles starting from Rizzi (2006) I have argued that freezing plays a key role in the explanation

---

4 As for the possibility of local subject questions in general, *qui est parti?*, *who left?*, etc., one of the “skipping devices” assumed in Rizzi & Shlonsky (2007) must be operative.
of classical ECP effects. If there is a subject criterion, the halting character of subject positions is immediately captured. The difficulty of extracting subjects, the prototypical case of which is the that-trace effect, can be made to follow from freezing. Subject–object asymmetries follow from the fact that there is a subject criterion but not (in typical cases) an object criterion.

In certain cases, the freezing approach is empirically advantageous compared to the ECP approach. We have seen a number of syntactic phenomena showing asymmetries (en cliticization, beaucoup extraction in French, etc.) in which a requirement of lexical government seems to be too weak, whereas a freezing analysis correctly captures the facts.

No attempt is made here (or in related work of mine) to capture the whole array of ECP phenomena in terms of freezing. For instance, the whole chapter of ECP effects at LF, and many of the “ECP extensions”, in Kayne’s (1984) sense are not addressed. Nevertheless, it is important to stress that some core ECP effects are naturally and advantageously amenable to an explanation in terms of tools provided by recent syntactic theorizing. This offers the promise that also other aspects of the vast and varied ECP phenomenology may regain the focus of attention and offer new grounds to test the explanatory capacities of current syntactic theory.

Abbreviations

ECP empty category principle   LF logical form
GB government-and-binding theory   PF phonetic form

UG Universal Grammar

Acknowledgements

This research was supported by the ERC advanced grant no. 340297 “SynCart”.

References


12 Rethinking the ECP: Subject–object asymmetries as freezing effects
