# Ángel Jiménez Fernández

# On the Composite Nature of Subject Islands: A Phase-Based Approach

#### **Abstract**

This paper focuses on the complex factors which render subject domains opaque to sub-extraction. Subjects have been held to be islands for extractability possibilities. Gallego & Uriagereka (2006) suggest that sub-extraction is banned from subjects when they occupy the specifier position of TP because TP is a phase in Romance. By contrast, I show that this is not the right constraint in languages such as Spanish or Italian, in which sub-extraction is licit from both post-verbal and pre-verbal subjects. In addition, English and other non-Romance languages also instantiate cases of sub-extraction from subjects, irrespective of their pre-verbal or post-verbal position. Building on Chomsky's (2008) notion of phase, I propose that DPs may be strong or weak phases depending on two major discourse-related factors; namely, Definiteness and Discourse-Linking. Thus, sub-extraction from a weak DP phase is possible if the DP is marked as indefinite and discourse-linked.

### 1. Introduction<sup>1</sup>

In this work I am concerned with the nature of subject islands. There has been a long-standing line of research within Generative Grammar since Ross's (1967) and Chomsky's (1973) first efforts to identify the different conditions which delimit the power of transformations. These constraints have been ever since comprised under the notion of island, which may be defined as a syntactic domain which bars extraction of a constituent out of it (Huang's [1982] extraction domains).

<sup>&</sup>lt;sup>1</sup>I thank Andrew Radford, Ignacio Bosque, Ian Roberts, Norbert Hornstein, Violeta Demonte, David Adger, Ángela Di Tullio, Amaya Mendikoetxea, Robert Borsley and Ana Ojea for their insightful comments on some parts of the present paper. I am also grateful to three anonymous reviewers of SKY Journal of Linguistics for their incalculable constructive criticism.

In line with Boeckx (2008), I take the view that not all islands are identical, claiming that varied factors influence the island effects that a given type of constituent arises. Concentrating on subject islands in opposition to objects (leaving aside the question of adjuncts), I explore the complex nature of extraction out of an island. Furthermore, not all types of island show an identical behaviour in respect of the extraction of their members. This has led linguists to draw a distinction between strong and weak islands (Cinque 1990; Postal 1998; Szabolcsi & den Dikken 2002). DPs have been claimed to be strong islands, especially if they are definite/specific (Fiengo & Higginbotham 1981; Manzini 1992, 1998) and if they are placed in subject position. In this work I implement the idea that the notion of island should be relativised as cross-linguistic data show that in some languages sub-extraction from a subject may be possible. This selective nature depicts DPs as weak islands. Examples such as (1) from Spanish illustrate the selective islandhood of subjects:<sup>2</sup>

<sup>2</sup>In the examples I use the trace-convention for movement instead of the copy-convention for reasons of space. An anonymous reviewer points out that examples such as (1) are not relevant for the claim that subjects are not always islands, since it contains unaccuative  $\nu$ Ps. Unaccusative subjects are underlying objects, so that they are not candidates for islandhood.

Following the classification provided by Levin & Rappaport Hovav (1995) and Mendikoetxea (1999), neither *creer* 'believe' nor *ser* 'be' stand as unaccusative verbs. One of the tests that Mendikoetxea (1999: 1583) offers as a diagnostic for unaccusativity consists in the possible occurrence of bare subjects:

(i) Siempre vienen mujeres. always come women 'Women always come.'

If *creer* or *ser* were unaccusatives, then they should allow for the occurrence of bare NP subjects:

(ii) \*Mujeres creen que libros de Juan son interesantes. women believe that books of John are interesting 'Women believe that John's books are interesting.'

Additionally, by definition, unaccusative verbs have only one argument which is assigned the semantic role of Theme. As far as *creer* 'believe' is concerned, it requires two arguments with their corresponding semantic roles.

In the light of these remarks, I do not find examples such as (1) irrelevant in order to show that sub-extraction out of subjects in pre-verbal position is plausible in Spanish.

<sup>(</sup>i) Signapus vienen muienes

(1) ¿De qué autori crees que varios libros  $t_i$  son interesantísimos? of which author believe-PRES.2SG that several books are very interesting 'Of which author do you believe that several books are very interesting?'

To put subject islands in perspective, the asymmetric behaviour of subjects and objects can be traced back to Huang's (1982) Condition on Extraction Domains (CED), according to which objects are transparent extraction domains, whereas subjects are opaque to extractability (on a par with adjuncts). This may well explain the difference in (2) (Chomsky 2008: 146):

(2) a. of which car<sub>i</sub> did they find the (driver, picture) t<sub>i</sub>?
b. \*of which car<sub>i</sub> did the (driver, picture) t<sub>i</sub> cause a scandal?

Throughout the history of Generative Grammar, two ways to explore the properties of islands may be clearly identified. Boeckx (2008) establishes this distinction, which I briefly outline. First, islands have been described as being an issue of the narrow syntax (Chomsky 2004, 2008). From this viewpoint they are part of the computational system and their impact on grammaticality is seen as derivational. For Chomsky (2008) subject DPs are phases, hence nothing can be extracted out of them in conformity with the Phase Impenetrability Principle (PIC).<sup>3</sup> In a similar vein, a very insightful approach to subject islands as a derivational phenomenon is Gallego & Uriagereka (2006), who claim that sub-extraction is blocked from a derived subject DP in Spec-TP, since TP is a phase in certain languages. I review this proposal below.

A second view takes islands to be conditions on the output of the narrow syntax, hence being applied on the product of derivations at the interfaces (Kayne 1984).<sup>4</sup> In accordance, islands are described as

<sup>&</sup>lt;sup>3</sup>I will come back to the notion of phase and the PIC in section 2.

<sup>&</sup>lt;sup>4</sup>Hornstein et al. (2007) make a distinction between LF-driven and PF-driven islands depending on whether the repairing phenomenon that circumvents the island is related to Logical Form or Phonological Form. Sluicing and resumption are two such repairing operations. Merchant (2001) has also claimed that, at least, some islands are PF-driven, while others are conditions at LF. An alternative view is found in Lasnik (2001) and Fox & Lasnik (2003), who suggest that there is no real reason to make such a

representational limitations. These two extreme poles can be reconciled depending on whether we take a uniform or heterogeneous approach to the nature of islands. Boeckx (2008) claims that the interface or narrow-syntax origin of the repairing mechanisms employed to mitigate the island character of a constituent proves an extremely useful tool to understand the very core of the existence of islands. To put it in other words, the extraction possibilities and the repairing strategies employed tell us whether islands emerge in the narrow syntax or at the interfaces.

In this connection, the interaction of islandhood and phenomena such as ellipsis or resumption has been vastly investigated (Boeckx 2003; Boeckx & Lasnik 2006; Hornstein et al. 2007; Merchant 2001). To illustrate the mitigating power of resumption, consider (3) (Boeckx 2008: 155):

- (3) a. \*Which woman did John laugh [after Bill kissed \_\_]
  - b. Which woman did John laugh [after Bill kissed her]

Adjuncts are felt to be islands, and as such nothing can be extracted out of them. However, if a resumptive pronoun is inserted in the gap that the extraction site leaves, the overall construction seems to be grammatically licensed. The problem arises when it comes to analyse ellipsis and resumption as applying in the syntax or at the interfaces. As my works proceeds, it will become clear that linguistic theory should characterise islands as involving both derivational and representational conditions.

In this work I explore the nature of islands and propose that the mitigating effects are a composite set of properties that may obviate the degradation of constructions when extraction out of them comes to play. I propose that subject islands are phases and that the phasehood of DP subjects arises from a set of intermingling discourse-related semantic properties, such as Definiteness and Discourse-Linking, independently from the derived position of the DP. Similar to  $\nu$ P, I take DPs to be weak phases but the combination of the factors just mentioned entitles a DP as a strong phase.

distinction. Finally, Hornstein & Uriagereka (2002a, 2002b) have detected LF properties on islands.

I suggest that the island-circumventing factors are discourse features. Hence, in line with Chomsky (2008), the distinction between weak and strong DP phases is reduced to feature visibility and PIC.<sup>5</sup> Consequently, subject islands as strong phases show both representational and derivational properties in that they exhibit interpretive traits which are relevant at the interfaces, but also drive the whole process of the derivation in the narrow syntax.

The organisation of the paper is as follows: i) section 2 focuses on the notion of phase and the phasal status of DPs; ii) section 3 reveals an intriguing challenge to subject islands, namely subject-islands are heterogeneous as regards sub-extraction; iii) in 4 I present Gallego & Uriagereka's (2006, 2007) phase-based approach to subject islands; iv) I identify some shortcomings in this approach in section 5, mainly concerning the distinction between derived and base-generated subjects; v) in section 6 I deal with the island-repairing factors of Definiteness and Discourse-linking; vi) section 7 offers a new phase-based approach to subject islands based on their composite nature and discourse-related features; and vii) the last section summarises my findings.

# 2. On phases and the phasal status of DPs

As advanced in the introduction, my analysis of subject islands relies on the phasal character of subject DPs. In order to make this proposal easier to follow, in this section I briefly present Chomsky's phase model and its technicalities. In addition, I deal with data and arguments that have recently been given to support the idea that DPs may be phases.

Chomsky (2001: 11–12) claims that in order to reduce computational load, derivations of syntactic structures proceed by phases. Phases are roughly cycles of syntactic computation that are sent to the semantic and phonological interfaces, where they receive a Logical Form (LF) interpretation and a Phonological Form (PF) interpretation, respectively.

<sup>&</sup>lt;sup>5</sup>A similar intuition underlies the analysis of islands offered by Truswell (2005: 6), except that he assumes a definition of strong islands based on the notion of multiple spell-out proposed by Uriagereka (1999) and Sabel (2002); namely, "A *strong island* is the non-projecting phrasal sister of a phrasal constituent." Accordingly, all subjects should be strong islands contrary to facts.

Once phases are transferred to these components, they are not accessible for further computation.

One of the main properties of phases is that they are impenetrable. Chomsky (2001: 5) makes this concrete under the Phase Impenetrability Principle (PIC): The head and complement of a phase cannot be accessed by an external probe; only the edge of a phase can be reached from outside the phase. In this connection, syntactic computations are guided by the operation of AGREE between a probe and a goal in order to ensure that all grammatical features are assigned a value and uninterpretable features are deleted before transfer to the interfaces.

The nature and number of phases is still a dark question. As Frascarelli (2006) herself admits, the notion and properties of phases are still an open issue for further research. Chomsky (2008) holds that CP and transitive  $v^*P$  (in opposition to unaccusative/passive  $v^*P$ ) are phases and leaves the door open to the inclusion of DPs in the list.

To illustrate how the phase system works, consider (4a) and its partial derivation in (4b):

(4) a. The band has won a new prize.

```
b. [CP[C \emptyset][TP[DP]] = [The band][T[Thas][v*P]] = [The band][The ba
```

Due to its uninterpretable features, T is an active probe which searches for a suitable goal. There are two candidates: the DP subject and the DP object. As is clear from the morpho-phonological form assigned to the auxiliary under T, this category agrees with the DP subject. The DP object is not accessed because, in compliance with PIC, it is in the complement domain of the  $v^*P$  phase. This phasal chunk has been transferred to the interfaces so that the DP object is not in the workspace of T. The EPP feature under T ensures that the category agreed with moves to Spec-TP.

As regards the properties of phases, there is no general consensus as to the uniform nature of all types of phases. To advance the strength of my proposal that subject islands are DP phases, I outline some of the arguments in favour of assigning phasal traits to DPs. Hicks (2009) holds that a phase is any syntactic structure corresponding to a proposition: CP is a phase as long as it includes tense and force;  $\nu$ P is a phase when it has a complete argument structure (transitive  $\nu$ \*Ps). By analogy, DPs can be phases in as much as they may have a complete argument structure:

### (5) Chomsky's publication of a new book

It is evident that DPs such as (5) contain an Agent subject and a Theme object. This is obviously reminiscent of the argumental structure of  $v^*Ps$  and the propositional character of CPs. As stated by Hicks (2009: 150), "the general tendency after Abney (1987) towards unifying the nominal and clausal architecture (DP and CP) would also be consistent with such an approach."

Many attempts have been made to unify phases and account for the phasal properties of DPs (see Hiraiwa 2005, Legate 2003, Matushansky 2005, Svenonius 2004, to mention just a few). The main trend has been to uncover LF-properties and PF-properties.

One crucial property of phases is the presence of a subject. For Chomsky, only those vPs which project a subject are phases. For Hicks (2009), DPs are phases when they have a subject. This definition is employed to explain binding relations within DPs.

#### (6) $John_i$ likes $[Bill_i]$ 's pictures of himself<sub>i.\*i</sub>].

The anaphor can only be interpreted as bound by Bill, the DP internal subject. Binding is based on the local domain and Hicks claims that DPs with a subject on their own are the local domain where binding applies. If the DP lacks a subject the binding domain extends to the next higher local domain, namely  $v^*P$ :

### (7) $John_i$ likes [pictures of himself\*\*<sub>i</sub>].

In other words, DPs may also be divided in two: strong phases and weak phases, depending on whether they contain an explicit subject or not. Hicks (2009) simply assumes that phrases are phases or non-phases, but this distinction basically corresponds to the difference between strong and weak

phases. If binding configurations are interpreted at LF, this is a good piece of evidence in favour of the LF-phase status of DPs.

As Frascarelli (2006) observes, Chomsky (2000) assumes that CP and  $v^*P$  are phases because they are semantically complete and "isolable" at the interfaces. It is self-evident that CPs are phonetically independent. Evidence for  $v^*P$  isolability is given by Chomsky (2000): pseudoclefting, fronting and response fragments:

- (8) a. What John did was [insult the dean].
  - b. John said that he would insult the dean and [insult the dean] he did.
  - c. [Me insult the dean]!?

Evidence for the phase status of DPs comes from extending Chomsky's (2000) tests to the DP (see also Matushansky 2005):

- (9) a. What John bought is [the last book by Chomsky].
  - b. [The last book by Chomsky] I haven't read yet.
  - c. Q: What did you buy?
    - A: [The last book by Chomsky].

Pseudo-clefting, fronting and response fragments show that DPs may be phonetically isolated, which is one of the properties of phases at PF. It is thus natural to conclude that DP may be a phase. However, Matushansky (2005) demonstrates that the phasal status of DPs is not uniform at PF and LF. Giusti (2006) also entertains that DPs are not fully independent LF-phases.

When dealing with diagnostic tests for phasehood, Matushansky (2005) classifies them into different types depending on whether they are related to phonology, semantics or syntax. As shown above, DPs show PF-and LF-isolability. Now, I concentrate on the phasal status of DPs in the syntax. Matushansky (2005) argues that what is transferred to the interfaces is the complement of phases. From this it follows that TP and VP cannot move while their corresponding phases, CP and  $v^*P$ , can. As regards DPs, the prediction is that if a DP is a phase, its complement (NP or N, in Abney's [1987] system) does not move in the syntax; conversely, the whole phasal DP may undergo movement as a whole, which is reflected at

PF. If quantifier fronting is treated with in terms of movement (Barbosa 2009), we can build an argument in favour of the phasal status of DPs, see (10a). Nevertheless, in languages such as Spanish the NP complement of a DP may be moved to the left periphery, which suggests that Matushansky's claim that the complements of DPs cannot move is not maintained in all languages, as illustrated in (10b):

```
(10) a. [DP Muchos pasteles]<sub>i</sub> comeré t<sub>i</sub> en la fiesta. many cakes eat-FUT.1SG at the party 'I will eat many cakes at the party.'
b. [NP Pasteles]<sub>i</sub> comeré muchos t<sub>i</sub> en la fiesta. cakes eat-FUT.1SG many at the party 'I will eat many cakes at the party.'
```

At a first sight, NP-preposing seems to contradict the phasal status of DPs. Recall that, according to PIC, the complement of a phasal head is not accessible from outside. In other words, C cannot attract any material in the phase domain. However, the edge of the phase is a position which may be used for further computation. This edge can be used as an escape hatch. Accordingly, the NP *pasteles* 'cakes' moves first to the specifier of DP and subsequently raises to spec-CP. If this analysis is correct, it sheds some light upon the phasal nature of DPs in that their behaviour in relation to internal movement is identical to that of CP and v\*P.

Now I turn to some morphological basis for the phasal status of DPs. Giusti (2006) proposes the existence of a more fine-grained structure in DPs, which includes discourse features, such as topic/contrast, and even EPP. In Albanian, adjectives occur in post-nominal position in the unmarked order. However, they can take a pre-nominal position if they are emphasised (Guisti 2006: 170):

```
(11) a. një grua tjetër e bukur a woman other nice
b. një e bukur grua tjetër a nice woman other
'another nice woman'
```

The derived order is obtained via movement to a KontrastPhrase in order to value the feature [+ Kontrast]. If movement is morphology-driven and if movement inside a specific phrase is symptomatic of its phasehood (see discussion in Matushansky 2005), the conclusion is that DPs may be phases.

All the above remarks point to the fact that DPs may be phases. Both from an interpretive (LF) and phonological (PF) point of view, DPs qualify as phases. One should be careful with this conclusion since the number and nature of phases is still a debatable point, as pointed out above. My contribution to this current debate is to propose that certain LF-related properties are crucial to turn a DP into a phase.

### 3. Presenting a challenging discovery

It is a standard assumption that sub-extraction is blocked out of subjects. This is commonly acknowledged as the object/subject asymmetry, which treats subjects as islands in respect of extractability possibilities. Accordingly, there have appeared many proposals to explain the paradigm in (12), from Lasnik & Saito (1992):

- (12) a. Who<sub>i</sub> did you hear [a story about  $t_i$ ]?
  - b. \*Who<sub>i</sub> did [a story about  $t_i$ ] amuse you?

The distinction between the behaviour of objects and subjects has been taken as the basis to claim that subjects ban sub-extraction, whereas objects do not. As stated in the introduction section, recent research has tried to explain the object/subject asymmetry from multiple perspectives. Some linguists have claimed that moved constituents do block sub-extraction, thereby accounting for the impossible sub-extraction for subjects as opposed to objects, when they are attracted to Spec-TP. Although using different explanatory and descriptive tools, this is the line pursued by Takahashi (1994), Gallego (2007) Gallego & Uriagereka (2006, 2007), Hong & An (2007), Stepanov (2007), among many others.<sup>6</sup>

<sup>&</sup>lt;sup>6</sup>On the basis of the derived or underlying subject status, Chomsky (2008) holds that whereas sub-extraction from an internal argument (including objects and

At the centre of all the afore-mentioned proposals is the prediction that if a subject does not undergo movement, it is still transparent to sub-extraction, similar to objects. In the light of the English and Spanish examples in (13), this prediction is borne out:

```
(13) a. Who<sub>i</sub> is there [a picture of t<sub>i</sub>] on the wall? (Stepanov 2007)
b. ¿De qué equipo<sub>i</sub> dices que han bailado
Of what team say-PRES.2SG that have-PRES.3PL danced
[DP cuatro participantes t<sub>i</sub>]?
four participants
```

'Which team do you say that four members of have danced?'
(Gallego & Uriagereka 2006)<sup>7</sup>

unaccusative/passive subjects) is licit, sub-extraction from external arguments is barred. See also Gallego (2007), where all theses factors are discussed.

<sup>7</sup>The verb *decir* 'say' in Spanish shows at least two different argument structures: i) it may select one single object; or ii) it may require one object and a prepositional object (PO):

- (i) Juan dijo que cuatro miembros del equipo habían
  John say-PAST.3SG that four members of-the team have-PAST.3PL
  bailado toda la noche.
  danced all the night
  'John said that four members of the team had danced all night long.'
- (ii) Juan dijo del equipo que cuatro miembros habían
  John say-PAST.3SG of-the team that four members have-PAST.3PL
  bailado toda la noche.
  danced all the night
  'John said about the team that four of their members had danced all night long.'

If the PO is interpreted as directly selected by the verb, as in (ii), no sub-extraction proper is at stake since in case the PO undergoes *wh*-movement, its source position is not within the subject DP of the subordinate clause. In fact, this PO already belongs within the matrix clause and its movement is not affected by any property of the *that*-clause.

Following this reasoning, sentence (13b) is ambiguous as it shows two different readings: i) the PP *de qué equipo* 'of what team' generates as complement of the noun *participantes* in the subordinate subject position, hence sub-extraction is involved; and ii) the PP originates as a complement of the matrix verb, in which case no sub-extraction is applied. The examples that Gallego & Uriagereka (2006) use involving the verb *decir* can always be felicitous in this second reading because there is no island

The conclusion drawn from these examples is that sub-extraction from subjects is permitted due to the fact that the relevant DPs remain *in situ*, within  $\nu$ P. In languages such as English, where the [EPP] feature is satisfied by moving the DP subject to Spec-TP, this conclusion always holds, except when some other mechanism such as expletive *there* is used to eliminate the [EPP]. Conversely, in languages such as Spanish, which may optionally leave subjects *in situ*, the permissive behaviour of postverbal subjects inhibits the islandhood of the relevant DP, thereby licensing sub-extraction. This is the difference that Gallego & Uriagereka (2006, 2007) detect in (14), taken from Uriagereka (1988):

trespassing. See also Broekhuis (2008) for the independent status of the preposed PP in Dutch examples.

(i) De qué autora; no sabes [[ qué traducciones ti]; C [tj han of which author not know-PRES.2SG which translations have-PRES.3PL ganado premios internacionales]]? won awards international

'Which author don't you know what translated books by have won international awards?'

Also Rizzi (2006: 114) discusses similar issues on extraction possibilities in relation to sentences such as (ii):

(ii) ?[CP][Di] quale autore $_j]$  C ti domandi [CP] [quanti libri  $t_j]_i$  C [TP] siano Of which author CL-2SG wonder how many books are stati censurati  $t_i]]]?$  been censored

'Which author do you wonder how many books by have been censored?'

However, the constructions in these two studies are different to the ones I am investigating in several respects. In both structures the subject contains two *wh*-operators. One of them raises to the subordinate CP; the second operator undergoes *wh*-movement to matrix CP. The two movements at issue are triggered by a Q-feature in each C. In my analysis, only the matrix C contains a Q-feature.

A second property which distinguishes Rizzi's structures from mine is that he applies extraction to passive subjects, hence derived subjects. As already noted in the main text there is no bar on extraction from derived subjects (Stepanov 2007).

<sup>&</sup>lt;sup>8</sup>An anonymous reviewer points out to me that Torrego (1985: 31) has already discussed data concerning extraction from subjects in Spanish:

```
parece
(14) a. ¿De qué conferenciantes; te
                                                           que me_z
                                   CL-2SG seem-PRES.3SG that CL-1SG go-PRES.3PL
         Of what speakers
        a impresionar_v [v*P[DP las propuestas t_i] t_z t_v]?
        to to-impress
                                the proposals
      b. *¿De qué conferenciantes; te
                                              parece
                                                             que [DP] las propuestas t_i ]_i
           Of what speakers
                                     CL-2SG seem-PRES.3SG that the proposals
                             a impresionar<sub>v</sub>[_{v*P} t_i t_z t_v]?
        CL-1SG go-PRES3PL to to-impress
         'Which speakers does it seem to you that the proposals by will impress me?'
```

The discovery I wish to reveal is the fact that preverbal subjects may also allow for sub-extraction in languages such as Spanish, provided that certain grammatical conditions are obeyed. Contra Gallego & Uriagereka (2006, 2007), I present data in (15) and (16) which demonstrate that sub-extraction from subjects is licit in Spanish regardless of the syntactic position they occupy:

- (15) a. ¿De qué cantante crees que son muy provocativas varias fotos? of which singer believe-PRES.2SG that are very provocative several photos
  b. ¿De qué cantante crees que varias fotos son muy provocativas? of which singer believe-PRES.2SG that several photos are very provocative 'Of which singer do you believe that several photos are very provocative?'
- (16) a. ¿De qué cantante parece que les han escandalizado of which singer seem-PRES.3SG that CL-3PL have-PRES.3PL shocked algunas fotos? some photos
  - b. ¿De qué cantante parece que algunas fotos les han of which singer seem-PRES.3SG that some photos CL-3PL have-PRES.3PL escandalizado? shocked

'Of which singer does it seems that some photos have shocked them?'

At least in Southern Peninsular Spanish these sentences are felicitous, which challenge Gallego & Uriagereka's claim that pre-verbally moved subjects are not candidates to permit sub-extraction. Note that in the

subordinate clauses in (15b) and (16b) their respective subjects *varias fotos* de qué cantante 'several photos of which singer' and algunas fotos de qué cantante 'some photos of which singer' have undergone movement to Spec-TP. Yet this does not yield an incorrect outcome.

Similarly, in English sub-extraction is also allowed (though marginally) even when the subject undergoes movement to Spec-TP, contrary to Chomsky's (2008) claim:<sup>9</sup>

- (17) a. ??Which singer did some pictures of shock the audience?
  - b. ??Which writer did a poem of shock the audience?

The EPP feature in (17) is satisfied by moving the DP subjects to Spec-TP. As such, this movement renders the subjects opaque for sub-extraction, hence predicting the ungrammaticality of (17), contrary to facts. This also calls into question the validity of Gallego & Uriagereka's proposal.

Lastly, it is also known that not only subjects are opaque to sub-extraction. Alongside subjects, objects are reluctant to allow *wh*-operators to trespass their DP boundaries under certain conditions. Stepanov (2007), building on Diesing (1992), offers examples in which sub-extraction from objects is blocked:

(18) \*Who<sub>i</sub> did John read every/all/most/the story/stories about  $t_i$ ?

Under the view that objects are transparent to sub-extraction, the ungrammaticality of (18) is unexpected.<sup>10</sup> Thus, the claim that subjects and

<sup>&</sup>lt;sup>9</sup>Chomsky (2008) discusses cases of sub-extraction with preposition pied-piping. Ian Roberts (p.c.) points out that preposition stranding yields slightly better results. He suggests that this is because of the unnatural character of pied-piping in at least these cases. I will not go into the reasons for this difference. In this work I consider both the pied-piping and P-stranding versions of these constructions. As expected, there is no general consensus among syntacticians on the acceptability of sub-extraction, though there is a common preference for P-stranding constructions.

<sup>&</sup>lt;sup>10</sup>Following a suggestion by Hornstein (p.c.), if indefinites involve a kind of restructuring that strong quantifiers cannot undergo, there may well be an expected difference here. In line with Diesing (1992), weak Q NPs are actually NPs whereas strong Q headed nominals are DPs. Thus, the results of sub-extraction are expected to be acceptable from NPs in clear contrast with DPs. As shown below, distinguishing between weak DP phases and strong DP phases also makes the correct predictions.

objects differ in terms of islandhood needs some revising. On the one hand, it is not the case that all subjects are opaque to sub-extraction, both cross-linguistically and language-particularly; on the other hand, it is untenable that all objects allow for internal movement.<sup>11</sup>

As suggested by Adger (p.c.), if we keep to the pied-piping construction and use *picture*-nouns, some examples may be correct:

(i) Of what<sub>i</sub> did John buy ?every/\*all/|\*most/\*the picture/pictures  $t_i$ ?

Davies & Dubinsky (2003) note this difference and adduce it to the semantic nature of *picture*-nouns in contrast with other nouns such as *story*. What is important is that not all types of sub-extraction yield a felicitous outcome, contrary to the standard view.

<sup>11</sup>Sabel (2002) holds that extraction out of a subject is barred due to the fact that DP subjects are barriers, as opposed to DP objects (cf. Chomsky 1986). It is again expected that sub-extraction should be licensed from DP objects, but banned from DP subjects. This prediction is not borne out in the light of the cross-linguistic data offered throughout my work.

Gallego & Uriagereka (2007) claim that objects in Spanish do not block sub-extraction and relate the extractability possibilities to Case assignment. Following Torrego (1998), they make a difference between objects which are preceded by the preposition a and objects which are not; otherwise, they are not case-assigned. On the basis of this distinction, Gallego & Uriagereka (2007: 64) hold that non-a-marked objects allow for sub-extraction, whereas a-marked objects do not (their grammatical judgement):

- (i) (?) ¿De qué artista se limpiaron ya los cuadros? of which artist CL clean-PAST.3PL already the paintings 'Which artist were the paintings by already cleaned up?'
- (ii) ?\* ¿De qué padres se limpió ya a los hijos? of which parents CL clean-PAST.3SG already to the children 'Of which parents were the children already cleaned up?'

Surprisingly, if the extraction domains are introduced by a different kind of D, these sentences appreciably improve and the deviance vanishes:

(iii); De qué artista se limpiaron ya algunos cuadros? of which artist CL clean-PAST.3PL already some paintings 'Which artist were some paintings by already cleaned up?'

As regards the ungrammaticality of (ii), it is not the case that all *a*-marked objects ban sub-extraction. This is clear in (iv), which suggests that Case assignment is not the reason why (ii) is degraded:

### 4. The syntactic position of subjects

In this section I deal with some derivational factors which influence the nature of subject islands. The crucial fact seems to be that subjects are islands when they occupy a derived position, and by extension, extraction out of subjects is allowed if they remain *in situ*. This line of research has been pursued by linguists such as Diesing (1992), Takahashi (1994), Lasnik & Saito (1992), Wexler & Culicover (1981), and more recently Stepanov (2007) and Gallego & Uriagereka (2007).

To start the discussion, passive subjects in English undergo movement to Spec-TP. Once the passive subject sits in Spec-TP, it is a derived subject in that it does not occupy the base-generated position. The prediction is that sub-extraction out of a passive subject should be banned. This is confirmed by (19), extracted from Stepanov (2007):

(19)  $?*Who_i$  was  $[a friend of t_i]_i$  arrested  $t_i$ ?

As is clear, extraction out of a passive subject is blocked in English. However, if the subject remains in its original position the degradation goes away, as instantiated in example (20), taken from Stepanov (2007):

(20) Who<sub>i</sub> is there [a picture of  $t_i$ ] on the wall?

The expletive *there* satisfies the EPP requirement of T, hence the logical subject does not undergo movement to Spec-TP and extraction is not blocked (Takahashi 1994; Stepanov 2007; Gallego & Uriagereka 2006, 2007). In Gallego & Uriagereka's system, the licensing conditions on extractability are linked to the freezing effects that Spec-TP is subject to. They explain the freezing effect of Spec-TP in terms of the Edge Condition, which states that "Syntactic Objects in phase edges become

<sup>(</sup>iv)?\*¿De qué padres has visitado a muchos amigos?

of which parents have-PRES.3SG visited to many friends

<sup>&#</sup>x27;Of which parents have you visited many friends?'

<sup>&</sup>lt;sup>12</sup>Davies & Dubinsky (2003) also arrive at the conclusion that extraction from subjects is banned in English due to the satisfaction of the EPP under T.

internally frozen" (Gallego & Uriagereka 2006: 5).<sup>13</sup> This Edge Condition accounts for the difference between (19) and (20) in that the DP *a picture of who* moves to Spec-TP only in (19), predicting that it gets frozen in this position, thereby blocking sub-extraction. By contrast, in (20) the Spec-TP is filled with the expletive, hence the DP *a picture of who* remains in its base-generated position, which enables it to permit sub-extraction.

Gallego & Uriagereka (2006, 2007) go further and suggest that the phase edge involved in these cases is the specifier of a  $\varphi$ -complete T. Chomsky (2008) holds that phases are CP and  $v^*P$ . From this it follows that TP is not a phase, at least in principle. In order to solve this difficulty, Gallego & Uriagereka (2006), building on Gallego (2007), propose the phenomenon of Phase Sliding, which basically consists of turning TP into a phase as a consequence of v-to-T movement in Romance (see also den Dikken 2007 on a similar idea based on extending phases). From this it follows that phases are still uniform cross-linguistically, so that CP and  $v^*P$  are phases in all languages; yet under certain conditions TP may be a phase in a specific language if little v undergoes v-to-T movement. In other words, TP inherits its phasehood from vP.

If TP may become a phase under certain circumstances in Romance, this predicts that no sub-extraction is allowed from the subject when it is placed in Spec-TP. To illustrate this prediction, Gallego & Uriagereka use the Spanish examples in (21), repeated for convenience:

```
(21) a. ¿De qué conferenciantes; te parece que mez van
Of what speakers CL-2sG seem-PRES.3sG that CL-1sG go-PRES.3PL

a impresionar, [v*P [DP las propuestas ti] tz tv]?
to to-impress the proposals
b. *¿De qué conferenciantes; te parece que [DP las propuestas ti] j mez
Of what speakers CL-2sG seem-PRES.3sG that the proposals CL-1sG

van a impresionar, [v*P tj tz tv]?
go-PRES.3PL to to-impress

'Which speakers does it seem to you that the proposals by will impress me?'
```

<sup>&</sup>lt;sup>13</sup>The explanatory power of this condition is essentially identical to Rizzi's (2006) Criterial Freezing. Also Hong & An (2007) employ the same strategy to distinguish between subjects and objects in respect of extractability.

The difference in terms of grammaticality is related to the fact that in (21b) the DP subject *las propuestas de qué conferenciantes* 'the proposals of which speakers' is an island because the whole DP has undergone movement to Spec-TP prior to sub-extraction to Spec-CP in the matrix clause. This previous movement entitles the whole TP as a phase edge via Phase Sliding (note that v moves to T). In accordance with the Edge Condition, the DP freezes at Spec-TP, thereby disallowing sub-extraction.

Conversely, in (21a) the DP subject *las propuestas de qué conferenciantes* 'the proposals of which speakers' stays *in situ*, which enables the higher probe C to see inside and attract the *wh*-operator.<sup>14</sup> In other words, the DP subject is not placed in a phase edge, thus there is no ban on sub-extraction to Spec-CP in the main clause.

So far, it seems that Gallego & Uriagereka's (2006) proposal is on the right track in that it employs the notion of phase as a primitive, hence deriving the notion of island. This proposal explains the subject/object asymmetries that my work is exploring in that sub-extraction is allowed out of objects, but disallowed out of subjects, except when these remain *in situ*. Note that contrary to other properties of islands, in Gallego & Uriagereka's system it is purely syntactic mechanisms that rule the island effects, namely edge phases and the Edge Condition. From this it follows that the distribution of islands is a narrow-syntactic phenomenon. This sheds light on Boeckx's (2008) claim that islands show both representational and derivational properties. In other words, the provisional conclusion is that a constituent is an island due to lack of interpretive content at the interfaces, or due to a specific syntactic position achieved in the narrow syntax.

Interestingly, Gallego & Uriagereka (2006, 2007) reduce the notion of island to the freezing effects of the specifier of a  $\phi$ -complete T. This raises the question as to the sub-extraction possibilities of a subject which is moved to the specifier of a defective T, for instance in ECM constructions. Chomsky (2008) makes a distinction between  $\phi$ -complete T and defective T in terms of sub-extraction. The relevant examples occur in (22):

<sup>&</sup>lt;sup>14</sup>The fact that post-verbal subjects show specific properties different from pre-verbal subjects has been vastly explored in the literature: Uribe-Etxevarria (1994); Ordóñez (1998, 2005); Cardinaletti (2004); Ortega-Santos (2008), among many others.

- (22) a. \*Of which car did [the (driver, picture) cause a scandal]?
  - b. Of which car did they believe the (driver, picture) to have caused a scandal?

T in ECM constructions is not  $\varphi$ -complete, hence its specifier does not qualify as a phase edge in (22b). Provided that in this case the Edge Condition is not operative, sub-extraction out of an ECM subject is allowed.<sup>15</sup> In strong contrast, in (22a) the subject occupies the specifier position of a  $\varphi$ -complete T so that it becomes a phase edge. Thus, the DP subject in (22a) freezes in Spec-TP and sub-extraction is banned.

### 5. Shortcomings of the phase-based approach

In this section I present theoretical and empirical arguments against Gallego & Uriagereka's (2007) phase-based approach to subject islands. Although I do agree that sub-extraction is connected to the notion of phases, it might be the case that it is not the phase nature of T in Romance that bans sub-extraction from subject DPs when they are in Spec-TP. Assuming a non-absolute definition of phasehood, a category may qualify as a phase depending on certain conditions. The idea is not new. In this respect, Sevdali (2009) shows that in Greek a CP may be a strong or weak phase due to the discourse properties of the head C. Also, Chomsky (2008) holds that  $\nu$ P is a selective phase in that only transitive  $\nu$ Ps stand for strong phases. In this line, I hint at the possibility that it is a combination of interpretive properties that make a DP a strong phase.

As mentioned earlier, DP subjects are not islands cross-linguistically. Actually, Stepanov (2007) gives examples of languages such as Hungarian and Palauan, among other languages, which do not block sub-extraction from a subject:

<sup>&</sup>lt;sup>15</sup>As Bianchi & Chesi (2008) note, for Kayne (1983), ECM subjects constitute left branch islands, thereby not allowing any kind of sub-extraction:

<sup>(</sup>i) \*[Which book]<sub>i</sub> do you believe [the first chapter of  $t_i$ ] to be full of lies?

Generally there is no consensus among speakers regarding the grammatical status of sub-extraction from ECM subjects. Because in my analysis I do not focus on the position occupied by ECM subjects, I leave this question aside.

Palauan (Georgopoulos 1991)

(23) Mary [a kltukl [el kmo ng-oltoir er a John \_\_]]
Mary R-clear COMP R-3SG-IMP-love John
'Mary, [that \_\_ loves John] is clear.'

Hungarian (É. Kiss 1987)

(24) *Melyik színésznőnek*<sub>i</sub> *gondolja János, hogy t*<sub>i</sub> *a fényképe meglett?* which actress's thinks Janos that the picture-her turned up 'Which actress does John think that a picture of \_ turned up?'

Furthermore, languages such as English also allow sub-extraction given that the subject is not in Spec-TP. This was the conclusion arrived at by Gallego & Uriagereka (2006), which is illustrated by the following example from Lasnik & Park (2003):

```
(25) a. [CP][Which\ candidate]_i\ were\ [TP\ there\ [vP\ [posters\ of\ t_i]\ all\ over\ the\ town]]]? b. *[CP][Which\ candidate]_z\ were\ [TP\ [posters\ of\ t_z]_i\ [vP\ t_i\ all\ over\ the\ town]]]?
```

Gallego & Uriagereka (2006) claim that sub-extraction from Spanish subjects is barred when they move to Spec-TP due to the phasal extension from v to T. Accordingly, in a language which consistently leaves v in situ, it is predicted that T is not a phasal head. Thus, sub-extraction from Spec-TP in English should be permitted, contrary to facts. If this line of reasoning is correct, the data in (25b) remains unexplained.

Gallego & Uriagereka (2007) try to build a solution by relating extractability possibilities to agreement, in line with Boeckx (2003) and Chomsky (2001, 2008). They explain that a subject gets frozen when full agreement holds between T and the relevant DP. This also poses questions as to the reason why in languages such as Spanish sub-extraction is allowed from Spec-TP.

Moreover, I have compiled examples from English which involve subextraction from a subject in Spec-TP and the outcome is not unacceptable:<sup>16</sup>

<sup>&</sup>lt;sup>16</sup>Actually, as Ian Roberts (p.c.) indicates, (26a) is only slightly degraded and (26b) is perfect with P-stranding. David Adger (p.c.) finds (27a) odd and shows no amelioration with respect to a DP introduced with the definite D *the*, whereas (27b) is fine and contrasts in terms of acceptability with definite DPs.

```
(26) a. ?[Which car]<sub>i</sub> did a driver of t<sub>i</sub> cause a scandal?b. [Which car]<sub>i</sub> did some pictures of t<sub>i</sub> cause a scandal?
```

```
(27) a. ??[Of which car]<sub>i</sub> did a driver t<sub>i</sub> cause a scandal?b. [Of which car]<sub>i</sub> did some pictures t<sub>i</sub> cause a scandal?
```

The conclusion drawn from this data is that in English Spec-TP is a position where sub-extraction may optionally apply (sometimes marginally). This is untenable in the light of the minimalist maxim that options are not allowed. Alternatively, rather than the specific position of subjects, it seems that internal properties of DPs are in charge of rendering them opaque to sub-extraction. I explore this alternative below.

A second problem for Gallego & Uriagereka's proposal is posed by the Spanish data they use to confirm that sub-extraction from post-verbal subjects is allowed, as opposed to pre-verbal subjects. Again, this is rightly predicted if Spec-TP is a phase edge as a consequence of the phasal properties inherited by T. The data is repeated in (28):

```
a. ¿De qué conferenciantes; te parece que mez van
Of what speakers CL-2sG seem-PRES.3sG that CL-1sG go-PRES.3PL

a impresionar [v*p [DP las propuestas ti] tz tv]?
to to-impress the proposals
b. *¿De qué conferenciantes; te parece que [DP las propuestas ti] mez
Of what speakers CL-2sG seem-PRES.3sG that the proposals CL-1sG

van a impresionar [v*p tj tz tv]?
go-PRES.3PL to to-impress
'Which speakers does it seem to you that the proposals by will impress me?'
```

There appears not to be a general consensus among Spanish speakers as to the grammaticality judgement of sentences such as (28). For instance, (28a) is degraded unless the determiner in the extraction site is replaced by a possessive D such as su 'their'. In addition, if the same substitution applies in (29b), the sub-extraction is strongly ameliorated. Note that the same strategy is used by Hungarian, as illustrated in (24):

<sup>&</sup>lt;sup>17</sup>Lasnik & Stowell (1991), Rizzi (2001), Falco (2007) deal with this data in terms of weak cross-over effects and observe that binding of the possessive pronouns in

```
(29) a. ¿De qué conferenciantes; te parece que mez van
Of what speakers CL-2SG seem-PRES.3SG that CL-1SG go-PRES.3PL

a impresionar, [v*P [DP sus propuestas ti] tz tv]?
to to-impress their proposals
b. ¿De qué conferenciantes; te parece que [DP sus propuestas ti] mez
Of what speakers CL-2SG seem-PRES.3SG that their proposals CL-1SG

van a impresionar, [v*P tj tz tv]?
go-PRES.3PL to to-impress

'Which speakers does it seem to you that the proposals by will impress me?'
```

Given that Spec-TP is a phase edge by Phase Sliding, sub-extraction is predicted to be blocked in (29b), contrary to facts. Moreover, sentences in (30) and (31) involve sub-extraction and no ban is put on it regardless of the base-generated or derived position of the subject:<sup>18</sup>

constructions similar to (29) is due to the specific nature of the *wh*-operator. If the *wh*-expression is non-specific, the binding relation does not obtain:

- (i) a. [Who the hell]<sub>i</sub> do (you say that) his<sub>?\*i/j</sub> students admire  $t_i$ ? Non-specific
  - b. [Which famous professor]<sub>i</sub> do (you say that) his<sub>i/j</sub> students admire  $t_i$ ? Specific

The core point seems to be that specificity (understood as Discourse-Linking) ameliorates *wh*-movement and provides with suitable workspace for binding the possessive pronoun. I return to the influence of D-linking on sub-extraction shortly. <sup>18</sup>An anonymous reviewer points out to me that sentences such as (30b) and (31b) are

<sup>18</sup>An anonymous reviewer points out to me that sentences such as (30b) and (31b) are degraded. In these examples, sub-extraction has been applied after moving the subject to Spec-TP. As mentioned in the main text, in at least certain varieties of Spanish all the examples are well-formed in as much as pragmatic factors such as length allow for such complex constructions.

Among syntacticians there is no general agreement on the well/ill-formedness of sentences such as (30) and (31). Ángela Di Tullio (p.c.) finds all four sentences grammatical, whereas Violeta Demonte (p.c.) and Amaya Mendikoetxea (p.c.), at a first sight, consider them incorrect. However, after close inspection, Demonte detects differences between examples in (a) and (b) and suggests that the (b)-examples improve if the premodifier *tanta* 'such' is replaced by a quantifier such as *mucha* 'much':

(i) ?¿De qué actriz varias fotos han causado mucha polémica? 'Of which actress have several photos caused much scandal?'

Interestingly, the sentences that Gallego & Uriagereka (2007) take as well-formed (examples in (a), with sub-extraction from post-verbal subject) do not show any

- (30) a. ¿De qué actriz han causado varias fotos tanta polémica? of which actress have-PERF.3PL caused several photos such a scandal
  - b. ¿De qué actriz varias fotos han causado tanta polémica? of which actress several photos have-PERF.3PL caused such a scandal 'Of which actress have several photos caused such a scandal?'
- (31) a. ¿De qué actriz parece que han causado varias fotos of which actress seem-PRES.3SG that have-PERF.3PL caused several photos tanta polémica? such a scandal
  - b. ¿De qué actriz parece que varias fotos han causado of which actress seem-PRES.3SG that several photos have-PERF.3PL caused tanta polémica? such a scandal

'Of which actress does it seem that several photos have caused such a scandal?'

If Spec-TP is a phase edge which does block sub-extraction in Spanish, sentences (30b) and (31b) should be incorrect. This prediction is not borne out.

Consider now the possible sub-extraction from objects in Spanish. One of the most prevalent characteristics of the subject/object asymmetry is that objects allow sub-extraction. This is also the view that Gallego & Uriagereka (2006, 2007) adopt on the basis of examples such as (32):

(32) ¿[De qué lingüista] ¡vais a leer muchos artículos t¡? of what linguist go-2PL to to-read many papers 'Which linguist are you going to read many papers by?'

Again, extraction out of objects posits some problems, since not all subextraction cases yield a grammatical output, as shown in (33):

amelioration. Instead, it is the (b)-examples that may get better under certain circumstances.

I am grateful to Violeta Demonte, Amaya Mendikoetxea, Ana Ojea and Ángela Di Tullio for their grammaticality judgements and further suggestions.

-

(33) \*¿[De qué lingüista] ¡vais a leer los artículos t¡? of what linguist go-2PL to to-read the papers 'Which linguist are you going to read the papers by?'

It is safe to conclude so far that sub-extraction from a DP is contingent on other factors which do not rely on the functional/positional status of the relevant DP. Therefore, provided that certain conditions are satisfied, sub-extraction from subjects and objects are ultimately felicitous regardless of the specific syntactic position that the DP occupies. In this respect, I concur with Ceplova (2001), Boeckx (2003) and Boeckx & Grohmann (2007) in assuming that a phase-based approach to subject domains is problematic in that if subject DPs are phases the conceptualization of phases is either too restricted or too permissive. Furthermore, Gallego & Uriagereka (2007) base their analysis of sub-extraction on the phasehood of TP (not on the phase properties of the relevant DP). In my approach, I elaborate a proposal founded on the phasal character of DPs, not on the phasal nature of the position that DPs occupy.

#### 6. Some constraints on internal sub-extraction

### **6.1** Definiteness effects and DP phases

Chomsky (2008) has identified DPs as phases. As Chomsky (2001, 2004) claims, the domain of a phase cannot be targeted by an outer probe in accordance with the Phase Impenetrability Principle. This is just a consequence of the Transfer process, by which a phase domain is sent to the phonological and semantic components to be assigned a phonological representation and a semantic representation, respectively. In this connection, once the domain of a definite DP has been transferred to the other components, nothing could be extracted out of it.

Radford (2009) explores the possibility that definite DPs are phases. To illustrate, consider sentences in (34), taken from Radford (2009):<sup>19</sup>

<sup>&</sup>lt;sup>19</sup>Davies & Dubinsky (2003) have proposed that objects in English are DPs, whereas subjects are only NPs. This explains why sub-extraction from objects is licensed as opposed to subjects. However, this proposal also poses some problems since, as illustrated in the main text, it is not the case that sub-extraction is allowed from all type of objects.

- (34) a. Who<sub>i</sub> were you reading a book about  $t_i$ ?
  - b. \*Who<sub>i</sub> were you reading the/this/that/his book about  $t_i$ ?

The difference of grammaticality in (34) is adduced to the definite character of the DP object in (34b), which will thus be classified as a phase.<sup>20</sup> One problem that this analysis poses is that it does not discriminate between the definite DP in (34b), barring *wh*-extraction and the definite DP in (35), which seemingly allows extraction in spite of the definite nature of the DP at issue.

(35) Which of these books did you design the covers of?

Both examples (34b) and (35) instantiate the use of definite DPs in object position, but only in (35) will extraction result in a correct sentence, although both DPs are phases due to their definite character.

In relation to the Definiteness Effects that I am dealing with here, Ticio (2006) describes possible extractions out of a DP depending on a three-fold classification of the extracted category in terms of objects, possessors and agents.<sup>21</sup> Dealing with Spanish, she suggests that only objects can be extracted out of a definite DP:

```
(36) a. *; [De qué autor]<sub>i</sub> has leído los libros t<sub>i</sub>? (agent) of which author have-PERF.2SG read the books 'Of which author have you read the books?'
```

b. \*¿[De quién] i has visto [las fotos de ese monte ti]? (possessor) of whom have-PERF.2SG seen the photos of that mountain 'Of whom have you seen the photos of that mountain?'

<sup>&</sup>lt;sup>20</sup>Definiteness effects on the extractability of DPs have been independently explored by Diesing (1992) and Davies & Dubinsky (2003), among others.

<sup>&</sup>lt;sup>21</sup>On previous approaches to the classification of Spanish DP constituents in terms of agents, possessors and objects and their different structural position within DP, see Torrego (1985), Ormazábal (1991) and Sánchez (1996). Giorgi & Longobardi (1991) also offer an analysis of extraction which is based on the type of argument that is included in the relevant DP.

```
c. ¿[De qué cantante] ¡ salieron publicadas las fotos t¡?<sup>22</sup> (object) of which singer were-PAST.3PL published the photos 'Of which singer were the photos published?'
```

The grammaticality of (36c) argues against an explanation of the impossibility of extraction in (36a–b) as a consequence of the definite nature of the DP. Note that the three sentences in (36) involve a definite DP. Contra Fiengo & Higginbotham (1981) and Storto (2000), among others, wh-movement out of definite DPs is not entirely excluded in the light of examples such as (36c). Interestingly, if an indefinite D such as varios/as 'several' substitutes for the definite D in (37) the ill-formedness disappears, provided that there is only one single argument present in the DP:

```
(37) a. ¿[De qué autor]<sub>i</sub> has leído varios libros t<sub>i</sub>? (agent) of which author have-PERF.2SG read several books 'Of which author have you read several books?'
```

- b.  $\[ [De\ qui\'en]_i\]$  has visto [varias fotos de ese monte  $t_i$ ]? (possessor) of whom have-PERF.2SG seen several photos of that mountain 'Of whom have you seen several photos of that mountain?'
- c. ¿[De qué cantante] i salieron publicadas varias fotos ti? (object) of which singer were-PAST.3PL published several photos 'Of which singer were several photos published?'

From the data in (37) a conclusion may be drawn that, regardless of the semantic relation between the noun and its prepositional complement, nondefinite DPs permit sub-extraction.

Moreover, Ticio (2006: 138) goes further when she asserts that Spanish definite DPs and Spanish specific DPs differ with respect to sub-extraction possibilities.

(38) a. \*¿[De qué autor]; has leído estos libros t;? (agent) of which author have-PERF.2SG read these books 'Of which author have you read these books?'

<sup>&</sup>lt;sup>22</sup>The verb *salir* 'come out' is unaccusative in Spanish, so that the subject *las fotos de qué cantante* 'the photos of which singer' originates as complement of VP, thereby behaving as an object.

- b.  $*_{\ddot{G}}[De\ qui\acute{e}n]_i\ has$  visto [estas fotos de ese monte  $t_i$ ]? (possessor) of whom have-PERF.2SG seen these photos of that mountain 'Of whom have you seen these photos of that mountain?'
- c. \*¿[De qué cantante]; salieron publicadas estas fotos t;? (object) of which singer were-PAST.3PL published these photos 'Of which singer were these photos published?'

The ungrammaticality of the examples in (38) with demonstratives suggests that there is no difference among agents, possessors and objects when Specificity Effects are concerned in Spanish, as all types of extraction will be banned in specific DPs irrespectively of the agent, possessor or object status of the moved category.

What seems to be prevalent in Ticio's (2006) approach to Spanish nominals is that all types of extraction involve movement out of a DP which occupies an (underlying) object position. One question arises at this point: What would happen if sub-extraction applied out of a DP in subject position? Sentences in (39) and (40) instantiate cases of extraction out of a DP subject:

- (39) a. ¿De qué cantante has dicho que son muy provocativas varias/las fotos? of which singer have-PERF.2SG said that are very provocative several/the photos 'Of which singer have you said that several/the photos are very provocative?'
  - b. \*¿De qué cantante has dicho que son muy provocativas estas fotos? of which singer have-PERF.2SG said that are very provocative these photos 'Of which singer have you said that these photos are very provocative?'
- (40) a. ¿De qué película has dicho que interrumpieron la conferencia of which film have-PERF.2SG said that interrupt-PAST.3PL the talk

varios/los directores? several/the directors

'Of which film have you said that several/the directors interrupted the talk?'

b.\*; De qué película has dicho que interrumpieron la conferencia of which film have-PERF.2SG said that interrupt-PAST.3PL the talk

estos directores? these directors

'Of which film have you said that these directors interrupted the talk?'

The grammaticality judgement of (39) clearly shows that extraction out of specific DP subjects is banned in Spanish, whereas extraction out of definite and indefinite DP subjects may be allowed under certain circumstances. This is surprising in the light of Huang's (1982) CED, according to which subjects are islands in that, as stated above with respect to English, they do not permit the extraction of any of their constituents.<sup>23</sup> This subject-island condition is illustrated in (40), according to which any extraction out of a DP subject is barred in Spanish, irrespective of the (non)definite/specific status. All the relevant examples in (39–40) improve appreciably when the extraction involves pied-piping of the whole DP subject, except with specifics, which is indicative of the islandhood of these DP subjects:

- (41) a. ¿Varias/las fotos de qué cantante has dicho que son several/the photos of which singer have-PERF.2SG said that are muy provocativas? very provocative
  - 'Several/the photos of which singer have you said are very provocative?'
  - b. \*¿Estas fotos de qué cantante has dicho que son muy provocativas? these photos of which singer have-PERF.2SG said that are very provocative 'These photos of which singer have you said are very provocative?'
- (42) a. ¿Varios/los directores de qué película has dicho que several/the directors of which film have-PERF.2SG said that interrumpieron la conferencia? interrupt-PAST3.PL the talk

'Several/the directors of which film have you said that interrupted the talk?

<sup>&</sup>lt;sup>23</sup>The precise definition of Huang's (1982) CED makes reference to proper government: only those subjects that are not properly governed by a lexical head are islands. As an anonymous reviewer comments, in languages such as Japanese, Spanish, Italian, etc., it was argued that subjects were governed. This leaves a door open to the possibility that in these languages, sub-extraction from subjects is plausible.

b. \*¿Estos directores de qué película has dicho que interrumpieron these directors of which film have-PERF.2SG said that interrupt-PAST.3PL la conferencia? the talk

'These directors of which film have you said that interrupted the talk?'

It is reasonable to conclude so far that Definiteness/Specificity effects arise in relation to extraction out of DPs irrespective of whether they are placed in object or subject position. Accordingly, Definite/Specific DPs are clearly islands and, as such, they may be dealt with in terms of phases. In this connection, Anti-definiteness may be seen as an island-circumventing factor. However, the phase-based approach analysis to DP islands is troublesome in that I have identified clear cases of extraction out of definite DPs in Spanish which yield a correct outcome, even if they are placed in subject position (see (39)–(42)). From this, two conclusions may be drawn: (i) The notion of island should be parameterised in order to capture typical cases of subject extractability in languages such as Spanish, in line with Boeckx (2003), Sabel (2002), Gallego & Uriagereka (2007), among others; (ii) The interpretation of DPs as (non)definite/specific is an interface issue, in that it is relevant at LF where semantic properties are subject to processing. This throws some light into the nature of islands since the circumventing feature seems to be an LF phenomenon, and following Boeckx's (2008) reasoning, subject islands are thus identified as representational conditions on syntactic objects. Other factors seem to be involved in repairing islands though, which I try to clarify in next section.

# **6.2** Discourse-linked operators

Linguists draw a distinction between two types of interrogatives: discourse-linked (D-linked) phrases such as *which man*, which implies the existence of a set of contextually determined entities (*men*) from which the speaker is asking for a choice, and non-D-linked interrogatives such as *who*, which carry no such implication (Pesetsky 1987; Cinque 1990; Enç 1991; Rizzi 2001; Frazier & Clifton 2002). Let's see what happens if *wh*-movement is applied to a sentence such as (43), from Aarts (1992: 47):

### (43) What did you design the covers of?

Sentence (43) is correct, especially if the interrogative operator *what* is interpreted as being D-linked, *i.e.* if *what* refers to a subset of a previously identified set in the context. This accounts for the grammaticality of sentences such as (44), in which the extraction conveniently affects a D-linked phrase (Radford, p.c.):

#### (44) Which of these books did you design the covers of?

The operator *what* in (43) has two interpretations depending on whether it is considered as a D-linked or as a non-D-linked phrase. As stated above, only when it is interpreted as D-linked will sentence (43) be completely felicitous. Note that in the above examples the extraction site is a definite DP, hence a phase, yet if the *wh*-operator is properly identified in the discourse the islandhood of these definite DPs is repaired.

Assuming the subject/object asymmetry as regards the extraction of a *wh*-operator, let's consider the extractability possibilities of D-/non-D-linked *wh*-constituents out of a DP in English:<sup>24</sup>

- (45) a. [Of which car]  $_i$  did they find the (driver, picture)  $t_i$ ? (No subject island + Definite DP + D-linked operator)
  - b. \*[Of which car] i did the (driver, picture) ti cause a scandal? (Subject island + Definite DP + D-linked operator) (From Chomsky [2008], repeated here for convenience)
- (46) a. ??? [Of what] i did they find the (driver, picture) ti? (No subject island + Definite DP + non-D-linked operator)
  - b. \*[Of what] i did the (driver, picture) ti cause a scandal? (Subject island + Definite DP + non-D-linked operator)

In the (a) sentences the extraction site for *wh*-movement occupies the object. As such, it allows for the extraction of a *wh*-operator provided this

<sup>&</sup>lt;sup>24</sup>I thank Ian Roberts, Norbert Hornstein, David Adger, Jane Arnold and Mary O'Sullivan for their grammaticality judgements. It must be stated that no general consensus has been achieved among native speakers of English and syntacticians. Even the acceptable examples quoted from Chomsky (2008) do not sound very good.

is D-linked. Interestingly, the situation is different when the extraction site is the subject of the whole sentence, since no amelioration is felt irrespectively of the D-linkedness or non-D-linkedness of the *wh*-operator. However, if the sub-extraction involves movement out of an indefinite DP, the sentence strongly improves, especially if the *wh*-operator is D-linked, as sentences in (47) illustrate:<sup>25</sup>

- (47) a. [Of which car]<sub>i</sub> did some pictures  $t_i$  cause a scandal?
  - b.  $??[Of what did]_i$  some pictures  $t_i$  cause a scandal?
  - c. (?)[Which car]<sub>i</sub> did some pictures of  $t_i$  cause a scandal?
  - d. [What did]<sub>i</sub> some pictures of  $t_i$  cause a scandal?<sup>26</sup>

This paradigm exhibits the fact that sub-extraction from a DP subject is licit given that extracted material is D-linked and the DP is indefinite. In any case, it should be clear that D-linking and definiteness are interface properties, since their influence is felt at LF once the derivation is transferred to be semantically processed. This leads me to conclude that island-effects are interface conditions.

## 7. A new phase-based approach to subject DPs

In this section I explore another possibility to explain the difference in terms of sub-extraction and its relation to the concept of islands. Implementing Chomsky's (2008) view, I suggest that all DPs are phases, on a par with CP and  $\nu$ P. However, some DPs are strong phases due to the combination of certain interface interpretive properties such as Definitess and D-Linking. Accordingly, only some DPs are islands, hence islandhood

<sup>&</sup>lt;sup>25</sup>The reason that Rizzi (2001) adduces to explain sub-extraction of D-linked *wh*-operators is that, as specific, they contain salient topic properties. Although I agree that information structure plays a role in licensing sub-extraction, I will not pursue this information-based approach here due to lack of space.

<sup>&</sup>lt;sup>26</sup>The P-stranding versions and the grammaticality judgement have been kindly given by Ian Roberts (p.c.). Concerning the pied-piping structures in (47a–b), David Adger detects a contrast between the non-D-linked and D-linked examples. As mentioned above, the operator *what* may have a D-linked reading, which explains why (47d) is well-formed.

is a derived notion. In this vein, what renders DPs opaque to sub-extraction is a complex of interface properties, and not only the derived position of DPs (contra Gallego & Uriagereka 2007). Actually, Chomsky (2008: 152) notes that "what yields the subject-island effect, it appears, is search that goes too deeply into a phase already passed, not the difference between base and surface position." Chomsky draws this conclusion from the grammatical status of sentences like (48):

(48) [Of which car] did they believe the (driver, picture)  $t_i$  to have caused a scandal?<sup>27</sup>

The core point about extractability possibilities is that they are ruled by a heterogeneous series of conditions. The availability of extraction cannot be accounted for by just proposing one single condition. To recapitulate, two factors influencing the extraction possibilities that I have considered are Definiteness and D-linking.<sup>28</sup>

These two conditions identify the possibilities of extraction out of DPs. Mind that I am concerned here with weak islands. This means that the grammaticality of the extractions under investigation is rather selective. The relative weakness of these islands is strengthened when the two factors combine, thereby obtaining a stronger island. The examples in (49)-(54) illustrate the emergence of a strong island when different combinations are taken into account, regardless of the subject/object asymmetry:

<sup>&</sup>lt;sup>27</sup>There is no general consensus as to the grammaticality of ECM constructions which involves sub-extraction from the subordinate subject. While Chomsky considers that sentences such as (48) are correct, Stepanov (2007) – quoting Chomsky (1973) and Kayne (1984) – holds that sub-extraction from an ECM is degraded on the basis of the example ??Who do you believe [a picture of t] to be on sale?

<sup>&</sup>lt;sup>28</sup>For a different list of constituents which induce island effects, see Szabolcsi & den Dikken (2002). I am aware that there are additional factors influencing the islandhood of a given constituent. One such factor may be preposition stranding (Chomsky 1986; Kayne 1984; Kuno 1973). The reason provided by Kuno (1973) lies on the NP-Incompleteness that defines the nominal expression left behind. Although this proposal sounds right, in this work I do not deal with the connection between islands and preposition stranding. I simply concentrate on two discourse properties which are responsible for the emergence of a strong island.

- (49) a. ?; De quién crees que son muy provocativas las fotos? of whom believe-PRES.2SG that are very provocative the photos (Subject island + Definite DP + non-D-linked operator) 'Of whom do you believe that the photos are very provocative?'
  - b. ¿De qué profesor crees que son muy provocativas las fotos? of which teacher believe-PRES.2SG that are very provocative the photos (Subject island + Definite DP + D-linked operator) 'Of which professor do you believe that the photos are very provocative?'
- (50) a. ?¿De quién crees que son muy provocativas algunas fotos? of whom believe-PRES.2SG that are very provocative some photos (Subject island + Indefinite DP + non-D-linked operator) 'Of whom do you believe that some photos are very provocative?'
  - b. ¿De qué profesor crees que son muy provocativas algunas fotos? of which teacher believe-PRES.2SG that are very provocative some photos (Subject island + Indefinite DP + D-linked operator) 'Of which professor do you believe that some photos are very provocative?'
- of what believe-PRES.2SG that have<sub>2SG</sub> met to the directors (No subject island + Definite DP + non-D-linked operator)

  'Of what do you believe that you have met the directors?'
  - b. ¿De qué película crees que has conocido a los directores? of what film believe-PRES.2SG that have-PERF.2SG met to the directors (No subject island + Definite DP + D-linked operator) 'Of which film do you believe that you have met the directors?'

(i)  $*_{\mathcal{C}}[De\ qu\'e\ estudiante]_i\ has$  criticado a los padres  $t_i$ ? of what student have-PERF.2SG criticized to the parents 'Which student have you criticized the parents of?'

Nevertheless, if the definite D *los* is replaced by the possessive D *sus*, the sub-extraction is repaired. In line with Falco (2007) it seems that possessive pronouns give rise to Weak Cross-Over effects, thereby accounting for the possibility of sub-extraction:

(ii) ¿[De qué estudiante]<sub>i</sub> has criticado a sus padres t<sub>i</sub>? of what student have-PERF.2SG criticized to his parents 'Which student have you criticized the parents of?'

<sup>&</sup>lt;sup>29</sup>As noted by Gallego & Uriagereka (2007), sub-extraction is not licensed when the object is introduced by the dative preposition a:

- (52) a. ???¿De qué crees que has conocido a unos directores? of what believe-PRES.2SG that have-PERF.2SG met to some directors (No subject island + Indefinite DP + non-D-linked operator) 'Of what do you believe that you have met some directors?'
  - b. ¿De qué película crees que has conocido a unos directores? of what film believe-PRES.2SG that have-PERF.2SG met to the directors (No subject island + Indefinite DP + D-linked operator) 'Of which film do you believe that you have met some directors?'
- (53) a. ¿De qué coche crees que encontraron al conductor/la foto? of what car believe-PRES.2SG that find-PAST.3PL to.the (No subject island + Definite DP + D-linked operator) 'Of which car do you believe they found the driver/picture?'
  - b. \*¿De qué coche crees que el conductor/la foto provocó of what car believe-PRES.2SG that the driver/the picture cause-PAST.3PL un escándalo?

a scandal

(Subject island + Definite DP + D-linked operator)
'Of which car do you believe the driver/picture caused a scandal?'

- (54) a. \*¿De qué crees que encontraron al conductor/la foto?
  of what believe-PRES.2SG that find-PAST.3PL to.the driver/the picture
  (No subject island + Definite DP + non-D-linked operator)
  'Of what do you believe they found the driver/picture?'
  - b. \*\*; De qué crees que el conductor/la foto provocó of what believe-PRES.2SG that the driver/the picture cause-PAST.3PL

un escándalo?

a scandal

(Subject island + Definite DP + non-D-linked operator)

'Of what do you believe that the driver/picture caused a scandal?'

Several describing generalisations derive from the data above:

1) The more D-linked a *wh*-operator, the more natural the resulting construction. This is one of the ameliorating strategies that Spanish employs to allow for the extraction of an operator from a definite DP, as examples in (51) illustrate. Similarly, English

- may circumvent a subject island by using D-linking, as shown throughout my work.
- 2) The CED on its own cannot account for some cases of extraction out of an island in Spanish, as examples (52)–(54) confirm. This can be taken as evidence against the CED as part of UG. This possibility is vastly explored by Stepanov (2007) in the light of Nunes & Uriagereka's (2000) nondiscrimination between complements and noncomplements. Also, Boeckx (2003) and Gallego & Uriagereka (2006, 2007) arrive at the same conclusion.
- 3) Regardless of their derived or base-generated position, subjects may allow for sub-extraction under certain circumstances. This is the crucial point in my work. It is not the case that subjects do not favour sub-extraction due to the fact that they occupy Spec-TP. It is rather their phasal nature that disallows sub-extraction.

All these generalisations may be accommodated in a principled way if some kind of phase-based analysis is adopted. As mentioned above, if DPs are phases as a consequence of a specific combination of properties such as D-linking and Definiteness, all the special traits of subject islands fall into place. This is the line I want to pursue here. Evidence in support of my analysis of DPs as selective phases comes from cross-linguistic and theoretical grounds.

In this connection, Sevdali (2009) discusses two types of non-finite clauses in Ancient Greek and two types of finite clauses in Modern Greek. Starting with Ancient Greek, she convincingly argues that the presence of discourse properties such as contrast renders a CP a strong phase. This explains why infinitival clauses with overt or null accusative subjects are strong phases, C\*Ps, whereas control infinitives are CPs, weak phases that permit case-agreement operations driven from outside.

As far as Modern Greek is concerned, Sevdali (2009) indicates that there are two types of finite *na*-clauses. If the subordinate clause may have either a controlled PRO or an explicit subject, the CP will be a strong phase. By contrast, if the *na*-clause can only take a controlled PRO as its subject, this CP will be a weak phase. The two examples that follow illustrate this distinction:

(55) O Manolis elpizi na erthi the Manolis-NOM hope-PRES.3SG na come-PRES.3SG

avrio (i Xristina) tomorrow (the Christina-NOM)

'Manolis hopes to come tomorrow'/'Manolis hopes that Christina comes tomorrow.'

(56) *O Manolis kseri na kolibai* the Manolis-NOM know-PRES.3SG *na*swim-PRES.3SG

\*(avrio) \*(i Xristina) (tomorrow) (the Christina-NOM)

'Manolis knows (how) to swim'/ but \*'Manolis knows how to swim Christina tomorrow.'30

The basic idea is that clauses allowing both PRO and a case-marked subject show discourse properties and they are analysed as strong C\*Ps; conversely, if they can only contain a controlled PRO and show no discourse properties it is because they are simple weak CPs. This is reminiscent of Chomsky's (2006, 2008) distinction between weak  $\nu$ P and strong  $\nu$ \*Ps and can be extended to all phases in all languages.

Following this line of reasoning, it will be optimal if all phasal heads are classified as weak or strong, hence making more prominent the strict parallelism that Chomsky advocates for. In this vein, DPs are strong phases (hence D\*Ps) when certain discourse-related properties intersect. If a DP is a strong phase it does block sub-extraction. What is crucial in this approach is that discourse features are relevant to decide whether a given category is a strong phase or not. Accordingly, it seems that discourse properties such as Definiteness and D-linking are in charge of turning DPs into D\*Ps.

The relevance of LF-related features for the phasehood of DPs is given a full account in Heck, Müller & Trommer (2008). They show that DPs may be phases in Scandinavian (Swedish and Danish) due to the presence of a Definiteness feature. For these linguists the [+ Def] feature is

<sup>&</sup>lt;sup>30</sup>The translations have been taken directly from Sevdali (2009). As noted by an anonymous reviewer, the ungrammatical translation can be built in as \*'Manolis knows how Christina to swim tomorrow.' Otherwise, the ungrammaticality may be thought to be caused by an improper transitive/causative use of *swim*. Greek subjects can stay *in situ* and this is presumably what Sevdali's gloss is intended to show.

sheltered under N, not under D. They assume that all DPs are phases and in order for the [+ Def] feature to be visible for computation, they propose feature movement to the edge of the DP. What is important for my analysis is the increasingly interface-driven character of syntactic operations (Grohmann 2008), since Definiteness and D-linking, two LF properties play a crucial role in the syntactic computation of sub-extraction.

Let me now illustrate how the notion of D\*P relates to extractability possibilities. In Chomsky's (2008) system, weak phases do not count for the purpose of Spell-Out or the Phase Impenetrability Principle in that a probe/goal Agree relation may be established between an external probe and any material in the complement of the weak phase. If this is on the right track, we have enough theoretical apparatus to explain why all cases of sub-extraction from a definite/non-D-linked DP are barred: they are strong phases and as such the complement has already been transferred to the interfaces so that a wh-operator in the complement of a D\*P cannot be targeted by C. On the other hand, weak phases are not Spell-Out domains. Therefore, non-definite/D-linked DPs are only weak DPs and subextraction of the wh-constituent is permitted, given that by the time this undergoes movement to Spec-CP it has not been transferred yet to the semantic and phonological components. If the distinction between DP/D\*P is on track, the grammaticality of English and Spanish sentences in (52) can be easily accommodated.

```
(57) a. Of which singer do you think that some pictures have shocked the audience?
```

```
b.¿De qué cantante te parece que algunas fotos han of which singer CL-2SG seem-PRES.3SG that some photos have-PERF.3PL escandalizado a la audiencia? shocked to the audience
```

'Of which singer does it seem to you that some photos have shocked the audience?'

The extraction domains in these two sentences are the DP some pictures of which singer and algunas fotos de qué cantante respectively. These DPs contain two features which are at stake when C comes to probe the internal wh-operators, namely [- def] and [D-linked]. Consequently, the DPs are only weak phases so that the [wh]-feature in their complement domain is

visible. This allows the edge feature under C to attract the *wh*-operator, yielding a grammatical result (Pesetsky & Torrego 2001).

To see more clearly how the derivation of cases of sub-extraction is drawn, let me make my proposal more explicit. The DP *algunas fotos de qué cantante* 'some pictures of which singer' has the featural structure informally drawn in (58), which is the starting point for the derivation of (57b):

First, the whole DP is moved into Spec-TP to satisfy the [EPP] feature. Once the matrix C is merged to TP, C probes and searches for a suitable goal in order to establish the AGREE relation. The [D-linked Wh] feature under the QP is visible at the CP cycle since the whole DP contains the feature [- def]. Recall that this DP is just a weak phase due to the [- def] feature. Accordingly, C may have access into this DP and agrees with the [D-linked] wh-feature. The edge feature (EF) in C triggers movement of the PP de qué cantante 'of which singer' to its specifier.

(59) 
$$[CP]$$
 de qué cantante  $[CP]$  TP ...  $[DP]$  algunas fotos  $[CP]$   $[DP]$   $[DP]$ 

When transferred to the semantic component, all the non-interpretable features have already been deleted and the derivation is assigned the right semantic interpretation.

By contrast, in (60a) the DP *las fotos de qué cantante* 'the pictures of which singer', although being marked as D-linked, contains a [+ def] feature which renders the whole DP impenetrable due to the fact that its phasehood has been strengthened. Yet, in case that the *wh*-operator lacks the [D-linked] feature, the outcome is even more degraded. This is illustrated in (60b). The reason again is found in the stronger phasehood of the corresponding DP.

```
(60) a. ???¿De qué cantante te parece que las fotos han of which singer CL-2SG seem-PRES.3SG that the photos have-PERF.3PL escandalizado a la audiencia? shocked to the audience

'Of which singer does it seem to you that the photos have shocked the audience?'
```

b. \*¿De quién te parece que las fotos han escandalizado of whom CL-2SG seem-PRES.3SG that the photos have-PERF.3PL shocked

a la audiencia?to the audience

'Of whom does it seem to you that the photos have shocked the audience?'

Concentrating on (60b), the DP subject *las fotos de quién* qualifies as a strong DP phase (hence D\*P), since it contains a [+ def] feature. This precludes any DP-internal feature from being visible for an outside probe. Consequently, the uninterpretable [wh-feature] in the matrix C remains unvalued and the derivation crashes because not all features may be interpreted in the semantic component.

As regards the representational/derivational nature of subject islands, the intuition is that both narrow-syntax and interface properties are crucial when treating a DP as a weak or strong phase. On the one hand, since the discourse-related features of Definiteness and Discourse-linking are conceived of as already present in the lexical array, they influence the computation of the relevant construction. From this it follows that a DP is opaque to sub-extraction if the features at issue make the DP a strong phase. In clear contrast, the DP is transparent to sub-extraction if the opposite discourse-related features interact and make the DP a weak phase. Recall that Definiteness and D-linking are interpretive features. From this perspective, a weak DP phase is licensed if interpreted correctly at LF.

As mentioned earlier, Chomsky (2008) notes that regardless of the base or surface position, subject-island effects arise when a probe searches for a goal within a phase that has already been transferred to the interfaces. My work lends further support to this claim in that sub-extraction is subject to the distinction between weak and strong DP phases and the degraded cases are explained by using a phase domain that has been already transferred to the other components of grammar, hence blocking any further computation. As a consequence, there remain uninterpretable

features under C which have not been valued, thereby driving the derivation to crash.

My analysis of subject islands as strong phases can be easily extended to other languages, which also gives further credit to it. I have already pointed out that Hungarian and Palauan are languages in which sub-extraction from subjects is licensed. In Italian, instances of sub-extraction are found that confirm the selective nature of DP subjects (Luca Grossi, p.c.):

(61) a. Di che autore credi che molti libri sono stati of which author believe-PRES.2SG that many books are-PERF.3PL been un successo?

a success

'Of which author do you believe that many books have been a success?'

b. *Di che autore credi che molti libri hanno causato tanta* of which author believe-PRES.2SG that many books have-PERF.3PL caused such *polemica?* a scandal

'Of which author do you believe that many books have caused such a scandal?'

In (61) the original subject DP *molti libri di che autore* 'many books of which author' is marked with the features [- def] and [D-linked]. Consequently, the whole DP is only a weak phase. Thus, the matrix C probes the *wh*-expression internal to DP and attracts if to Spec-CP, thereby satisfying the EF. Interestingly, the *wh*-operator may undergo movement to Spec-CP despite having previously moved to Spec-TP in the embedded clause. Again, this situation confirms that sub-extraction is not connected with the base-generated or derived nature of DP subjects.

Although dealing with topicalisation, another type of A'-movement, Broekhuis (2008: 63) points out that in Dutch sub-extraction from subject DPs is perfectly acceptable independently of the syntactic position that they occupy, as shown in the following examples (capitals are indicative of contrast):

- (62) a. Van DEZE fabriek hebben de werknemers gisteren het werk onderbroken. of this factory have the employees yesterday the work interrupted 'Of this factory, the employees interrupted their work yesterday.'
  - b. Van DEZE school hebben alle leerlingen verleden jaar de marathon gelopen. of this school have all the pupils last year the marathon run 'Of this school, all the pupils run the marathon last year.'

Note that the displaced PP is marked as [D-linked], which renders the whole subject DP a weak phase, thereby permitting sub-extraction. Accordingly, Dutch also provides a further argument in favour of my analysis of sub-extraction in terms of phases.

Finally, as brought out to me by Ignacio Bosque (p.c.), relative clauses in Spanish also constitute a good type of construction to test sub-extraction and demonstrate the phasehood of DPs. Chomsky (2008) already analysed cases of relative clauses in cleft-constructions in which sub-extraction seems to be banned (Chomsky's grammaticality judgement):

- (63) a. It was the CAR (not the TRUCK) of which [they found the (driver, picture)]
  - b. \*it was the CAR (not the TRUCK) of which [the (driver, picture) caused a scandal]

Chomsky assumes that the ill-formedness in (63) is due to some subjectisland effect. Notice that in (63a) sub-extraction has taken place out of a DP object, whereas in (63b) the relative operator has been extracted from a DP subject, thereby yielding an unacceptable outcome. As demonstrated throughout my work, subject extraction is licensed in *wh*-constructions in other languages. One sub-type of *wh*-construction is the relative clause. In this respect, Stepanov (2007: 92) observes that in Turkish sub-extraction is licit in relative clauses:

(64) a.  $[Op_i[Ahmet-in \ t_i \ git-me-si]-nin \ ben-i \ \ddot{u}z-d\ddot{u}-\ddot{g}-\ddot{u}]$  ev. Ahmet-GEN go-INF-AGR-GEN I-ACC sadden-PAST-COMP-AGR house Lit. 'The house [which [that Ahmet went to \_ ] saddened me].'

b. [Op<sub>i</sub> [pro [[t<sub>i</sub> anne-si]-nin herkes-le konuş-tu-ğ-u]-nu-m mother-AGR-GEN everyone-with talk-PAST-COMP-AGR-ACC

duy-du-ğ-um] adam. hear-PST-COMP-AGR man

Lit. 'The man [whose I heard [that [ \_ mother] talked to everyone]].'

In Spanish, sub-extraction out of a subject has been shown to yield acceptable structures in *wh*-interrogatives. As regards relative clauses, we should expect the same results. Sub-extraction of the relative operator out of the subject of the relative clause gives rise to sentences that are grammatically perfect (Bosque's grammaticality judgement):

- (65) a. La actriz de la que han causado varias fotos una gran polémica the actress of whom have-PERF.3PL caused several pictures a huge scandal
  - b. La actriz de la que varias fotos han causado una gran polémica the actress of whom several pictures have-PERF.3PL caused a huge scandal 'the actress of whom several pictures have caused a huge scandal'
- (66) a. La actriz de la que parece que han causado varias fotos the actress of whom seem-PRES.3SG that have-PERF.3PL caused several pictures una gran polémica
  - a huge scandal
  - b. *La actriz de la que parece* que varias fotos han causado the actress of whom seem-PRES.3SG that several pictures have-PERF.3PL caused *una gran polémica* 
    - a huge scandal

'the actress of whom it seems that several pictures have caused a huge scandal'

In both DPs the relative operator has been moved out of DP subject which is marked as [- def]. This subject may follow the verb, as in (65a) and (66a). In that case, the subject remains *in situ*. However, it can also precede the verb, in which case it undergoes movement to Spec-TP and it is at this stage that sub-extraction of the operator takes place. If the DP subject is marked as [+ def], sub-extraction is blocked. The reason is that in that case the DP is a phase and the [wh]-feature of the relative operator is too deeply inside the phase as to be the goal of an outside probe.

- (67) a. \*La actriz de la que han causado las fotos una gran polémica the actress of whom have-PERF.3PL caused the pictures a huge scandal
  - b. \*La actriz de la que las fotos han causado una gran polémica the actress of whom the pictures have-PERF.3PL caused a huge scandal 'the actress of whom the pictures have caused a huge scandal'
- (68) a. \*La actriz de la que parece que han causado las fotos the actress of whom seem-PRES.3SG that have-PERF.3PL caused the pictures una gran polémica
  - a huge scandal
  - b. \*La actriz de la que parece que las fotos han causado the actress of whom seem-PRES.3SG that the pictures have-PERF.3PL caused una gran polémica
    - a huge scandal

'the actress of whom it seems that the pictures have caused a huge scandal'

As is evident, the syntax of relative clauses also supports my proposal that LF-related features determine the phasal status of DPs.

## 8. Conclusions

In this work I have focused on the interaction of discourse-related features such as Definiteness and Discourse-Linking as the basis to render a DP a strong phase and account for the subject-island effects which arise under certain circumstances. I have proved that sub-extraction is licit when a subject DP is a weak phase, regardless of the syntactic position it occupies. In such a situation, C may penetrate down to the phase domain to probe the *wh*-operator and agree with it. Then, the edge feature under C attracts the *wh*-operator to Spec-TP. Nevertheless, when the subject is marked as definite and non-D-linked, it turns into a strong D\*P, thereby blocking sub-extraction since the [wh]-feature is not visible for C to establish an agreement relation. From this viewpoint, subject islands emerge in the narrow syntax. However, due to the specific interpretive properties a given DP is also processed as an island in the interfaces. Data from Spanish, English, Italian, Hungarian and Dutch favour my phase-based approach to subject islands.

In addition, my analysis also supports Chomsky's (2008) view that there is a strict parallelism among all phases since on a par with CP and  $\nu$ P, DP may be a weak or strong phase.

## References

- Aarts, Bas (1992) *Small Clauses in English: The Nonverbal Types*. Berlin/New York: Mouton de Gruyter.
- Abney, Steven (1987) The English Noun Phrase in its sentential aspect. PhD Dissertation, MIT.
- Bianchi, Valentina & Chesi, Cristiano (2008) Phases, strong islands, and computational nesting. *Studies in Linguistics* 2: 81–100.
- Boeckx, Cedric (2003) *Islands and Chains: Resumption as Stranding*. Amsterdam: John Benjamins.
- —— (2008) Islands. *Language and Linguistics Compass* 2(1): 151–167.
- Boeckx, Cedric & Grohmann, Kleanthes K. (2007) Putting phases in perspective. *Syntax* 10(2): 204–222.
- Boeckx, Cedric & Lasnik, Howard (2006) Intervention and repair. *Linguistic Inquiry* 37(1): 150–155.
- Broekhuis, Hans (2008) Extraction from subjects: some remarks on Chomsky's *On phases*. In Hans Broekhuis, Riny Huybregts, Norbert Corver, Ursula Kleinhenz & Jan Koster (eds.), *Organizing Grammar: Studies in Honor of Henk van Riemsdijk*, pp. 59–68. Berlin/New York: Mouton de Gruyter.
- Cardinaletti, Anna (2004) Towards a cartography of subject positions. In Luigi Rizzi (ed.), *The Structure of CP and IP: The Cartography of Syntactic Structures*, vol. 2, pp. 115–165. New York: Oxford University Press.
- Ceplova, Markéta (2001) Minimalist islands restricting P-features. Unpublished Ms., MIT.
- Chomsky, Noam (1973) Conditions on transformations. In Stephen R. Anderson & Paul Kiparsky (eds.), *A Festschrift for Morris Halle*, pp. 232–286. New York, NY: Holt, Rinehalt, and Winston.
- —— (1986) Barriers. Cambridge, MA: MIT Press.
- (2000) Minimalist inquiries: The framework. In Robert Martin, David Michaels & Juan Uriagereka (eds), *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*, pp. 89–155. Cambridge, MA: MIT Press.
- —— (2001) Derivation by phase. In Michael Kenstowicz (ed.), *Ken Hale: A Life in Language*, pp. 1–50. Cambridge, MA: MIT Press.
- —— (2004) Beyond explanatory adequacy. In Adriana Belletti (ed.), *Structures and Beyond: The Cartography of Syntactic Structures*, vol. 3, pp. 104–131. New York: Oxford University Press.
- —— (2006) Approaching UG from below. Unpublished Ms., MIT.

- (2008) On phases. In Robert Freidin, Carlos Otero & María Luisa Zubizarreta (eds.), *Foundational Issues in Linguistic Theory*, pp. 133–166. Cambridge, MA: MIT Press.
- Cinque, Guglielmo (1990) Types of A'-Dependencies. Cambridge, MA: MIT Press.
- Davies, William & Dubinsky, Stanley (2003) On extraction from NPs. *Natural Language & Linguistic Theory* 21: 1–37.
- den Dikken, Marcel (2007) Phase extension: Contours of a theory of the role of head movement in phrasal extraction. *Theoretical Linguistics* 33: 1–42.
- Diesing, Molly (1992) Indefinites. Cambridge, MA: MIT Press.
- É. Kiss, Katalin (1987) Configurationality in Hungarian. Dordrecht: D. Reidel.
- Enç, Murvet (1991) The semantics of specificity. Linguistic Inquiry 22: 1–26.
- Falco, Michelangelo (2007) The licensing of pronominal features in WCO and OPC. In Luis Eguren & Olga Fernández Soriano (eds.), *Coreference, Modality, and Focus*, pp. 19–44. Amsterdam/Philadephia: John Benjamins.
- Fiengo, Robert & Higginbotham, James (1981) Opacity in NP. *Linguistic Analysis* 7: 395–421.
- Fox, Danny & Lasnik, Howard (2003) Successive cyclic movement and island repair: The difference between sluicing and VP ellipsis. *Linguistic Inquiry* 34: 143–154.
- Frascarelli, Mara (2006) Phases of Interpretation. Berlin: Mouton de Gruyter.
- Frazier, Lyn & Clifton, Charles (2002) Processing "d-linked" phrases. *Journal of Psycholinguistic Research* 31(6): 633–659.
- Gallego, Ángel (2007) Phase theory and parametric variation. PhD dissertation, Universitat Autònoma de Barcelona.
- Gallego, Ángel & Uriagereka, Juan (2006) Sub-extraction from subjects: A phase-based theory. Ms., Universitat Autònoma de Barcelona/University of Maryland.
- (2007) Conditions on sub-extraction. In Luis Eguren & Olga Fernández Soriano (eds.), *Coreference, Modality, and Focus*, pp. 45–70. Amsterdam/Philadephia: John Benjamins.
- Georgopoulos, Carol (1991) Syntactic Variables: Resumptive Pronouns and Binding in Palauan. Dordrecht: Kluwer.
- Giorgi, Alessandra & Longobardi, Giuseppe (1991) *The Syntax of Noun Phrases*. Cambridge: Cambridge University Press.
- Giusti, Guiliana (2006) Parallels in clausal and nominal peripheries. In Mara Frascarelli (ed.), *Phases of Interpretation*, pp. 163–186. Berlin: Mouton de Gruyter.
- Grohmann, Kleanthes K. (2008) Copy modification and the architecture of the grammar. Paper presented at the Mediterranean Syntax Meeting 2, Boğaziçi University.
- Heck, Fabian; Müller, Gereon & Trommer, Jochen (2008) A phase-based approach to Scandinavian definiteness marking. In Charles B. Chang & Hannah J. Haynie (eds.), *Proceedings of the 26th West Coast Conference on Formal Linguistics*, pp. 226–233. Somerville, MA: Cascadilla Proceedings Project.
- Hicks, Glyn (2009) *The Derivation of Anaphoric Relations*. Amsterdam/Philadephia: John Benjamins

- Hiraiwa, Ken (2005) Dimensions of symmetry in syntax: Agreement and clausal architecture. PhD dissertation, MIT.
- Hong, Soo-Min & An, Dong-Hwan (2007) On subject extraction: Parallel movement and chain uniformity. *Journal of Language Sciences* 14(3): 159–171.
- Hornstein, Norbert; Lasnik, Howard & Uriagereka, Juan (2007) The dynamics of islands: Speculations on the locality of movement. *Linguistic Analysis* 1(2): 149–175.
- Hornstein, Norbert & Uriagereka, Juan (2002a) Reprojection. In Samuel David Epstein & T. Daniel Seely (eds.), *Derivation and Explanation in the Minimalist Program*, pp. 106–132. Malden, MA: Blackwell.
- (2002b) Labels and projections: A note on the syntax of quantifiers. In Juan Uriagereka (ed.), *Derivations: Exploring the Dynamics of Syntax*, pp. 115–135. London: Routledge.
- Huang, Cheng-Teh James (1982) Logical relations in Chinese and the theory of grammar. PhD dissertation, MIT.
- Kayne, Richard S. (1983) Connectedness. Linguistic Inquiry 14: 223–249.
- —— (1984) *Connectedness and Binary Branching*. Dordrecht: Foris.
- Kuno, Susumu (1973) Constraints on internal clauses and sentential subjects. *Linguistic Inquiry* 4: 363–385.
- Lasnik, Howard (2001) When can you save a structure by destroying it? In Minjoo Kim & Uri Strauss (eds.), *Proceedings of North East Linguistic Society* 31, pp. 301–320. Amherst, MA: GLSA.
- Lasnik, Howard and Myung-Kwan Park (2003) The EPP and the subject condition under sluicing. *Linguistic Inquiry* 34: 649–660.
- Lasnik, Howard & Saito, Mamoru (1992) *Move α: Conditions on its Application and Output.* Cambridge, MA: MIT Press.
- Lasnik, Howard & Stowell, Tim (1991) Weakest crossover. *Linguistic Inquiry* 22: 687–720.
- Legate, Julie (2003) Some interface properties of the phase. *Linguistic Inquiry* 34(3): 506–516.
- Levin, Beth & Rappaport Hovav, Malka (1995) *Unaccusativity at the Syntax-Lexical Interface*. Cambridge, MA: MIT Press.
- Manzini, M. Rita (1992) *Locality: A theory and some of its empirical consequences*. Cambridge, MA: MIT Press.
- —— (1998) A minimalist theory of weak islands. In Peter Culicover & Louise McNally (eds.), *Syntax and Semantics 29: The Limits of Syntax*, pp. 185–209. San Diego: Academic Press.
- Matushansky, Ora (2005) Going through a phase. In Martha McGinnis & Norvin Richards (eds.), *Perspectives on Phases*. MIT Working Papers in Linguistics 49. MITWPL.
- Mendikoetxea, Amaya (1999) Construcciones inacusativas y pasivas. In Ignacio Bosque & Violeta Demonte (eds.), *Gramática Descriptiva de la Lengua Española*, pp. 1575–1629. Madrid: Espasa.

- Merchant, Jason (2001) The Syntax of Silence. Oxford: Oxford University Press.
- Nunes, Jairo & Uriagereka, Juan (2000) Cyclicity and extraction domains. *Syntax* 3: 20–43.
- Ordóñez, Francisco (1998) Post-verbal asymmetries in Spanish. *Natural Language and Linguistic Theory* 16: 313–346.
- —— (2005) Two Specs for postverbal subjects: Evidence from Spanish and Catalan. Unpublished Ms, SUNY Stony Brook.
- Ormazábal, Javier (1991) Asymmetries on Wh-movement and some theoretical consequences. Unpublished Ms, University of Connecticut.
- Ortega-Santos, Iván (2008) Projecting Subjects in Spanish and English. PhD dissertation, University of Maryland.
- Pesetsky, David (1987) Wh-in-situ: Movement and unselective binding. In Eric Reuland & Alice ter Meulen (eds.), *The Representation of (In)definitess*, pp. 98–129. Cambridge, MA: MIT Press.
- Pesetsky, David & Torrego, Esther (2001) T-to-C movement: Causes and consequences. In Michael Kenstowicz (ed.), *Ken Hale: A Life in Language*, pp. 355–426. Cambridge, MA: MIT Press.
- Postal, Paul M. (1998) *Three Investigations of Extraction*. Cambridge, MA: MIT Press. Radford, Andrew (2009) *Analysing English Sentences*. Cambridge: CUP.
- Rizzi, Luigi (2001) Reconstruction, weak island sensitivity and agreement. In Carlo Cecchetto, Gennaro Chierchia & Maria Teresa Guasti (eds.), *Semantic Interfaces: Reference, Anaphora and Aspect*, pp. 145–176. Chicago: CSLI.
- —— (2006) On the form of chains: Criterial positions and ECP effects. In Lisa Cheng & Norbert Corver (eds.), *Wh-Movement: Moving On.* Cambridge, MA: MIT Press.
- Ross, John R. (1967) Constraints on variables in syntax. PhD dissertation, MIT. [Published 1986 as *Infinite Syntax!* Norwood, NJ: Ablex.]
- Sabel, Joachim (2002) A minimalist analysis of syntactic islands. *The Linguistic Review* 19: 271–315.
- Sánchez, Liliana (1996) Syntactic structure in nominals: A comparative study of Spanish and Southern Quechua. PhD dissertation, University of Southern California.
- Sevdali, Christina (2009) Control into CPs: When finiteness does not matter. *MIT Working Papers in Linguistics* 57: 238–266.
- Stepanov, Arthur (2007) The end of CED? Minimalism and extraction domains. *Syntax* 10(1): 80–126.
- Storto, Gianluca (2000) On the structure of indefinite possessives. In Brendan Jackson & Tanya Matthews (eds.), *Proceedings of Semantics and Linguistics Theory* X, pp. 1–18. Ithaca, NY: CLC Publications.
- Svenonius, Peter (2004) On the edge. In David Adger, Cécile de Cat & George Tsoulas (eds.), *Peripheries: Syntactic Edges and their Effects*, pp. 261–287. Dordrecht: Kluwer.

- Szabolcsi, Anna & den Dikken, Marcel (2002) Islands. In Lisa Cheng & Rint Sybesma (eds), *The Second State-of-the-Article Book*, pp. 123–241. Berlin: Mouton de Gruyter.
- Takahashi, Daiko (1994) Minimality of movement. PhD dissertation, University of Connecticut, Storrs.
- Ticio, M. Emma (2006) Locality conditions in Spanish DPs. In Timothy L. Face & Carol A. Klee (eds.), *Selected Proceedings of the 8th Hispanic Linguistics Symposium*, pp. 137–153. Somerville, MA: Cascadilla Proceedings Project.
- Torrego, Esther (1985) On empty categories in nominals. Ms., University of Massachusetts.
- —— (1998) The Dependencies of Objects. Cambridge, MA: MIT Press.
- Truswell, Robert (2005) Phase heads, multiple spell-out, and a typology of islands. In Nuria Yáñez-Bouza (ed.), *Proceedings of the Fourteenth Manchester Postgraduate Conference in Linguistics*, pp.125–141. Manchester: Department of Linguistics and English Language.
- Uriagereka, Juan (1988) On government. PhD dissertation, University of Connecticut.
- —— (1999) Multiple spell-out. In Samuel D. Epstein & Norbert Hornstein (eds.), *Working Minimalism*, pp. 251–282. Cambridge, MA: MIT Press.
- Uribe-Etxebarria, Myriam (1994) Interface licensing conditions on negative polarity items: A theory of polarity and tense interactions. PhD dissertation, University of Connecticut.
- Wexler, Kenneth & Culicover, Peter (1981) Formal Principles of Language Acquisition. Cambridge, MA: MIT Press.

## Contact information:

Ángel Jiménez Fernández
Departamento de Filología Inglesa (Lengua Inglesa)
Facultad de Filología
C/ Palos de la Frontera s/n
41004 Sevilla
Spain
e-mail: ajimfer(at)us(dot)es