

Why Questions Break the Residual V2 Restriction (in Basque and Beyond)

Aritz Irurtzun
CNRS-IKER

Contents

1	Introduction	1
2	Standard question and focalization strategies	2
3	Why	12
3.1	The special behavior of <i>why</i> -questions	13
3.2	An early acquired and early set pattern	20
3.3	Antisuperiority effects	22
3.4	Why+focus	25
3.5	Long, short, and clausal construals	28
3.6	Clausal <i>zergatik</i>	31
4	Analysis and discussion	37
5	Conclusions	41

1 Introduction

Why questions –interrogative sentences that inquiry about reasons, causes, and purposes– are particular in many respects. In this paper I address the main properties of *why*-questions in Basque and contrast them with the patterns attested cross-linguistically. I show that there are two main construals (constructions with and without V2) and that they are accompanied by different semantic nuances. Adopting [Shlonsky & Soare’s \(2011\)](#) richly articulated CP, I propose that in *why* questions with V2, the interrogative phrase is first-merged in Spec-ReasonP and then moved successive cyclically,

which is accompanied by movement of the verb (T-to-C movement), whereas in non-V2 constructions the interrogative phrase is externally merged in a very high position, where it is frozen, and takes scope over the whole clause.

The paper is organized as follows: Section 2 offers a brief introduction to the syntax of interrogatives (and foci) in Basque. Section 3 then addresses the particular behavior of *why*-questions in this language from a comparative perspective. In Section 4 I present my analysis of the different construals and finally Section 5 closes the chapter with the conclusions.

2 Standard question and focalization strategies

Basque is both a SOV and a ‘discourse configurational’ language, which means that even if the neutral word order is SOV, alternate word orders are also grammatical, but with a marked information structure. For instance, an informationally neutral statement would have the word order in (1), that is, SOV. An alternative word order such as the SVO of (2), even if grammatical, would be unacceptable as an informationally neutral sentence (this rather corresponds to a focalization over the subject (see below)):

- (1) Jonek ura edan du.
Jon water drink AUX
Jon drank water.
- (2) # Jonek edan du ura.
Jon drink AUX water
Jon drank water.

In a similar vein, subject *wh*-questions cannot maintain the neutral SOV word order (3) and necessarily display adjacency between the interrogative phrase and the verb (4). Otherwise it generates strong ungrammaticality (more so than in Spanish, cf. Dold (2018)):

- (3) * Nork ura edan du?
who water drink AUX
Who drank water?
- (4) Nork edan du ura?
who drink AUX water
Who drank water?

These patterns are generally analyzed as instances of *wh*-movement followed by the verb, which constitutes a ‘residual V2’. ‘Residual V2’ is defined by Rizzi (1996, 64) as “such construction-specific manifestations of I-to-C movement in a language (like English and the modern Romance languages except Rætho-Romansch) which does not generalize the V2 order to main declarative clauses”.

In Basque, this property generalizes to both embedded and matrix interrogative clauses, and just as in (4), in example (5) we observe a leftward position of the interrogative phrase followed by O-V inversion in the embedded clause, and S-V inversion in the matrix clause. Failing to display residual V2 in either embedded (6), matrix (7), or both clauses (8) produces ungrammaticality (see [Iruartzun, 2016](#), for an overview of the syntax of interrogatives):

- (5) [Nork](#) esan du Jonek [edan duela ura]?
 who say AUX Jon drink AUX.C water
 Who did Jon say that drank water?
- (6) *[Nork](#) esan du Jonek [ura edan duela]?
 who say AUX Jon water drink AUX.C
 Who did Jon say that drank water?
- (7) *[Nork](#) Jonek esan du [edan duela ura]?
 who Jon say AUX drink AUX.C water
 Who did Jon say that drank water?
- (8) *[Nork](#) Jonek esan du [ura edan duela]?
 who Jon say AUX water drink AUX.C
 Who did Jon say that drank water?

Alternatively, the whole embedded clause can be fronted, but again this requires adjacency between the interrogative phrase and the verb in the embedded clause, as well as adjacency between the whole embedded clause and the matrix verb (9). This is known as a ‘clausal pied-piping’ construction ([Ortiz de Urbina \(1989\)](#), *et seq.*). Again, failing to render residual V2 in either embedded (10), matrix (11), or both clauses (12) generates ungrammaticality:

- (9) [[Nork](#) edan duela ura] esan du Jonek?
 who drink AUX.C water say AUX Jon
 Who did Jon say that drank water?
- (10) *[Nork](#) ura edan duela] esan du Jonek?
 who water drink AUX.C say AUX Jon
 Who did Jon say that drank water?
- (11) *[Nork](#) edan duela ura] Jonek esan du?
 who drink AUX.C water Jon say AUX
 Who did Jon say that drank water?
- (12) *[Nork](#) ura edan duela] Jonek esan du?
 who water drink AUX.C Jon say AUX
 Who did Jon say that drank water?

Interestingly focalization displays the very same pattern in Basque. As advanced above, focus on the subject necessarily affects the word order and instead of the neutral SOV, SVO is obtained, with residual V2 and adjacency between the focal phrase and the verb (*cf. i.a.* [De Rijk \(1978\)](#)):

- (13) [Peiok]_F edan du ura.
 Peio drink AUX water
 [Peio]_F drank water.

Such a pattern, again, generalizes to embedded clauses and the same restrictions that we saw for interrogatives hold (compare the focalization data in (14-17) with the question data in (5-8):

- (14) [Peiok]_F esan du Jonek [edan duela ura].
 Peio say AUX Jon drink AUX.C water
 Jon said that [Peio]_F drank water.
- (15) * [Peiok]_F esan du Jonek [ura edan duela].
 Peio say AUX Jon water drink AUX.C
 Jon said that [Peio]_F drank water.
- (16) * [Peiok]_F Jonek esan du [edan duela ura].
 Peio Jon say AUX drink AUX.C water
 Jon said that [Peio]_F drank water.
- (17) * [Peiok]_F Jonek esan du [ura edan duela].
 Peio Jon say AUX water drink AUX.C
 Jon said that [Peio]_F drank water.

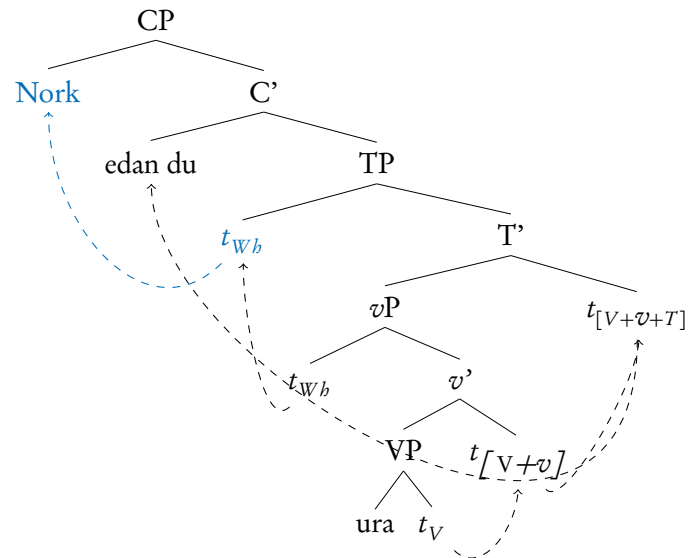
The clausal pied-piping construction also exists for focalizations, with the same restrictions with respect to the necessity of residual V2 (compare (126-21) with (9-12)):

- (18) [[Peiok]_F edan duela ura] esan du Jonek.
 Peio drink AUX.C water say AUX Jon
 Jon said that [Peio]_F drank water.
- (19) * [[Peiok]_F ura edan duela] esan du Jonek.
 Peio water drink AUX.C say AUX Jon
 Jon said that [Peio]_F drank water.
- (20) * [[Peiok]_F edan duela ura] Jonek esan du.
 Peio drink AUX.C water Jon say AUX
 Jon said that [Peio]_F drank water.
- (21) * [[Peiok]_F ura edan duela] Jonek esan du.
 Peio water drink AUX.C Jon say AUX
 Jon said that [Peio]_F drank water.

The standard analysis of *wh*-questions in Basque posits *wh*-movement to Spec-CP, which is then followed by T-to-C movement to check the Q-feature in a Spec-Head configuration –giving rise to the residual V2 configuration (see Ortiz de Urbina (1989) *et seq.*). Thus, sentence (22a) with a *wh*-question on the subject receives the analysis in (22b):

- (22) a. **Nork** edan du ura?
 who drink AUX water
 Who drank water?

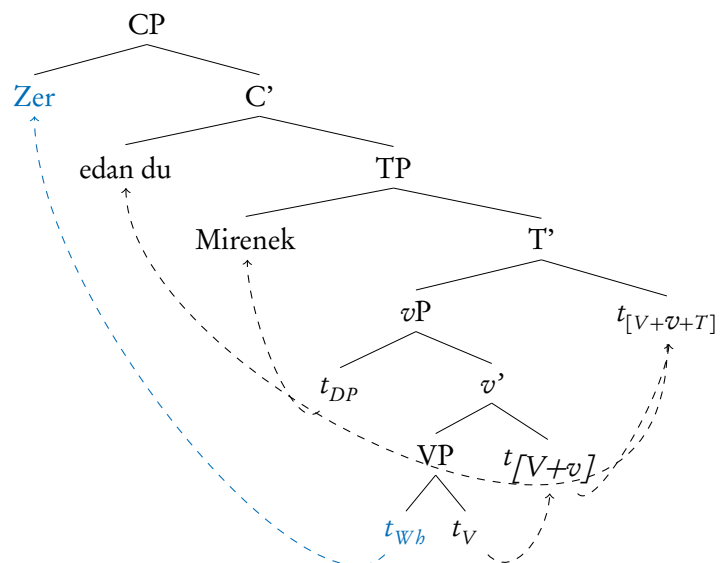
b.



Likewise, an interrogative on the direct object as in (23a), showing V-S inversion would have the structure in (23b):

- (23) a. **Zer** edan du Mirenek?
 what drink AUX Miren
 What did Miren drink?

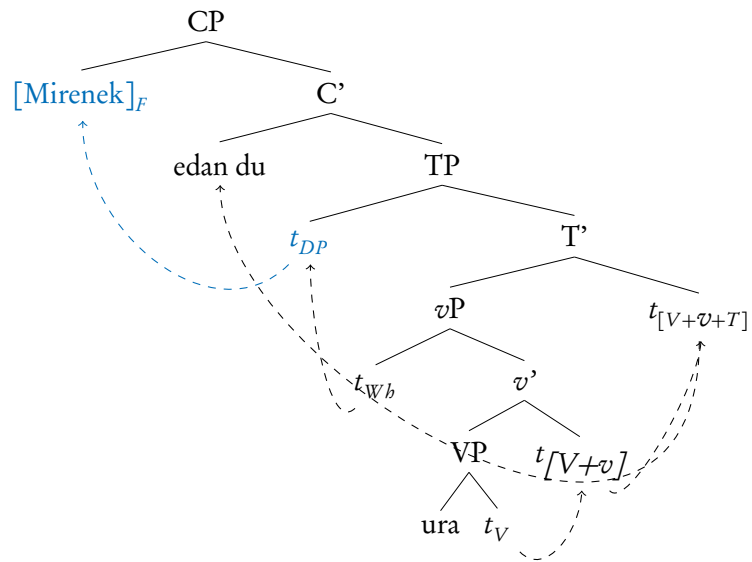
b.



As the patterns for subject (24a) and object (25a) focalizations are analogous to those of *wh*-questions, their corresponding structures are also parallel (compare (24b) and (25b) with (22b) and (23b)):

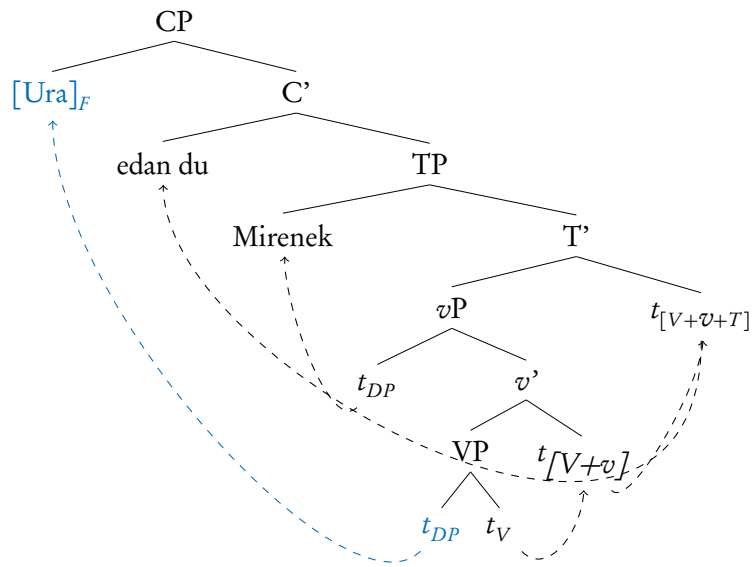
- (24) a. [Mirenek]_F edan du ura.
 Miren drink AUX water
 [Miren]_F drank water.

b.



- (25) a. [Ura]_F edan du Mirenek.
 water drink AUX Miren
 Miren drank [water]_F.

b.

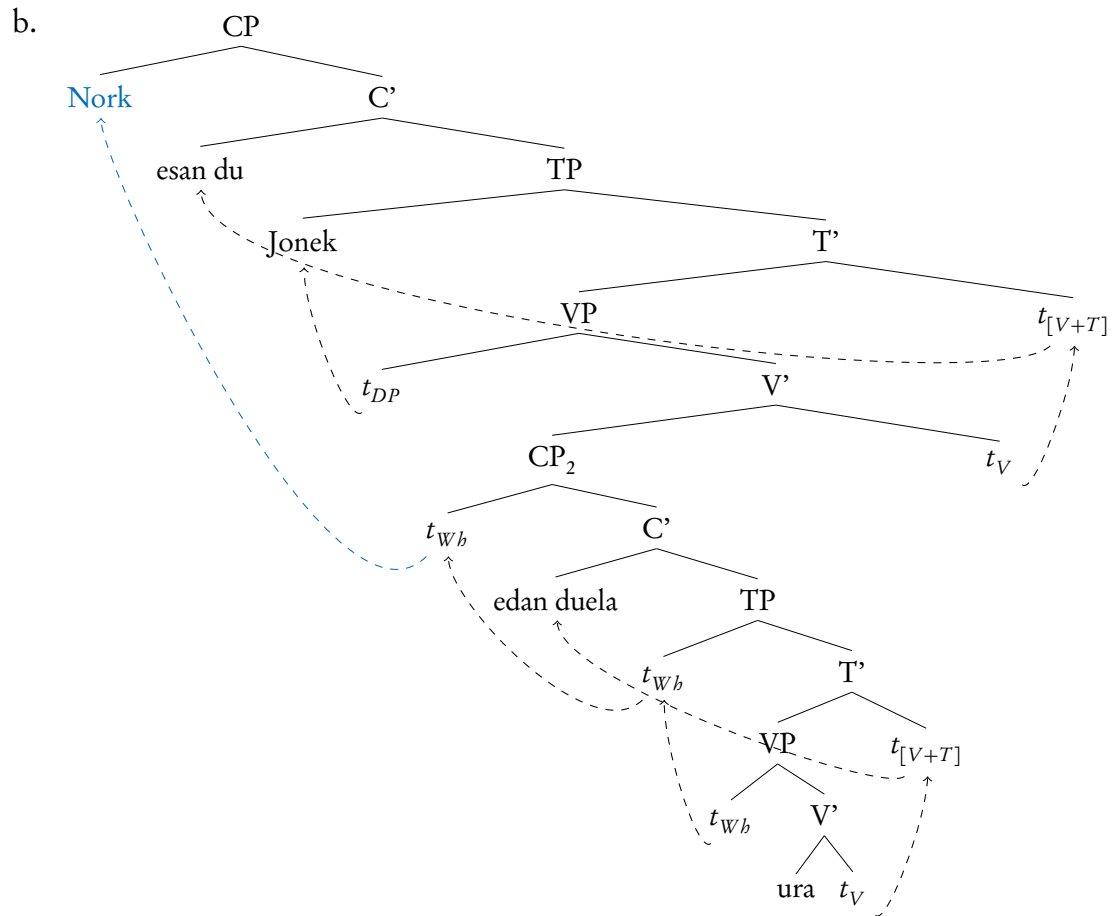


As I advanced, extraction out of embedded clauses takes place the same way successive cyclically, with the consequence that the residual V2 configuration is repeated in both clauses. Therefore, departing from the basic statement in (26), question (27a) on the subject of the embedded clause displays O-V inversion in the embedded clause as well as S-V inversion in the matrix clause, which derives from the structure depicted in (27b): the interrogative phrase undergoes movement to the specifier of the embedded CP first, followed by T-to-C movement of the embedded verb (which renders O-V inversion), and then it is extracted to the specifier of the matrix CP, which triggers again T-to-C movement of the matrix verb, resulting in S-V inversion:¹

- (26) Jonek [Peiok ura edan duela] esan du.
 Jon Peio water drink AUX.C say AUX
 Jon said that Peio drank water.

- (27) a. **Nork** esan du Jonek edan duela ura?
 who say AUX Jon drink AUX.C water
 Who did Jon say that drank water?

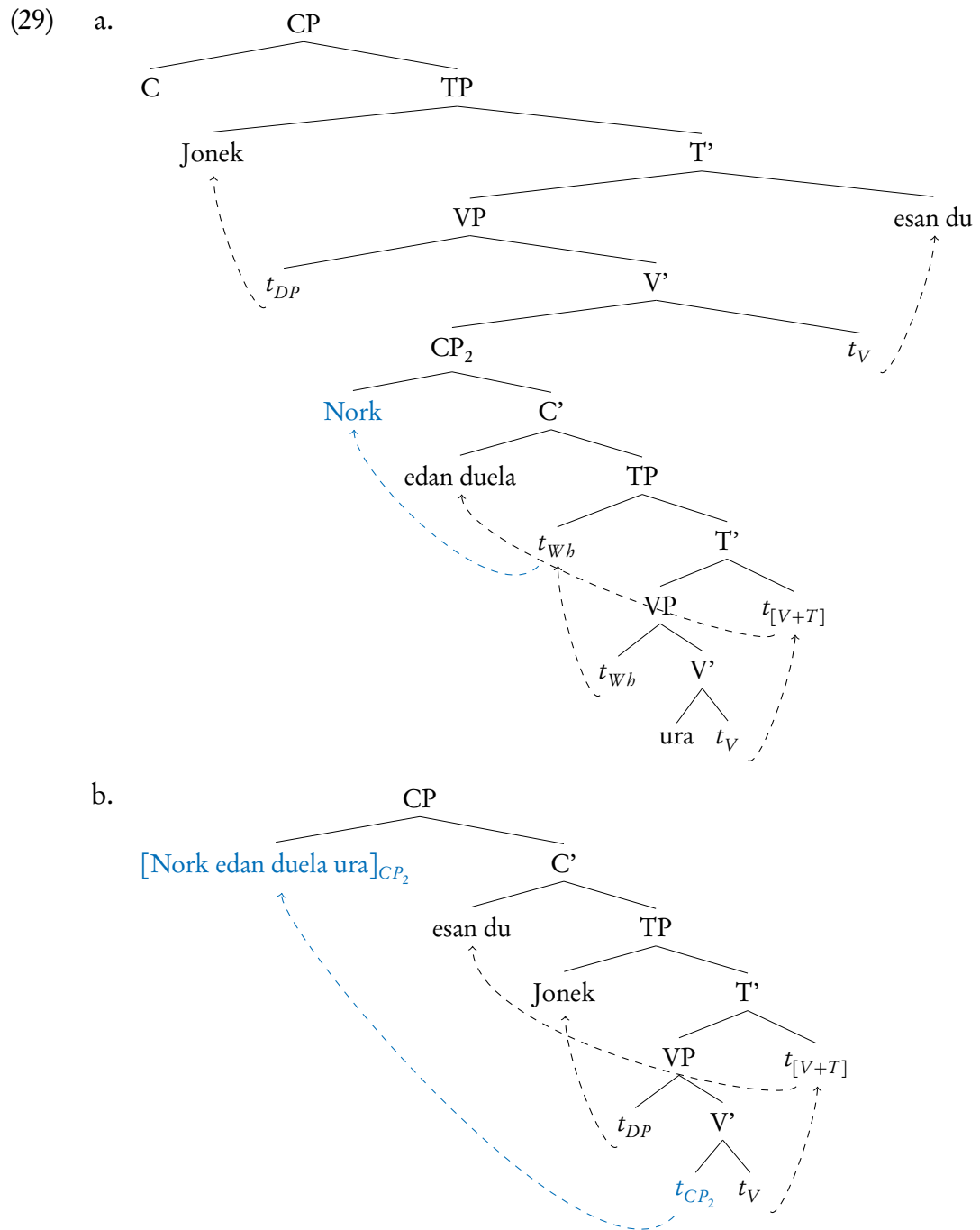
¹Again, the pattern for long-distance focalizations is the same. In the interest of space, I omit such examples and tree-structures.



Finally, the first step of the pied-piping strategy that we saw in (9) –repeated here as (28)– is identical to that of the long-distance extraction: the *wh*-phrase is extracted to Spec-CP of the embedded clause, triggering movement of the verb (29a). However, the second step is different since then the whole embedded clause is extracted to the specifier of the matrix CP, which is followed by movement of the matrix verb to C (rendering the residual V2 effect), (29b):²

- (28) [Nork edan duela ura] esan du Jonek?
 who drink AUX.C water say AUX Jon
 Who did Jon say that drank water?

²Here again I omit the examples and tree-structures for focalizations, as the displacements and syntactic configurations of each construal are identical for *wh*-constructions and focalizations.



Additional evidence in support of this analysis comes from the fact that extraction out of islands such as coordinate structures (30), adjuncts (31), left branches (32), and complex NPs (33) is fully deviant (again, the pattern is the same for both *wh*-constructions and focalizations):

- (30) a. Jonek [salda eta legatza] nahi ditu.
Jon stock and hake want AUX
Jon wants stock and hake.
- b. **Zer* nahi ditu Jonek [salda eta *t*]?
what want AUX Jon stock and
Lit. What does Jon want stock and?
- c. **Zer* nahi ditu Jonek [*t* eta legatza]?
what want AUX Jon and hake
Lit. What does Jon want and hake?
- (31) a. Jon [abestia entzun duelako] poztu da.
Jon song hear AUX.because get.happy AUX
Jon got happy because he heard the song.
- b. **Zer* poztu da Jon [*t* entzun duelako]?
what get.happy AUX Jon hear AUX.because
Lit. What did Jon got happy because he heard?
- c. **[Abestia]_F* poztu da Jon [*t* entzun duelako].
song.ART get.happy AUX Jon hear AUX.because
Lit. Jon got happy because he heard [the song]_F.
- (32) a. Mirenek [Jonen liburua] irakurri du.
Miren Jon.GEN book read AUX
Miren read Jon's book.
- b. **Noren* irakurri du Mirenek [*t* liburua]?
whose read AUX Miren book
Whose book did Miren read?
- c. **[Jonen]_F* irakurri du Mirenek [*t* liburua].
Jon.GEN read AUX Miren book
Miren read [Jon's]_F book.
- (33) a. [Jonek liburu bat idatzi duelako zurrumurrua] entzun duzu.
Jon book one write AUX.C.P rumour hear AUX
You heard the rumour that Jon wrote a book.
- b. **Zer* entzun duzu [Jonek *t* idatzi duelako zurrumurrua]?
what hear AUX Jon write AUX.C.P rumour
Lit. What did you hear the rumour that Jon wrote?
- c. **[Liburu bat]_F* entzun duzu [Jonek *t* idatzi duelako zurrumurrua]
book one hear AUX Jon write AUX.C.P rumour
You heard the rumour that Jon wrote [a book]_F.

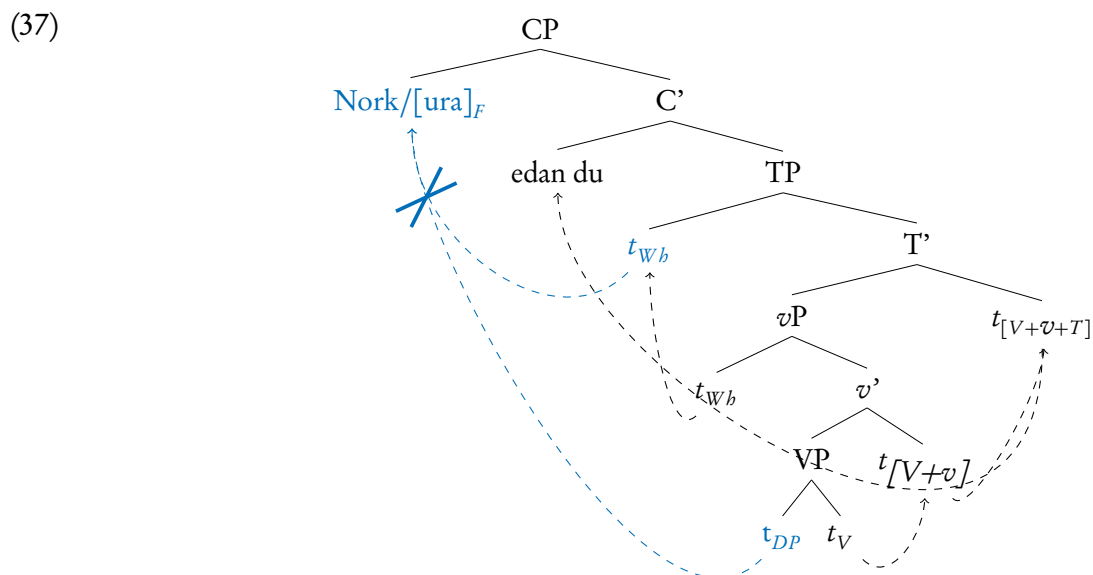
Summarizing, the main characteristics of *wh*-questions and focalizations are the following ones:

- Movement of the focus/ *Wh*-phrase to Spec-CP.
- T-to-C movement.
- Focus/ *Wh*-movement is cyclic.
- The usual restrictions on extraction (islands) apply.

Now, since the syntax of foci and *wh*-questions is uniform, constructions combining a *wh*-question and a focalization are ungrammatical. This is illustrated by the ungrammaticality of both *Wh*»*Foc* (35) and *Foc*»*Wh* (36) with respect to the grammaticality of the simpler *wh*-question in (34):

- (34) **Nork** edan du ura?
 who drink AUX water
 Who drank water?
- (35) * **Nork** [*ura*]_F edan du?
 who drink AUX
 Who drank [*water*]_F?
- (36) * [*Ura*]_F **nork** edan du?
 who drink AUX
 Who drank [*water*]_F?

From the Principles and Parameters model, such pattern has been analyzed as a clash deriving from two elements (the *wh*-phrase and the focal phrase) targeting the same position:



In the next section I will analyze the special behavior of Basque *why*-questions from a cross-linguistic perspective. I will argue that in many languages *why*-questions behave in particular ways with respect to other argument and adjunct questions and, furthermore, that often *why*-questions show a special syntax in patterned ways. In particular, I will show that when fronted they tend to not require the otherwise general adjacency to the verb (hence, no V2) and that they are compatible with foci. I will propose that such patterns derive from a very high merger of the causal interrogative.

3 Why

A longstanding observation is that not all *wh*-phrases necessarily require adjacency to the verb. In particular, researchers such as Mitxelena (1981); Uriagereka (1999) or Etxepare & Ortiz de Urbina (2003) observed that “the operator-verb adjacency is occasionally absent, especially with *zergatik* ‘why’ and other causal *wh*-words” (Etxepare & Ortiz de Urbina, 2003, 465). For instance (38b) is (indeed) grammatical, alongside the general –residual V2– construction of (38a):

- (38) a. *Zergatik* hil zuen zaldunak herensugea?
 why kill AUX knight dragon
 Why did the knight kill the dragon?
- b. *Zergatik* zaldunak herensugea hil zuen?
 why knight dragon kill AUX
 Why did the knight kill the dragon?

That is, in contrast to (38a), in (38b) we do not observe O-V inversion, but contrary to what we saw in examples such as (3) and the following, the result is grammatical with a *why*-question.

There are some small dialectal differences with respect to the morphology of the interrogative element employed for *why*-questions: Southern dialects employ *zergatik* which is composed of the interrogative item *zer* ‘what’ + a motivative postposition, producing a *wh*-item with the value of ‘why’. This element can be reinforced in some varieties with the addition of an inessive marker *-n*, producing *zergatikan* with no apparent change in meaning. In Northern dialects the interrogative item corresponding to *why* is composed of *zer* ‘what’ + the possessive *-ren* + the destinative postposition *-dako*: *ze(re)ndako* = ‘why/for what purpose’.

In Northern dialects (which are the dialects that have *wh-in-situ* alongside *wh*-movement (cf. Duguine & Irurtzun, 2014)) lack of V2 in *why*-questions is more widely available than in Southern dialects. Thus, it is not uncommon to find in Northern Basque examples such as (39) or (40):³

³Example (39) extracted from Jean Etxepare’s *Buruxkak*; example (40) from a tweet by user @lamis-carreb.

- (39) Zendako itxurari ez darraiko bethi izana?
 why appearance.DAT NEG follow always being
 Why doesn't always come the being after the appearance?
- (40) Zendako nere aintzinian jarri zira?
 why me.POSS front.INESS sit AUX
 Why did you sit down in front of me?

3.1 The special behavior of *why*-questions

The particular behavior of *why*-questions is by no means an idiosyncratic property of Basque.⁴ For instance, a classical observation on the literature in French syntax is that French has both *wh in situ* and *wh*-movement questions and while all *wh*-phrases tend to behave similarly in both constructions, *pourquoi* 'why' escapes this general pattern (Kayne, 1972; de Cornulier, 1974; Kayne & Pollock, 1978; Rizzi, 1990; Hamann, 2000). To begin with, the basic word order in French being SVO, in the *wh in situ* strategy *wh*-phrases tend to occupy the same position as the phrase they substitute. See for example the case of a direct object in (41):

- (41) a. Tu veux [un vin]. [French]
 you want one wine
 You want a wine.
- b. Tu veux *quoi*?
 you want what
 What do you want?

The same happens with adjuncts such as *où* 'where':

- (42) a. Tu vas [au restaurant]. [French]
 you go to.ART restaurant
 You are going to the restaurant.
- b. Tu vas *où*?
 you go where
 Where are you going?

However, *pourquoi* cannot appear in the same position as its corresponding phrases (43b); it must obligatorily appear in the left periphery (43c):

- (43) a. Tu verses de l'eau [parce que tu as soif] [French]
 you pour PART ART.water because you have thirst
 You pour water because you are thirsty.
- b. *Tu verses de l'eau *pourquoi*?
 you pour PART ART.water why
 Why do you pour water?

⁴See the contributions in this volume, and Buell's, and Bonan and Shlonsky's in particular.

- c. *Pourquoi* tu verses de l'eau?
 why you pour PART ART.water
 Why do you pour water?

In *wh*-movement questions the general SVO word order is altered with S-V inversion, ending up in a residual V2 configuration, both in argument (44a) and adjunct questions (44b-44c):

- (44) a. *Que* fait Pierre? [French]
 what does Pierre
 What is Pierre doing?
 b. *Comment* votera Pierre?
 how vote.FUT Pierre
 How will Pierre vote?
 c. *Quand* votera Pierre?
 when vote.FUT Pierre
 When will Pierre vote?

However, in questions with *pourquoi*, subject-verb inversion is deviant (45a), and a *pourquoi*-subject-verb order is grammatical (45b).⁵

- (45) a. */[?] *Pourquoi* votera Pierre? [French]
 why vote.FUT Pierre
 Why will Pierre vote?
 b. *Pourquoi* Pierre votera?
 why Pierre vote.FUT
 Why will Pierre vote?

Last, note that this only affects *why*-question *pourquoi* since a question on a segmentally homophonous adjunct such as *pour quoi* 'for what' (seeking an answer such as 'For the abolition of the law') is ungrammatical if it does not involve subject-verb inversion (46a-46b):

- (46) a. *Pour quoi* votera Pierre? [French]
 for what vote.FUT Pierre
 What will Pierre vote for?
 b. **Pour quoi* Pierre votera?
 for what Pierre vote.FUT
 What will Pierre vote for?

The general pattern in Spanish *wh*-questions (an SVO language too) is that *wh*-questions involve *wh*-movement followed by movement of the verb (and hence, residual V2):

⁵Speakers tend to prefer a variant of (45b) with a clitic attached to the verb (*Pourquoi Pierre votera-t-il?*) but since this is optional I kept the bare (45b) for a better comparison with (45a) and (46a).

- (47) a. *Qué* querían esos dos? [Spanish]
 what wanted those two
 What did those two want?
 b. **Qué* esos dos querían?
 what those two wanted
 What did those two want?

But questions with *por qué* ‘why’ can optionally avoid the V2 pattern, and alongside the regular V2 patterns of (48a) and (49a), non-V2 constructions are also grammatical with *por qué* (47b), (48b) (Torrego, 1984; Uriagereka, 1988; Kaiser et al., 2019):⁶

- (48) a. *Por qué* quiere Juan salir antes que los demás? [Spanish]
 why wants Juan go.out before than the rest
 Why does John want to leave before the rest?
 b. *Por qué* Juan quiere salir antes que los demás?
 why Juan wants go.out before than the rest
 Why does John want to leave before the rest?
 (49) a. *Por qué* ha dicho Pedro que Juan quiere a María? [Spanish]
 why AUX say Pedro that Juan loves María
 Why did Pedro say that Juan loves María?
 b. *Por qué* Pedro ha dicho que Juan quiere a María?
 why Pedro AUX say that Juan loves María
 Why did Pedro say that Juan loves María?

Something similar happens in Italian. In general, *wh*-movement has to be accompanied by V2. Hence (50a) with no S-V inversion (no residual V2) is ungrammatical whereas (50b) with adjacency between the *wh*-phrase and the verb is perfectly grammatical:

- (50) a. **Che cosa* Maria ha detto? [Italian]
 what thing Maria AUX said
 What did Maria say?
 b. *Che cosa* ha detto Maria?
 what thing AUX said Maria
 What did Maria say?

The same pattern is observed with adjunct *wh*-phrases such as *dove* ‘where’ (51a) and *come* ‘how’ (51b), which require a residual V2 construal. However, this is not the case with *perché* ‘why’, which appears in a clause initial position as in (51c), not triggering movement of the verb (Rizzi, 1996, 2001):

- (51) a. *Dove* è andato Gianni? [Italian]
 where AUX went Gianni
 Where did Gianni go?

⁶In Spanish orthography it is customary to write clause-initial inverted question marks. However, in order to avoid confusion with acceptability judgements such question marks were omitted here.

- b. *Come* è partito Gianni?
how AUX left Gianni
How did Gianni leave?
- c. *Perché* Gianni è venuto?
why Gianni AUX came
Why did Gianni come?

In New Testament Greek argument questions on direct objects *tí* require *wh*-movement and residual V2 (52a), but *why*-questions with *dià tí* do not, and are perfectly grammatical with a *Wh»Subj* word order (52b-52c) (Kirk, 2012).

- (52) a. è: *tí* dó:sei ánt^h ro:pos antállagma tê:s psukhê:s autoû? [NTG]
or what give man price the soul his
Or what price will a man give for his soul?
- b. *Dià tí* toûto tò múron ouk epráthe:?
why this the ointment NEG sell
Why was this ointment not sold?
- c. *Dià tí* hoi methe:taí sou parabaínousin tê:n parádosin tô:n presbutéro:n?
why the disciple your transgress the teaching the elder
Why do your disciples transgress the teaching of the elders?

Likewise, in Romanian *wh*-phrases tend to require residual V2 construals, as shown with the pair in (53) (Shlonsky & Soare, 2011):

- (53) a. *Cui* i-a cumpărat Ion un CD? [Romanian]
who.DAT him-has bought Ion a CD
For whom did Ion buy a CD?
- b. **Cui* Ion i-a cumpărat un CD?
who.DAT Ion him-has bought a CD
For whom did Ion buy a CD?

However, such a restriction does not hold for *de ce* ‘why’, which allows both V2 (54a) and non-V2 (54b) construals:

- (54) a. *De ce* a cumpărat Ion un CD pentru el? [Romanian]
why has bought Ion a CD for him
Why did Ion buy him a CD?
- b. *De ce* Ion a cumpărat un CD pentru el?
why Ion has bought a CD for him
Why did Ion buy him a CD?

In Syrian Arabic the basic VSO word order is changed into *Wh-V-S* in *wh*-questions in general, as represented in (55), from Sulaiman (2017):

- (55) a. *shw* ĥaka basem? [Syrian Arabic]
what said.3SG.M Basel
What did Bassel say?

- b. **shw* basem haka?
 what Bassel said.3SG.M
 What did Bassel say?

However, the Wh-V adjacency becomes optional in questions with *lesh* ‘why’. Compare the grammaticality of (56b) with the deviance of (55b):⁷

- (56) a. *lesh* tdday?-et mary? [Syrian Arabic]
 why upset-3SG.F Mary
 Why did Mary get upset?
 b. *lesh* mary tdday?-et?
 why Mary upset-3SG.F
 Why did Mary get upset?

The pattern of Singapore English *wh*-questions is a bit different. In this language, *wh*-phrases show optional fronting and may be optionally accompanied by a question particle *ah* in either sentence-final or second position (Yeo, 2010):

- (57) a. You buy *what* ah? [Singapore English]
 you buy what Q
 What did you buy?
 b. *What* you buy ah?
 what you buy Q
 What did you buy?
 c. *What* ah, you buy?
 what Q you buy
 What did you buy?

But *why*-questions show the particularity that *why* has to be obligatorily in the clause-initial position, otherwise, the sentence is clearly deviant (58c):⁸

- (58) a. *Why* John like Mary ah? [Singapore English]
 why John like Mary Q
 Why does John like Mary?
 b. *Why* ah John like Mary?
 why Q John like Mary
 Why does John like Mary?
 c. *John like Mary *why* ah?
 John like Mary why Q
 Why does John like Mary?

Similarly, in Persian in general, *wh*-phrases can remain *in situ* (59a), or move to the focus position (59b-59c) (Kahnemuyipour, 2001; Karimi, 2005):

⁷ Actually, the translation provided for V2 (56a) and non-V2 (56b) in Sulaiman (2017, 328) is ‘What did upset Mary?’ but from the discussion and the context it is clear that it is a *why*-question.

⁸ *how* also behaves similarly in this language.

- (59) a. Kimea diruz ketâb-ro **be ki** dâd? [Persian]
 Kimea yesterday book-râ to who gave
 Who did Kimea give the book to yesterday?
- b. Kimea **be ki** diruz ketâb-ro dâd?
 Kimea to who yesterday book-râ gave
 Who was it that Kimea gave the book to yesterday?
- c. **Be ki** Kimea diruz ketâb-ro dâd?
 to who Kimea yesterday book-râ gave
 Who was it that Kimea gave the book to yesterday?

The only exception to this general pattern is *cherâ* ‘why’, which obligatorily surfaces in the left periphery (even in an echo-reading) (60b-60c):

- (60) a. Ali bâ Maryam ezdevâj kard chon dust-esh dâsht. [Persian]
 Ali with Maryam marry AUX because friend-her AUX
 Ali married Maryam because he loved her.
- b. *Ali bâ Maryam ezdevâj kard **cherâ**?
 Ali with Maryam marry AUX why
 Why did Ali marry Maryam?
- c. Ali **cherâ** bâ Maryam ezdevâj kard?
 Ali why with Maryam marry AUX
 Why did Ali marry Maryam?

In Krachi *wh*-phrases can appear both *in situ* (61a) as well as in the left periphrastic focus position (61b) (Kandybowicz & Torrence, 2012):

- (61) a. ⚭fíw ε-mò **bwatéo momo**? [Krachi]
 woman AGR-kill.PST chicken which
 Which chicken did the woman slaughter?
- b. **Bwatéo momo** jí ⚭fíw ε-mò?
 chicken which FOC woman AGR-kill.PST
 Which chicken did the woman slaughter?

This is general across all *wh*-phrases (*nse* ‘who’, *ne* ‘what’, *ɲfré* ‘where’, *kemeké* ‘when’, *nené* ‘how’...), with the exception of *nání* ‘why’, which unlike the rest, cannot surface in the clause-internal position (62a) and has to be necessarily fronted (62b):

- (62) a. *⚭fíw ε-mò **bwatéo nání**? [Krachi]
 woman AGR-kill.PST chicken why
 Why did the woman slaughter the chicken?
- b. **Nání** jí ⚭fíw ε-mò **bwatéo**?
 why FOC woman AGR-kill.PST chicken
 Why did the woman slaughter the chicken?

In Irish a different externalization pattern suggests a similar underlying structure: in this language, there are two different overt complementizers, *aN* –which is used when Spec-CP is occupied directly by external merge–, and *aL* –which is used when Spec-CP is filled *via* internal merge, *i.e.* movement (McCloskey, 2002, 2003). Thus, for instance, when the clause hosts A'-binding of a moved element, it is headed by *aL* (63a), but when it hosts A'-binding of a resumptive pronoun, it is headed by *aN* (63b):

- (63) a. an ghirseach a ghoid na síogaí [Irish]
 the girl *aL* stole the fairies
 the girl that the fairies stole away.
 b. an ghirseach a-r ghoid na síogaí í
 the girl *aN*-past stole the fairies her
 the girl that the fairies stole away.

Interestingly for our discussion, *why*-questions behave differently with respect to the other *wh*-questions. Whereas argument and adjunct *wh*-questions in general surface with the *aL* complementizer (64a), *why*-questions necessarily surface with *aN* (65):

- (64) a. Cá fhad a bhí tú ann? [Irish]
 how long *aL* be.past you there
 How long were you there?
 b. *Cá fhad a raibh tú ann?
 how long *aN* be.past you there
 How long were you there?
 (65) a. Cad chuige a ndeachaigh tú ann? [Irish]
 why *aN* went you there
 Why did you go there?
 b. *Cad chuige a chuaigh tú ann?
 why *aL* went you there
 Why did you go there?

This pattern extends beyond the simple ‘why’ *cad chuige* to more complex reason questions composed with *cén fáth* ‘what reason’:

- (66) Cén fáth a-r dúirt tú sin? [Irish]
 what reason *aN*-[PAST] say you that
 Why did you say that?
 (67) *Cén fáth a dúirt tú sin?
 what reason *aL* say you that
 Why did you say that?

In Sinhala (Kishimoto, 2018) a specificity of *wh*-adjuncts such as *æi* ‘why’ and *mokə* *də* ‘why’ is that they are not associated with a separable Q particle: *æi* must stand alone (68a), and the Q element *də* appearing in *mokə də* is fixed in a position next to

the *wh*-word, and cannot be separated from it (contrary to what happens with other *wh*-elements). Thus the ungrammaticality of (68a) and (68b):⁹

- (68) a. *Ranjit [Chitra æi də aawe kiɪ] dannəwa. [Sinhala]
 Ranjit Chitra why Q came.E that know.A
 Ranjit knows why Chitra came.
 b. *Ranjit [Chitra moka aawa də] kiɪəla] dannəwa.
 Ranjit Chitra why came.A Q that know.A
 Ranjit knows why Chitra came.

All in all, the broad picture that we obtain from this quick cross-linguistic comparison is that across languages of different families and types *why*-questions display different patterns with respect to the rest of *wh*-questions. In particular, the *wh*-phrase corresponding to *why* tends to be in a higher position and (as a consequence) it tends not to require movement of the verb in the languages that generally have it (residual V2) and/or not to display the characteristic properties of *wh*-movement.^{10,11}

3.2 An early acquired and early set pattern

The relative difference between *why*-questions and other *wh*-questions with respect to the requirement (or not) of residual V2 is an early acquired pattern. For instance, Barreña (1995) reports the following data on the acquisition of Basque:

- (69) a. Hau nok ipini dau? (2;04;24)
 this who put AUX
 Who put this?
 b. Nok apurtu dau holakue? (2;04;24)
 who break AUX like.this
 Who broke the one like this one?
 c. Nun daoɹ nire egurrek? (2;06;05)
 where are my woods
 Where are my woods?

As can be seen, by the age of 2;04 subject questions (69a-69b) as well as locative adjunct questions (69c) trigger movement of the verb attracting it to be right-adjacent

⁹Remarkably, these *wh*-adjuncts are restricted to have short-distance scope.

¹⁰In Khmer too, “The behavior of *mec* ‘how, why’ is comparable to that of the similarly homophonous words in Chinese (Li & Thompson, 1980, 523–4) and other languages. When the word means “how”, it appears wherever manner adverbs are found, but when it means “why” it typically occurs clause initially. (Haiman, 2011, 234). Likewise with Dhao *ngaa tao* ‘why’, which “[u]nlike other interrogatives, *ngaa tao* never occurs in clause final position.” (Balukh, 2020, 133).

¹¹See also the behavior of Zulu *ngani* ‘why’ which surfaces postverbally as other *wh*-phrases, but requires its preceding verb to be appear in disjoint or neutral form, whereas all other *wh*-phrases require a preceding verb in conjoint or neutral form. Buell (2011) builds on this evidence to propose that even if it surfaces postverbally, *ngani* is in the CP area (an Int° head). This contrasts with the rest of *wh*-phrases, which appear inside the VP.

to the *wh*-phrase. This contrasts sharply with the patterns of *why*-questions, which even at much later ages do not necessarily show V2 effects:

- (70) a. **Zeatik** honek jo ein bi dau trena? (2;08;13)
 why this hit do have AUX train
 Why does this one have to hit the train?
- b. **Zegatik** Arantzan zalduna bota dozu? (3;01;12)
 why Arantza.POSS knight throw AUX
 Why did you throw Arantza's knight?
- c. **Zegatik** azeriek untzie jaten dabie? (3;09;11)
 why foxes rabbit eat AUX
 Why do foxes eat rabbit?

Similar differential patterns in acquisition have also been found in other languages such as English (Labov & Labov, 1978; Stromswold, 1990; Berk, 2003; Thornton, 2004, 2008; Conroy & Lidz, 2007), French (Hamann, 2000, 2006), or Korean (Ko, 2006) and Japanese (Ikeda et al., 2019) among others.

Besides, so far I have been mentioning contemporary Basque data, but absence of V2 in *why*-questions is attested across all the history of Basque literature. In Ancient and (Post-)Classical Basque *wh*-movement was only optionally accompanied by V2 (Ortiz de Urbina, 1989; Aldai, 2011; Duguine & Irurtzun, 2014), but by far the *wh*-phrase that appears the less often left-adjacent to the verb (*i.e.* in a V2 construction) is *zergatik*.¹² Observe the examples below from 16th century Lazarraga (71a) and Leizarraga (71b); or the 17th century Axular (71c), Haranburu (71d), or Belapeyre (71e):

- (71) a. **Cegaiti** lauoi bardin on erechi ez deraustegu euren amoreetan?
 why four.DAT equally well consider NEG cause their loves.in
 Why didn't you make the four of them love each other?
- b. **Cergatic** haur hunela blasphemio erraiten ari da?
 why this thus blasphemy saying PROG AUX
 Why is this one saying blasphemies like that?
- c. **Cergatic** bada gorputceco eritasunagatic eguiten duçuna, eztuçu
 why then body.from sickness.for do AUX.REL.ART NEG.AUX
 arimacoagatic eguiten?
 soul.for do.FUT
 Why then won't you do for the sickness of the soul, that that you do for the sickness of the body?

¹²For instance, studying a sample from the New Testament translations by Leizarraga (1571), Haraneder (c. 1740) and Etcheandey (1999) (the Gospels of Matthew and John), Aldai (2011) finds 14 examples of *wh*-V non-adjacency *vs.* 7 examples of adjacency with *zergatik*, but only 1 example of non-adjacency with *zer* *vs.* 15 examples of adjacency, or no examples of non-adjacency *vs.* 5 examples of adjacency with *non* 'where'.

- d. *Cergatic* ni neror bakharric vtzten nauçu?
 why me me.contrast alone leave AUX
 Me, why do you leave me alone?
- e. *Cergatic* egun oroz goiçan, egüerditau, eta arraxen hirourna
 why day all morning.at noon.at and afternoon.at three.each
 cegnu khaldi emaiten dira eliçan?
 sign hit give AUX church.at
 Why are everyday, in the morning, by noon and in the afternoon three
 signs given at the church?

3.3 Antisuperiority effects

A remarkable observation made in the literature on *why*-questions is that they may be subject to ‘antisuperiority effects’ in multiple *wh*-constructions. Such is, for instance, the case of Japanese (cf. *i.a.* Hornstein, 1995; Takita & Yang, 2014). In Japanese, as is well known, *wh*-phrases do not need to front and can remain *in situ*, which renders SOV word order (72a). However, Japanese has scrambling operations that can produce alternative word orders such as OSV, where the object is scrambled over the subject (72b). Both sentences are perfectly grammatical as there is no superiority effect:

- (72) a. *Dare-ga nani-o* katta no? [Japanese]
 who-NOM what-ACC bought Q
 Who bought what?
- b. *Nani-o_i dare-ga t_i* katta no?
 what-ACC who-NOM bought Q
 Who bought what?

However, in multiple *wh*-constructions involving *naze* ‘why’ sentences such as (73a) –where *nani* ‘what’ precedes *naze* ‘why’– are perfectly grammatical, but sentences such as (73b) –where *naze* precedes *nani*– are ungrammatical. Such a restriction is known as an ‘antisuperiority effect’:

- (73) a. *Taroo-ga nani-o naze* katta no? [Japanese]
 Taroo-NOM what-ACC why bought Q
 What did Taroo buy why?
- b. **Taroo-ga naze nani-o* katta no?
 Taroo-NOM why what-ACC bought Q
 What did Taroo buy why?

Korean displays a similar behavior: the word order *why*»*what* is ungrammatical (74a), whereas *what*»*why* is perfectly grammatical (74b) (Jeong, 2003):

- (74) a. **Wae mwues-ul* ne-nun sa-ess-ni? [Korean]
 why what-ACC you-TOP buy-PAST-Q
 Why did you buy what?

- b. Mwues-ul wae ne-nun sa-ess-ni?
 what-ACC why you-TOP buy-PAST-Q
 Why did you buy what?

Tibetan is also a language with scrambling operations whereby *wh*-phrases can surface in different positions. For instance, *garebyadnas* ‘why’ can either precede (75a) or follow (75b) the direct object:

- (75) a. Bkrashis-lags-gi gyag garebyadnas gzigs-gnang-pa-red? [Tibetan]
 Tashi-HON-ERG yak why buy-HON-PAST-AGR
 Why did Tashi buy a yak?
 b. Bkrashis-lags-gi garebyadnas gyag gzigs-gnang-pa-red?
 Tashi-HON-ERG why yak buy-HON-PAST-AGR
 Why did Tashi buy a yak?

But if we substitute the direct object with interrogative *gagi* ‘which/what’, the antisuperiority effect arises and while the order *what*»*why* is grammatical (76a), the order *why*»*what* is not (76b) (Richards, 1997):

- (76) a. Bkrashis-lags-gi gagi garebyadnas gzigs-gnang-pa-red? [Tibetan]
 Tashi-HON-ERG which why buy-HON-PAST-AGR
 Why did Tashi buy what?
 b. *Bkrashis-lags-gi garebyadnas gagi gzigs-gnang-pa-red?
 Tashi-HON-ERG why which buy-HON-PAST-AGR
 Why did Tashi buy what?

The same pattern can be found in Hungarian, where in multiple *wh*-questions *ki* ‘who’ must precede *miért* ‘why’ (Kiss, 2002):

- (77) a. Ki miért hazudott? [Hungarian]
 who why lied
 Who lied why?
 b. *Miért ki hazudott?
 why who lied
 Who lied why?

In Romanian too, *de ce* ‘why’ may co-occur with other *wh*-phrases in multiple *wh*-questions, but always following them, as shown in (78) and (79) (Shlonsky & Soare, 2011):

- (78) a. Cine de ce a plecat? [Romanian]
 who why has left
 Who left and why?
 b. *De ce cine a plecat?
 why who has left
 Who left and why?

- (79) a. Pe cine **de ce** ai întrebat despre accident? [Romanian]
 ACC who why have asked about accident
 Who did you ask about the accident and why?
 b. ***De ce** pe cine ai întrebat despre accident?
 why ACC who have asked about accident
 Who did you ask about the accident and why?

In Basque, we find an analogous behavior: the word order in (80a), where *zer* ‘what’ precedes *zergatik* ‘why’ is grammatical, but the reverse word order, illustrated in (80b) is deviant:

- (80) a. **Zer** erosi duzu **zergatik**?
 what buy AUX why
 Why did you buy what?
 b. ?**Zergatik** erosi duzu **zer**?
 why buy AUX what
 Why did you buy what?

Takita & Yang (2014) provide an analysis of Japanese antisuperiority facts that treat *naze* as a “defective” element and where the antisuperiority effect is an illicit case of feature valuation of C, which is induced by the defective feature specification of *naze*. Extending their analysis to the Basque data is a nontrivial task, since both languages differ in the interrogative strategy employed (*wh in situ* in Japanese; *wh*-movement in Basque) and in the availability of ‘free’ scrambling. I would like to conjecture that the deviance of constructions such as (80b) may be rather due to discursive reasons: first, note that they improve in embedded contexts. Thus, both (81a), with the order *What*»*Why*, and (81b), with *Why*»*What*, are perfectly grammatical:

- (81) a. Esadazu [**zer** erosi duzun **zergatik**].
 tell.me what buy AUX.C why
Lit. Tell me what you bought why.
 b. Esadazu [**zergatik** erosi duzun **zer**].
 tell.me why buy AUX.C what
Lit. Tell me why you bought what.

But furthermore, multiple *wh*-constructions seem to be subject to the D-linking requirement of the leftmost *wh*-phrase (Bolinger, 1978), which is generally harder to satisfy for *why*-questions. In order to show the D-linking requirement of the leftmost *wh*-phrase, Bolinger (1978) provides the following paradigm:

- (82) It’s nice to have all those times scheduled, but when are you doing what?
 (#...but what are you doing when?)
 (83) It’s nice to have all those activities ahead of you, but what are you doing when?
 (#...but when are you doing what?)

That is, when the set that the *wh*-phrase stands for is discursively given, it can be fronted to the leftmost position. Extending this analysis to the Basque data, we can observe that D-linking the *zergatik*-phrase ameliorates substantially the acceptability of *zergatik*»*zer* patterns (compare example (84) with (80b)):

- (84) Gauza horiek guztiak erosteko arrazoi asko eman dituzu, baina *zergatik*
 things those all buy.for reason many give AUX but why
 erosi duzu *zer*?
 buy AUX what
 You gave many reasons for buying all those things, but why did you buy what?

So, all in all, there may not be an absolute ‘antisuperiority’ restriction on multiple *wh*-constructions with *zergatik*, and the pattern observed in Basque and in other languages may be due to discursive factors.

3.4 Why+focus

We saw in Section 2 that in Basque *wh*-questions and focalizations are in general incompatible with each other. This is not a particularity of Basque, but one that has been observed language after language. In Italian, for instance, there is no possible combination of a focalization with a *wh*-question such as *a chi* ‘to whom’, and both *Wh*»*Foc* and *Foc*»*Wh* word orders are ungrammatical (Rizzi, 2001):

- (85) a. * *A chi* [questo]_F hanno detto? [Italian]
 to whom this AUX said
 To whom did they say [this]_F?
 b. * [Questo]_F *a chi* hanno detto?
 this to whom AUX said
 To whom did they say [this]_F?

There is a caveat though, since *why*-questions tend to allow for focalizations, as Rizzi (2001) observes. However, in these combinations *perchè* ‘why’ must precede the focal phrase:

- (86) a. *Perchè* [questo]_F avremmo dovuto dirgli? [Italian]
 why this AUX should say.CL
 Why should we have said [this]_F to him?
 b. * [Questo]_F *perchè* avremmo dovuto dirgli?
 this why AUX should say.CL
 Why should we have said [this]_F to him?

This is in line with what other researchers have observed in other languages. As Partee (1991, 171) puts it, “WHY-questions are focus-sensitive in a way that other WH-questions are not”. But interestingly, language after language the same patterns of combination seem to emerge. In Romanian for instance, *wh*-questions in general cannot be combined with focalizations in any order (Shlonsky & Soare, 2011):

- (87) a. * **Cui** [un CD]_F i-ai cumpărat? [Romanian]
 who.DAT a CD him-AUX bought
 To whom is it a CD that you bought?
 b. * [Un CD]_F **cui** i-ai cumpărat?
 a CD who.DAT him-AUX bought
 To whom is it a CD that you bought?

However, *de ce* ‘why’ can be combined with focalized constituents, but only with the *Why»Foc* order:

- (88) a. **De ce** [un CD]_F ai cumpărat pentru el? [Romanian]
 why a CD AUX bought for him
 Why is it a CD that you bought for him?
 b. * [Un CD]_F **de ce** ai cumpărat pentru el?
 a CD why AUX bought for him
 Why is it a CD that you bought for him?

A similar pattern is also observed in Japanese. Endo (2015) provides the following paradigm combining a *wh*-question and a focalization with *dake* ‘only’:

- (89) a. **Nan-de** [John-dake]_F naiteiru no? [Japanese]
 why John-only crying Q
 Why is only John crying?
 b. ?? [John-dake]_F **nan-de** naiteiru no?
 John-only why crying Q
 Why is only John crying?

As can be seen, the order *Why»Foc* is grammatical (89a), whereas the order *Foc»Why* is deviant (89b).¹³

A similar pattern can be observed in Hungarian, which is known for having a designated preverbal focus position, like in Basque (Kiss, 2002; Horvath, 2013). In this language, both *wh*-movement and focus movement are taken to target the same position (the immediately preverbal one) and hence, their combination produces ungrammaticality (in any of the logically possible orders):

- (90) a. * [Pétert]_F **kinek** mutatad be? [Hungarian]
 Péter.ACC who.DAT showed.2SG PRT
 To whom did you introduce [Péter]_F?
 b. * **Kinek** [Pétert]_F mutatad be?
 who.DAT Péter.ACC showed.2SG PRT
 To whom did you introduce [Péter]_F?

However, *wh*-questions with *miért* ‘why’ can be naturally combined with focalizations, provided they have the *Why»Foc* word order:

¹³See, however, Miyagawa (2017) for discussion on the strength of this evidence.

- (91) a. **Miért** [Péter]_F mutatad be Marinak? [Hungarian]
 why Péter.ACC showed.2SG PRT Mari.DAT
 Why was it [Péter]_F that you introduced to Mari?
- b. * [Péter]_F **miért** mutatad be Marinak?
 Péter.ACC why showed.2SG PRT Mari.DAT
 Why was it [Péter]_F that you introduced to Mari?

Finally, the same pattern appears in New Testament Greek (Kirk, 2012): *why*-questions can be combined with focal elements such as *kaì humeîs* ‘also/even you’, but they always appear in the *Why»Foc* order:

- (92) **Dià tí** kaì humeîs parabáinete tè:n entolè:n toû t^beoû dià tè:n parádosin
 why also you transgress the commandment the God by the tradition
 humô:n? [NTG]
 your
 Why do also you transgress the commandment of God by your tradition?

And the pattern in Basque is exactly the same: the general impossibility of combining *wh*-questions and focalizations disappears with *zergatik*, and complex questions+focalizations can be formed. Such sentences share the same word order pattern that we saw for the other languages whereby the interrogative item has to precede the focus:

- (93) a. **Zergatik** [Peiok]_F eman die albistea?
 why Peio give AUX news
 Why was it Peio that gave them the news?
- b. * [Peiok]_F **zergatik** eman die albistea?
 Peio why give AUX news
 Why was it Peio that gave them the news?

Furthermore, the V2 pattern may not be kept in such constructions and the verb may surface right-adjacent to the focal element (V3, as in (93a)). Nonetheless, placing the verb in the position after *zergatik*, that is, sandwiched between *zergatik* and the focus is also grammatical:

- (94) **Zergatik** eman die [Peiok]_F albistea?
 why give AUX Peio news
 Why was it Peio that gave them the news?

As expected, leaving the verb in the sentence-final position is ungrammatical (as this would go against the general pattern of both *wh*-questions and focalizations):

- (95) * **Zergatik** [Peiok]_F albistea eman die?
 why Peio news give AUX
 Why was it Peio that gave them the news?

Last, note that *zergatik*-questions are also compatible with a particular type of focalization construal whereby instead of the sentence-initial position, the focus appears in the sentence-final position:¹⁴

- (96) a. *Zergatik* esan die Peiok [egia]_F?
 why tell AUX Peio truth
 Why did Peio tell them [the truth]_F?
 b. *Zergatik* esan die egia [Peiok]_F?
 why tell AUX truth Peio
 Why did [Peio]_F tell them the truth?

The fact that *why*-questions can be combined with focalizations makes them context and contrast-sensitive in a way that other *wh*-questions are not, as the philosophical literature has discussed (Partee, 1991; Bromberger, 1993; Cox, 2019). As a matter of fact, question (96a) can be naturally answered with an answer such as “because lying would be problematic”, but not with an answer such as “because he was the only one around at the moment”. On the contrary, (96b) clashes with an answer like “because lying would be problematic”, but it is completely natural with an answer such as “because he was the only one around at the moment”.

In conclusion, as in other languages, the syntax of *why*-questions in Basque seems to be characterized by a couple of particularities. It can behave like any other *wh*-question but (i) in some construals it does not generate the otherwise general residual V2, and (ii) it can be combined with focalizations (always with the *Why»Foc* word order). All this suggests that *zergatik* can be merged very high in the structure, so much so that the position of the verb may not be affected by it. This is what I will explore in the next sections proposing that there are two different construals: one where *why* is merged below the core complementizer area (in SpecReasonP) and then moved to IntP successive cyclically (generating V2 effects), and another one where *why* is directly merged in SpecIntP, where it is frozen, and does not generate any V2 effects.

3.5 Long, short, and clausal construals

Regarding interpretation, the cross-linguistic literature on *why*-questions has uncovered that often *why*-questions can be ambiguous between the so-called short *vs.* long construals (cf. Cattell (1978); Ko (2005); Shlonsky & Soare (2011)). For example, question (97) could be interpreted in either of the two following ways:

- (97) Why did you ask her to resign?
 1. What is the reason *x* such that for *x*, you asked her to resign?
 e.g. Because I didn’t want to just tell her. (Short Construal)

¹⁴These constructions tend to have a reinforced contrastive reading (see Ortiz de Urbina (2002) for discussion and a derivational proposal).

2. What is the reason x , such that you asked her to resign for that particular reason x ?
e.g. I asked her to resign because of her health, not because of her intelligence... (Long Construal)

The explicit syntax that Basque deploys in *wh*-questions helps disambiguating potential ambiguities with respect to short *vs.* long construals. In particular, the cyclicity of verbal movement and clausal pied-piping provides evidence of the extraction site of adjuncts which, in principle, can be extracted out of either the matrix or the embedded clauses. For instance, question (98) with *noiz* ‘when’ is eminently a question over the eventuality described in the matrix clause (*i.e.* *when*→*think*, not *when*→*finish*) since there is verbal movement in the matrix clause (signalled by S-V inversion) but not in the embedded clause (and hence no object-verb inversion):

- (98) *Noiz* pentsatzen du Jon [gerra bukatuko dela]?
 when think AUX Jon war finish AUX.C
 When is it that Jon thinks that the war will finish?

Alternatively, in (99) the question is over the eventuality described in the embedded clause (*i.e.* *when*→*finish*), since there is verbal movement in both clauses, which signals that the extraction of *noiz* took place from the embedded clause:

- (99) *Noiz* pentsatzen du Jon [bukatuko dela gerra]?
 when think AUX Jon finish AUX.C war
 According to Jon, when will the war finish?

Since clausal pied-piping is an alternative to long-distance extraction, it can only arise in questions on the embedded clause (*i.e.* *when*→*finish*), as shown in (100):

- (100) [*Noiz* bukatuko dela gerra] pentsatzen du Jon?
 when finish AUX.C war think AUX
 According to Jon, when will the war finish?

The behavior of *zergatik* is the expected one given this syntax: the extraction site can be tracked down in the word order. Thus, the interpretation of (101a) is that of a short construal (*why*→*think*), in (101b) we have a long construal (*why*→*finish*) and in (101c) a long construal:¹⁵

- (101) a. *Zergatik* pentsatzen du Jon [gerra bukatuko dela]?
 why think AUX Jon war finish AUX.C
 Why is it that Jon thinks that the war will finish?
 b. *Zergatik* pentsatzen du Jon [bukatuko dela gerra]?
 why think AUX Jon finish AUX.C war
 According to Jon, why will the war finish?

¹⁵Even if it is generally judged grammatical as such, example (101c) is more natural with *gerra* topicalized over *zergatik*.

- c. [Zergatik bukatuko dela gerra] pentsatzen du Jonek?
 why finish AUX.C war think AUX Jon
 According to Jon, why will the war finish?

(101a) shows S-V inversion in the matrix clause but no O-V inversion in the embedded clause, as a consequence, its reading is that of a question on the matrix clause (*why*→*think*). (101b) is an instance of long-distance extraction (with V2 in both clauses) and as a consequence its reading is that of a question on the embedded clause (*why*→*finish*). Last, the clausal pied-piping construction of (101c) is also a question on the embedded clause, hence its reading is *why*→*finish*.

Likewise, island structures provide good environments to assess the association between word order and interpretation. As we said in Section 2, *wh*-movement in Basque is impossible out of adjunct clauses. Thus, departing from the base in (102a), *wh*-extraction of the direct object such as in (102b) produces ungrammaticality:

- (102) a. Jon [abestia entzun duelako] poztu da.
 Jon song hear AUX.because get.happy AUX
 Jon got happy because he heard the song.
 b. *Zer poztu da Jon [*t* entzun duelako]?
 what get.happy AUX Jon hear AUX.because
 Lit. What did Jon got happy because he heard?

Such behavior helps explaining the pattern in (103) below. Example (103a) with V-S inversion in the matrix clause and no inversion in the embedded clause is perfectly grammatical because *zergatik* can only be understood as coming from the matrix clause (hence its effect in the movement of the matrix verb, and its lack thereof in the embedded clause). Accordingly, its intended meaning questions why Jon got happy because he heard a song. Contrariwise, movement of the embedded verb over the direct object *abestia* in (103b) could only be due to a residual V2 of the movement of *zergatik* from the embedded clause to successive cyclically move to the specifier of the matrix clause (triggering again movement of the matrix verb). However, such an extraction out of an adjunct clause is illicit, hence the ungrammaticality of (103b):

- (103) a. Zergatik poztu da Jon [*pro* abestia entzun duelako]?
 why get.happy AUX Jon song hear AUX.because
 Why did Jon got happy because he heard the song?
 b. *Zergatik poztu da Jon [entzun duelako *pro* abestia]?
 why get.happy AUX Jon hear AUX.because song
 Jon got happy because why did he hear the song?

Now, regarding the ambiguity that we saw in (97) on the short *vs.* long construal, a Basque variant as in (104a) is also ambiguous between the short and long construal readings (even though the short construal reading is more prominent). This derives from the fact that having *zergatik* in clause-initial position, and then followed by the matrix and the embedded verbs can correspond to extraction from any of the

clauses; both construals provide the same word order. However, a word order such as (104b) can only correspond to a long construal reading, which would be derived *via* the clausal pied-piping operation: first, movement of the *wh*-word to the specifier of the embedded CP, movement of the verb of the embedded clause to C (V2), and then movement of the whole embedded clause to the specifier of the matrix clause, which is again followed by movement of the matrix verb. It is the clausal pied-piping movement that produces the [_{CP} *wh* V] V word order:

- (104) a. *Zergatik* eskatu zenion uko egiteko?
 why ask AUX renunciation do
 Why did you ask her/him to resign?
- b. *Zergatik* uko egiteko eskatu zenion?
 why renunciation do ask AUX
 Why did you ask her/him to resign?

However, there is a third reading that questions such as (104b) can get: a ‘clausal’ reading. In the next sections I will argue that example (104b) can correspond to the clausal-pied piping construction (with its associated reading), but also to a construction where *zergatik* is directly first-merged in the left periphery, and there is no verbal movement whatsoever (thus, similar to the construction that we saw in (38b) and to the ones attested in other languages). In such cases, the interrogative element takes clausal scope and a reason interpretation similar to that of English *how come* or *why is it that*.

3.6 Clausal *zergatik*

The clausal reading, as indicated by the name, does not inquire about the reasons of the VP/*v*P but about the whole TP. Hence it does not question on the motives of the subject/initiator. It rather questions why is it that the whole eventuality described by the sentence took place (thus, it is similar to the meaning of English *how come*).

Imagine the following scenario:

- (105) *Context:* We leave Leire –an infant– with a caregiver for the morning. We know that she is very tired, since she spent a large part of last night awake and playing, but we need the caregiver to keep Leire from getting asleep, otherwise in the afternoon it will be difficult to get her to bed. When we come home at noon, we see that Leire is asleep.

In such a context, it would be pointless to ask a question like (106a), for we know the answer (she got asleep because she did not have enough sleep the night before and she was very sleepy). However, (106b) questions over the whole clause/ eventuality, and therefore it is coherent in this context as a recrimination of how come such an event took place, even if it was meant not to:

- (106) a. # *Zergatik* lokartu da Leire?
 why get.asleep AUX Leire
 Why did Leire get asleep?
 [≈what were the reasons for her sleepiness?]
- b. *Zergatik* Leire lokartu da?
 why Leire get.asleep AUX
 Why did Leire get asleep?
 [how come she got asleep]

So in this pattern we would be in front of a very high *zergatik*, reminiscent of those that have been proposed for other languages.

Besides, regular *why*-questions tend to display ambiguity between a reason and a purpose reading (Stepanov & Tsai, 2008; Chapman & Kučerová, 2016). This is illustrated in example (107), from Chapman & Kučerová (2016):

- (107) Why did she resign?
- a. ✓ *Purpose*: For what purpose did they resign? In order to earn more money next year.
- b. ✓ *Reason*: What was the reason for their resigning? Because they got a pay cut.

In Chapman & Kučerová's (2016) analysis, reason *why*-s are base-generated as TP adjuncts whereas purpose *why*-s are base-generated as adjuncts of CausP (a *v*P functional layer corresponding to agentive predicates having a volitional requirement).¹⁶

The same pattern arises in Basque with the V2 construal of *zergatik*:

- (108) *Zergatik* utzi du lana?
 why leave AUX work
 Why did they resign?
- a. ✓ *Purpose*: For what purpose did they resign? In order to earn more money next year.
- b. ✓ *Reason*: What was the reason for their resigning? Because they got a pay cut.

However, this ambiguity disappears with high *zergatik*, and the only available reading is the *how come* (reason) reading (≈ *Why is it that they resigned?*):

- (109) *Zergatik* lana utzi du?
 why work leave AUX
 How come they resigned?
- a. # *Purpose*: For what purpose did they resign? In order to earn more money next year.

¹⁶This explains the fact that only agentive dynamic predicates allow for both readings; in particular neither passives, unaccusatives, nor locative existential predicates allow for purpose readings.

- b. ✓ *Reason*: What was the reason for their resigning? Because they got a pay cut.

Very often, the clausal and the reason reading are hard to distinguish, as agents/initiators are which bring about the eventualities. However, we will see below that the clausal construction is particularly employed in conjectural or wondering situations such as rhetorical questions that wonder about the state of affairs and they often imply a counter-expectation (see Tsai (2008) for similar evidence).

Likewise, I mentioned in Section 3 that *zergatik* is composed of *zer* ‘what’ + “motivative” *-gatik*. Thus, a construction with *zergatik* can be ambiguous between being construed with the lexicalized *zergatik* ‘why’ or with the syntagmatic *zer-gatik* “what for”. Below I provide some examples that show that regular constructions with V2 can display this ambiguity (the ‘a’ examples in (110) to (112)), thus they can be answered either with causes or with PP complements. However, constructions with high *zergatik* (the ‘b’ examples) cannot; they only accept the causal reading, as shown in (110b-C), (111b-C), and (112b-C).^{17,18}

- (110) a. A. *Zergatik* gaisotu da Miren?
 why/what.for get.sick AUX Miren
 Why did Miren get sick? OR What did Miren get sick for?
 B. Ez babesteagatik.
 not protecting.for
 Because she took no precautions.
 C. Gaixotasun profesionalagatik.
 illness profesional.for
 Of an occupational disease.
- b. A. *Zergatik* Miren gaisotu da?
 why Miren get.sick AUX
 How come Miren got sick?
 B. Ez babesteagatik.
 not protecting.for
 Because she took no precautions.
 C. # Gaixotasun profesionalagatik.
 illness profesional.for
 Of an occupational disease.

¹⁷Some of these examples are substandard and probably calques from Spanish.

¹⁸In Central Basque, ambiguous structures can be disambiguated by means of prosody: *zergátik* with the regular peninitial stress corresponds to the lexicalized “why” whereas *zérgatik* with initial stress to the syntagmatic “what for”.

- (111) a. A. *Zergatik* kondenatu dute Jon?
 why/what.for convict AUX Jon
 Why did they convict Jon? OR What did they convict Jon of?
- B. Bere aurkako frogak zituztelako.
 his against.GEN proofs have.because
 Because they had proofs against him.
- C. Ogasun publikoaren aurkako delituengatik.
 estate public.GEN against.GEN crimes.for
 Of crimes against public finances.
- b. A. *Zergatik* Jon kondenatu dute?
 why Jon convict AUX
 How come they convicted Jon?
- B. Bere aurkako frogak zituztelako.
 his against.GEN proofs have.because
 Because they had proofs against him.
- C. # Ogasun publikoaren aurkako delituengatik.
 estate public.GEN against.GEN crimes.for
 Of crimes against public finances.
- (112) a. A. *Zergatik* aldatu du Peiok oparia?
 why/what.for change AUX Peio present.ART
 Why did Peio change the present? OR What did Peio change the present for?
- B. Ez zitzaiolako gustatzen.
 not AUX.because like
 Because he didn't like it.
- C. Beste batengatik.
 other one.for
 For another one.
- b. A. *Zergatik* Peiok oparia aldatu du?
 why Peio present change AUX
 How come Peio changed the present?
- B. Ez zitzaiolako gustatzen.
 not AUX.because like
 Because he didn't like it.
- C. # Beste batengatik.
 other one.for
 For another one.

Further evidence in favor of the high merger of this *zergatik* is that the word order of an embedded clause cannot be affected by it. Thus, example (113a) is acceptable with initial *zergatik* followed by the rest of the elements of the clause in their base

generated position, that is: S [O V]_{CP} V. However, (113b) with O-V inversion in the embedded clause is not, since this inversion could only be triggered by a previous A'-movement to the specifier of the embedded CP. In (113b) it cannot be the case that *zergatik* originated in the embedded clause, for it would have triggered V2 in both the embedded and the matrix clauses. Thus the ungrammaticality of (113b). Alternatively, if *zergatik* was directly merged in the left periphery of the matrix clause we should not observe O-V inversion in the embedded clause (and hence, we would obtain (113a)):

- (113) a. *Zergatik* Jonek [gerra bukatuko dela] pentsatzen du?
 why Jon war finish AUX.C think AUX
 {Why is it that/How come} Jon thinks that the war will finish?
- b. **Zergatik* Jonek [bukatuko dela gerra] pentsatzen du?
 why Jon finish AUX.C war think AUX
 {Why is it that/How come} Jon thinks that the war will finish?

However, the restriction on verb movement with ‘high’ *zergatik* is not an absolute one; it only holds if the movement can only be derivative of a purported movement of *zergatik* itself (as in the case of the embedded verb in (113b)), which clashes with the idea that *zergatik* was directly introduced in the left periphery. Verb movements are fine, provided that they piggy back on another displacement operation (say, a focalization), as in the case of (114a), where high *zergatik* is combined with a focalization on the subject of the matrix clause (which in consequence shows O-V inversion). Again, (114b) is ungrammatical because besides the O-V inversion of the matrix clause, it also displays O-V inversion in the embedded clause. The former is due to the focalization of the subject (as in (114a)), but the latter could only be due to a *wh*-displacement in the embedded clause that generated the V2. The unavailability of such a movement suggests that in this construction *zergatik* is introduced directly high in the left periphery:

- (114) a. [?]*Zergatik* [Jonek]_F pentsatzen du [gerra bukatuko dela]?
 why Jon think AUX war finish AUX.C
 {Why is it that/How come} Jon thinks that the war will finish?
- b. **Zergatik* [Jonek]_F pentsatzen du [bukatuko dela gerra]?
 why Jon think AUX finish AUX.C war
 {Why is it that/How come} Jon thinks that the war will finish?

Besides focalized constituents, high *zergatik* can also be combined with topics and other intervening material which suggests its very high position in the clause. We already saw an example from Haranburu (17th c.) in (71d), repeated here as (115) where *zergatik* is followed by a reduplicated topical pronoun:

- (115) *Cergatic* ni neror bakharric vtzten nauçu?
 why me me.contrast alone leave AUX
 Lit. {Why is it that/How come}, me, you leave me alone?

But it can also be followed by series of adjuncts, as in example (71e), repeated here as (116) from Belapeyre (17th c.):

- (116) **Cergatic** egun oroz goiçan, egüerditau, eta arraxen hirourna cegnu
 why day all morning.at noon.at and afternoon.at three sign
 khaldi emaiten dira eliçan?
 hit give AUX church.at
Lit. {Why is it that/How come} everyday, in the morning, by noon and in
 the afternoon three signs given at the church?

Also by vocatives, as in example (117), from Uriarte (19th c.):

- (117) **Zergatik**, Jauna, urrutira alde egin dezu, desanparatu nazu denborarik bearrenean,
 why Lord far.to side do AUX abandon AUX time.of need.most.at
 naigabea?
 desperation.at
 Why standest Thou afar off, o Lord? Why hidest Thou Thyself in times of
 trouble?

In contemporary Basque high *zergatik* can even be associated with clauses with overt complementizers, as in the Basque variants of the reportative clauses analyzed by Etxepare (2010) for Spanish. Etxepare studies a set of constructions of Spanish which are characterized by having an overt initial complementizer *que*, which adds a reportative reading to the sentence. Thus, there is a difference between examples (118a) and (118b):

- (118) a. Si viene mi madre, el tabaco es tuyo. [Spanish]
 if comes my mother the tobacco is yours
Lit. If my mother comes, the tobacco is yours.
 b. Si viene mi madre, que el tabaco es tuyo.
 if comes my mother C the tobacco is yours
Lit. If my mother comes, that the tobacco is yours.

Etxepare (2010) imagines the following scenario: two teenagers (A and B) are secretly smoking in a room. Suddenly, fearing that his/her mother could show up and find out, A tells B (118a): *Si viene mi madre, el tabaco es tuyo*. By saying that, A asks B to act as if the tobacco was B's, if A's mother comes. However, by saying (118b), A asks B something more than just pretense: A asks B to *say* that the tobacco is B's. If B doesn't say so, B will not be complying with A's request.

Basque also has analogous constructions (119A), and high *zergatik* can appear with them (119B), as in the following scenario, where *nirea* 'mine' intervenes between *zergatik* and the verb:

- (119) A. Nire ama etortzen ba-da, tabakoa zurea de-la.
 my mother come if-BE tobacco yours BE-C
 If my mother comes, (say) that the tobacco is yours.

- B. *Zergatik* nirea de-la?
 why mine BE-C
 {Why is it that/How come} (I/we should say) that it is mine?

By employing the complementizer *-la* (which surfaces attached to the verb) speaker A asks B to *say* that the tobacco is B's, but again by employing *-la* B complains asking why should she/he say that it is hers/his.

All this bears testimony to the fact that what I called *high zergatik* is very high in the clausal structure.

4 Analysis and discussion

Given that the *wh*-phrase corresponding to the meaning of *why* has such a particular behavior across languages, several authors have proposed that it is directly merged in a high position in the clause (see *i.a.* Hornstein, 1995; Rizzi, 2001; Ko, 2005). More recently, Shlonsky & Soare (2011) propose a richly articulated complementizer structure at the left periphery of the clause, arguing that English *why* (or Romanian *de ce*) is externally merged in Spec-ReasonP and then moved to Spec-IntP whereas *how come* is directly externally merged in Spec-IntP:

(120) IntP > TopP > FocP > WhP > ReasonP...

Building on this richly articulated left periphery, I would like to propose that the differential patterns that we observed for Basque *zergatik* are due to the following two types of construals where *zergatik* participates:

1. *Zergatik* externally merged at Spec-ReasonP (above NegP); then moving to IntP successive cyclically followed by the verb and generating residual V2 effects. This gives rise to the canonical Short and Long construals.¹⁹
2. *Zergatik* directly merged at Int⁰, where it is frozen, takes clausal scope and as a consequence does not generate any verb movement (*i.e.*, high *zergatik*).

In this respect, high *zergatik* is a complementizer with the same syntax as English *how come*. There is no *wh*-movement and therefore no V2 effect in consequence (Collins, 1991).

It is also similar to Basque *nola* 'how', that beyond being a *wh*-adjunct in questions (121) can also serve as a complementizer in embedded clauses, not triggering movement of the verb. Thus, in (121a) we observe *nola*-verb adjacency and we have an embedded manner interpretation (hence the grammaticality of (121b) with *galdetu* 'ask' in the matrix clause, a question-embedding verb):

¹⁹I leave the question open as to whether in purpose questions *zergatik* is introduced lower, as suggested by Chapman & Kučerová (2016).

- (121) a. Begira *nola* estali duen Jonek oparia.
 look how wrap AUX Jon present
 Look how Jon wrapped the present.
- b. Galdetu *nola* estali duen Jonek oparia.
 Ask how wrap AUX Jon present
 Ask how Jon wrapped the present.

In (122a) on the contrary we have complementizer *nola*; it implies a factive-like meaning and does not attract the verb. This complementizer cannot be combined with a question-embedding verb (122b):

- (122) a. Begira *nola* Jonek oparia estali duen.
 look how Jon present wrap AUX
 Observe the fact that Jon wrapped the present.
- b. *Galdetu *nola* Jonek oparia estali duen.
 question how Jon present wrap AUX
 Question the fact that Jon wrapped the present.

Returning to *zergatik*, I mentioned already that high *zergatik* has an interpretation akin to that of English *how come* in that it generates speculative, wondering questions on how come the eventuality described by the clause happened. There is another interesting fact about *how come*: contrary to *why*, it always takes highest scope (cf. Collins (1991)). In (123a) the surface order *why*- \forall is ambiguous between the $Wh \gg \forall$ and the $\forall \gg Wh$ readings; however, (123b) with *how come* can only be interpreted with frozen scope $Wh \gg \forall$:

- (123) a. Why was every candidate elected?
 $Wh \gg \forall$ / $\forall \gg Wh$
- b. How come every candidate was elected?
 $Wh \gg \forall$ / * $\forall \gg Wh$

If as I proposed high *zergatik* is externally merged in the same position as *how come*, the prediction would be that it should have a similar behavior with respect to scopal properties. Indeed, we observe the very same pattern: (124a) with V2 is ambiguous between the $Wh \gg \forall$ and the $\forall \gg Wh$ readings, but (124b) with high *zergatik* only has the $Wh \gg \forall$ reading whereby *zergatik* takes highest scope:²⁰

- (124) a. *Zergatik* etorri dira ikasle hauek guztiak?
 why come AUX student these all
 Why did all these students come?
 [$Wh \gg \forall$ / $\forall \gg Wh$]

²⁰As a matter of fact, high *zergatik* can also be substituted by another *wh*-element with the same meaning that takes highest scope: *nolatan*.

- b. *Zergatik* ikasle hauek guztiak etorri dira?
 why student these all come AUX
 {Why is it that/How come} all these students came?
 [*Wh*»*∀* / **∀*»*Wh*]

Likewise, English *how come* –as opposed to *why*– cannot appear in multiple *wh*-constructions, as represented in (125) (see Ochi (2004) for discussion). The same pattern is attested in Basque with high *zergatik*, which renders ungrammatical results when combined with another interrogative phrase. Compare the grammaticality of ‘regular’ *zergatik* with either multiple fronting in (126a) or with single fronting+*in situ wh* in (126b), with the ungrammaticality of high *zergatik* in (126c):²¹

- (125) a. Why did John eat what?
 b. *How come John ate what?
- (126) a. ? *Zergatik* jan du Jonek *zer*?
 why eat AUX Jon what
 Why did Jon eat what?
 b. ?? *Zergatik zer* jan du Jonek?
 why what eat AUX Jon
 Why did Jon eat what?
 c. * *Zergatik* Jonek jan du *zer*?
 why Jon eat AUX what
 {Why is it that/How come} Jon ate what?

If high *zergatik* is a complementizer directly inserted in Int°, it follows that it is incompatible with any construction that targets that very same position.

Last, disjoint causal questions have a pair-list reading. Thus, question (127A) may ask about the reasons one may have for choosing either of the options, where (127B) could provide a coherent answer to it. (127A) can also be naturally answered with a single pair that focuses on the whole disjunct (127C):

- (127) A. *Zergatik* hautatuko zenuke bata ala bestea?
 why choose.FUT AUX one.ART XOR other.ART
 Why would you choose the one or the other?
- B. Bata merkeagoa delako, bestea hobeagoa delako.
 one.ART cheaper is.C.because other.ART better is.C.because
 The one because it is cheaper, the other one because it is better.
- C. Biak ezin ditudalako erosi.
 two.ART impossible AUX.C.because buy
 Because I can’t buy both of them.

²¹ As I explained in Section 3.3, ‘antisuperiority-violating’ constructions such as (126a) and (126b) are inherently degraded in matrix constructions.

Under the first reading, *zergatik* interacts with each of the disjuncts and generates a set of questions (a set of sets of propositions). Under the second reading, *zergatik* takes the whole clause as its sister and generates a single question.²²

As can be expected, the pair-list reading is unavailable with high *zergatik* (as represented in (128B)), the only available reading being the single-pair (128C):

- (128) A. *Zergatik* bata ala bestea hautatuko zenuke?
 why one.ART XOR other.ART choose.FUT AUX
 {Why is it that/How come} you would choose the one or the other?
- B. # Bata merkeagoa delako, bestea hobea delako.
 one.ART cheaper is.C.because other.ART better is.C.because
 The one because it is cheaper, the other one because it is better.
- C. Biak ezin ditudalako erosi.
 two.ART impossible AUX.C.because buy
 Because I can't buy both of them.

The high merger of *zergatik* directly in Int^o makes it take scope over the whole clause as such. It is unsurprising then that high *zergatik* is particularly employed in rhetorical questions.²³ Below are some examples of rhetorical questions with high *zergatik*:²⁴

- (129) a. *Zergatik* mundua hain gaizki banatua dago?
 why world so badly share AUX
 {Why is it that/How come} the world is shared so badly?

²²This is similar to what happens in polarity questions with disjunction such as “*Do you want tea or coffee?*” which can be interpreted with a polar interpretation (one single question for which a possible answer could be “Yes.”), or with an alternative interpretation (a pair/series of questions for which a possible answer could be “Tea.”).

²³In this respect, it seems to be different from English *how come*: Fitzpatrick (2005) and Conroy (2006), when analyzing English *why* and *how come* mention examples such as (ia) and (ib), claiming that they show that *why*, as opposed to *how come* can be employed in forming rhetorical questions (for example, (ia) could be part of an exchange where someone asks, ‘Did John leave?’ and the response is ‘No, why would John leave?’, but (ib) cannot be used in such a case):

- (i) a. Why would John leave?
 b. *How come John would leave?

However note that these facts could be analyzed in a different manner: if *how come*, like Basque high *zergatik* is factive and takes the whole clause in its scope –hence asking about the whole eventuality– it could not perform a question (rhetorical or not) inquiring about John’s motives for leaving. Whether *how come* cannot be really employed in rhetorical questions is a matter that deserves further investigation.

²⁴Example (129a) taken from a Fotolog entry, example (129b) from the Wikipedia webpage of Ingmar Bergman’s film *Det sjunde inseglet*, example (129c) from a tweet by user @Iraultza8m (which employs the reinforced dialectal morphological variant *zeatikan* < *zergatik* + *-n* (inessive)), example (129d) from a tweet by user @EuskalHedabide and example (129e) from a tweet by user @beatxo.

- b. **Zergatik** Jainkoak ez du bere izatearen froga edo erantzunik ematen?
 why God neg AUX its nature proof or response give
 {Why is it that/How come} God doesn't provide answers or proof of
 its nature?
- c. **Zeatikan** beti neska bat mutil baten jarrera matxistaz kexatzen
 why always girl one boy one.POSS attitude macho.INSTR complain
 danen atea bar da beste mutil bat esanez "not all man"?
 AUX.C get.out have AUX other boy one saying "not all man"
 {Why is it that/How come} whenever a girl complains about the macho
 attitude of a boy another boy comes saying "not all man [SIC]"?
- d. **Zendako** herri batetan gure hizkuntza, bigarren hizkuntza da?
 why town one.in our language second language BE
 {Why is it that/How come} our language is a second language in a town?
- e. **Zendako** beti eni tokatzen zait aspiragailua pasatzea?
 why always me.to touch AUX hoover pass.to
 {Why is it that/How come} it is always my turn to do the hoovering?

Rather than an innovation of recent years, such different behavior can be observed already in the Classical Basque literature.

5 Conclusions

Why-questions are special in many respects. As I showed, their syntactic pattern is particular in many languages, but their particularity seems to be homogeneous across them: they tend to surface at the left edge of the clause (outscoping topics, foci, etc.) and in languages generally requiring V2 this restriction disappears with *why*-questions. Also, they seem to be able to be first-merged in different positions in the clausal spine, and the very nature of infinite causal links makes it impossible to provide a fully exhaustive answer to a *why*-question. In consequence, any *why*-question can be answered with a series of propositions, each explaining further the information provided by the previous one: Q: *Why did John eat a sandwich?*, A: *Because he was hungry. And there was a sandwich in the plate. And there was no one around. And he knew that he was not going to have anything else until late. And...* This is particular of *why*-questions, as questions on arguments (*Who?*, *What?*...) or other adjuncts (*When?* *Where?*) do not allow such infinity.²⁵

The literature is converging on the idea that elements such as *why* can be first-merged very high in the structure. Here I discussed evidence from Basque in support of this vision, providing evidence that we should distinguish different types of *why*-questions: (i) lower *why*-questions (for reasons and purposes) that show cyclicity ef-

²⁵Maybe *how* also allows series of answers (even if they are more bounded than with *why*). This is another feature linking *why* and *how* together (cf. the discussion on Basque *nola(tan)* or English *how come* above).

fects and residual V2, and (ii) high *why*-questions where the interrogative element is a complementizer directly merged at Int^o and taking the whole clause as a complement. These are elements like English *how come* or Basque high *zergatik*, which are frozen and do not generate V2 effects.

Acknowledgements

Many thanks to M. Duguine, R. Etxepare, C. Mounole and an anonymous reviewer for their helpful comments. Many thanks also to G. Soare for her invitation. This research was funded by the following grants: ANR-18-FRAL-0006 UV2 (ANR-DFG), ANR-17-CE27-0011 BIM (ANR), and PGC2018-096870-B-I00 and FFI2017-87140-C4-1-P (MINECO).

References

- Aldai, Gontzal. 2011. Wh-questions and SOV languages in Hawkins' (2004) theory: Evidence from Basque. *Linguistics* 49(5). 1079–1135. doi:[10.1515/ling.2011.030](https://doi.org/10.1515/ling.2011.030).
- Balukh, Jermy Imanuel. 2020. *A Grammar of Dhao: An Endangered Austronesian Language in Eastern Indonesia*. Amsterdam: LOT Dissertation Series.
- Barreña, Andoni. 1995. *Gramatika jabekuntza-garapena eta haur euskaldunak*. Bilbao: University of the Basque Country UPV/EHU.
- Berk, Stephanie. 2003. Why *why* is different. In Barbara Beachley, Amanda Brown & Frances Conlin (eds.), *BUCLD 27: Proceedings of the 27th annual Boston University Conference on Language Development*, 127–137. Sommerville: Cascadilla Press.
- Bolinger, Dwight. 1978. Asking more than one thing at a time. In Henry Hiz (ed.), *Questions*, 107–150. Dordrecht: Springer. doi:[10.1007/978-94-009-9509-3_4](https://doi.org/10.1007/978-94-009-9509-3_4).
- Bromberger, Sylvain. 1993. *On What We Know We Don't Know: Explanation, Theory, Linguistics and How Questions Shape Them*. Chicago.
- Buell, Leston Chandler. 2011. Zulu *ngani* 'why': Postverbal and yet in CP. *Lingua* 121(5). 805–821. doi:[10.1016/j.lingua.2010.11.004](https://doi.org/10.1016/j.lingua.2010.11.004).
- Cattell, Ray. 1978. On the source of interrogative adverbs. *Language* 54(1). 61–77.
- Chapman, Cassandra & Ivona Kučerová. 2016. Two base-generated positions of why: Evidence from English. Talk delivered at the DogDays syntax workshop at University of Toronto. Toronto, ON, August 18, 2016.
- Collins, Chris. 1991. *Why* and *How come*. *MIT Working Papers in Linguistics* 15. 31–45.
- Conroy, Anastasia. 2006. The semantics of *how come*: A look at how factivity does it all. In Nina Kazanina, Utako Minai, Philip J. Monahan & Heather Lee Taylor (eds.), *University of Maryland Working Papers in Linguistics* 14, 1–24. College Park.
- Conroy, Anastasia & Jeffrey Lidz. 2007. Production/comprehension asymmetry in children's *why* questions. In Alyona Belikova, Luisa Meroni & Mari Umeda (eds.),

- Proceedings of the 2nd Conference on Generative Approaches to Language Acquisition North America (GALANA)*, 73–83. Sommerville: Cascadilla Proceedings Project.
- Cox, Ryan. 2019. How why-interrogatives work. *Synthèse* doi:[10.1007/s11229-019-02364-w](https://doi.org/10.1007/s11229-019-02364-w).
- de Cornulier, Benoît. 1974. "Pourquoi" et l'inversion du sujet non clitique. In Christian Rohrer, Nicolas Ruwet & Benoît de Cornulier (eds.), *Actes du Colloque Franco-Allemand de Grammaire Transformationnelle. Vol. 1. Études de syntaxe*, 139–163. Tübingen: Max Niemeyer.
- Dold, Simon. 2018. *Basque-Spanish Language Contact: An Empirical Study on Word Order in Interrogatives*. Konstanz: Universität Konstanz dissertation.
- Duguine, Maia & Aritz Irurtzun. 2014. From obligatory Wh-movement to optional Wh in situ in Labourdin Basque. *Language* 90.1. e1–e30. doi:[10.1353/lan.2014.0006](https://doi.org/10.1353/lan.2014.0006).
- Endo, Yoshio. 2015. Two ReasonPs: What are*(n't) you coming to the United States for? In Ur Shlonski (ed.), *Beyond Functional Sequence: The Cartography of Syntactic Structures, Volume 10*, 220–231. Oxford: Oxford University Press.
- Etxepare, Ricardo. 2010. From hearsay evidentiality to samesaying relations. *Lingua* 120(3). 604–627. doi:[10.1016/j.lingua.2008.07.009](https://doi.org/10.1016/j.lingua.2008.07.009).
- Etxepare, Ricardo & Jon Ortiz de Urbina. 2003. Focalization. In José Ignacio Hualde & Jon Ortiz de Urbina (eds.), *A Grammar of Basque*, 465–522. Berlin: Mouton de Gruyter.
- Fitzpatrick, Justin. 2005. The Whys and How Comes of presupposition and NPI licensing in questions. In John Alderete, Chung hye Han & Alexei Kochetov (eds.), *Proceedings of the 24th West Coast Conference on Formal Linguistics*, 138–145. Somerville: Cascadilla Proceedings Project.
- Haiman, John. 2011. *Cambodian khmer*. Amsterdam & Philadelphia: John Benjamins.
- Hamann, Cornelia. 2000. The acquisition of constituent questions and the requirements of interpretation. In Marc-Ariel Friedemann & Luigi Rizzi (eds.), *The Acquisition of Syntax: Studies in Comparative Developmental Linguistics*, 170–235. London & New York: Routledge.
- Hamann, Cornelia. 2006. Speculations about early syntax: the production of Wh-questions by normally developing French children and French children with SLI. *Catalan Journal of Linguistics* 5. 143–189.
- Hornstein, Norbert. 1995. *Logical Form: From GB to Minimalism*. Oxford: Wiley-Blackwell.
- Horvath, Julia. 2013. Focus, exhaustivity and the syntax of Wh-interrogatives: The case of Hungarian. In Johan Brandtler, Valéria Molnár & Christer Platzack (eds.), *Approaches to Hungarian: Volume 13: Papers from the 2011 Lund conference*, 97–132. Amsterdam & Philadelphia: John Benjamins.
- Ikeda, Kanako, Tomohiro Fujii & Kyoko Yamakoshi. 2019. Why in the left periphery in Child Japanese: Evidence from children's word order. In Tatiana Bondarenko, Colin Davis, Justin Colley & Dmitry Privoznov (eds.), *Proceedings of the 14th Workshop on Altaic Formal Linguistics (WAFL14)*, 91–102. Cambridge: MIT Working Pa-

- pers in Linguistics 90.
- Irurtzun, Aritz. 2016. Strategies for argument and adjunct focalization in Basque. In Beatriz Fernández & Jon Ortiz de Urbina (eds.), *Microparameters in the Grammar of Basque*, 243–263. Amsterdam and Philadelphia: John Benjamins.
- Jeong, Youngmi. 2003. Deriving Anti-Superiority effects: Multiple *wh*-questions in Japanese and Korean. In Cedric Boeckx & Kleanthes K. Grohmann (eds.), *Multiple Wh-fronting*, 131–140. Amsterdam & Philadelphia: John Benjamins.
- Kahnemuyipour, Arsalan. 2001. On *wh*-questions in Persian. *Canadian Journal of Linguistics/Revue canadienne de linguistique* 46(1–2). 41–61. doi:[10.1017/S000841310001793X](https://doi.org/10.1017/S000841310001793X).
- Kaiser, Georg A., Klaus von Heusinger & Svenja Schmid. 2019. Word order variation in Spanish and Italian interrogatives. The role of the subject in ‘*why*’-interrogatives. In Natascha Pomino (ed.), *Proceedings of the IX Nereus International Workshop “Morphosyntactic and semantic aspects of the DP in Romance and beyond”*, 69–90. Konstanz: Universität Konstanz.
- Kandybowicz, Jason & Harold Torrence. 2012. Krachi *wh*- in situ: a question of prosody. In Jaehoon Choi, E. Alan Hogue, Jeffrey Punske, Deniz Tat, Jessamyn Schertz & Alex Trueman (eds.), *Proceedings of the 29th West Coast Conference on Formal Linguistics*, 362–370. Somerville: Cascadilla.
- Karimi, Simin. 2005. *A Minimalist Approach to Scrambling: Evidence from Persian*. Berlin: Mouton de Gruyter.
- Kayne, Richard S. 1972. Subject Inversion in French Interrogatives. In Jean Casagrande & Bohdan Saciuk (eds.), *Generative Studies in Romance Languages*, 70–126. Rowley: Newbury House.
- Kayne, Richard S. & Jean-Yves Pollock. 1978. Stylistic inversion, successive cyclicity, and Move NP in French. *Linguistic Inquiry* 9(4). 595–621.
- Kirk, Allison. 2012. Word order variation in New Testament Greek *wh*-questions. In Ans M.C. van Kemenade & Nynke de Haa (eds.), *Historical Linguistics 2009: Selected papers from the 19th International Conference on Historical Linguistics, Nijmegen, 10-14 August 2009*, 293–314. Amsterdam & Philadelphia: John Benjamins.
- Kishimoto, Hideki. 2018. Some asymmetries of long distance scope assignment in Sinhala. In Kunio Nishiyama, Hideki Kishimoto & Edith Aldridge (eds.), *Topics in Theoretical Asian Linguistics Studies in Honor of John B. Whitman*, 73–96. Amsterdam & Philadelphia: John Benjamins.
- Kiss, Katalin É. 2002. *The Syntax of Hungarian*. Cambridge: Cambridge University Press.
- Ko, Heejeong. 2005. Syntax of *Why-in-situ*: Merge into [SPEC, CP] in the overt syntax. *Natural Language & Linguistic Theory* 23(4). 867–916. doi:[10.1007/s11049-004-5923-3](https://doi.org/10.1007/s11049-004-5923-3).
- Ko, Heejeong. 2006. On the structural height of reason *wh*-adverbials: Acquisition and consequences. In Lisa Lai-Shen Cheng & Norbert Corver (eds.), *WH-Movement: Moving on*, 319–350. Cambridge: MIT Press.
- Labov, William & Teresa Labov. 1978. Learning the syntax of questions. In *Recent*

- advances in the psychology of language: Formal and experimental approaches*, 1–44. New York & London: Plenum Press. doi:[10.1007/978-1-4684-2532-1_1](https://doi.org/10.1007/978-1-4684-2532-1_1).
- Li, Charles N. & Sandra A. Thompson. 1980. *Mandarin Chinese*. Berkeley: University of California Press.
- McCloskey, James. 2002. Resumption, successive cyclicity, and the locality of operations. In Samuel David Epstein & T. Daniel Seely (eds.), *Derivation and Explanation in the Minimalist Program*, 184–226. Oxford: Blackwell.
- McCloskey, James. 2003. Working on Irish. *GLOT International* 7(3). 63–72.
- Mixelena, Koldo. 1981. Galdegaia eta mintzagaia euskaraz. In *Euskal linguistika eta literatura: Bide berriak*, 57–81. Bilbao: Deustuko Unibertsitatea.
- Miyagawa, Shigeru. 2017. *Agreement Beyond Phi*. Cambridge: MIT Press.
- Ochi, Masao. 2004. *How Come* and other adjunct *wh*-phrases: A cross-linguistic perspective. *Language and Linguistics* 5. 29–57.
- Ortiz de Urbina, Jon. 1989. *Some Parameters in the Grammar of Basque*. Dordrecht: Foris Publications.
- Ortiz de Urbina, Jon. 2002. Focus of correction and remnant movement in Basque. In Joseba A. Lakarra & Xabier Artiagoitia (eds.), *Erramu boneta: A festschrift for Rudolph P.G. de Rijk*, 511–524. Bilbao: Euskaltzaindia.
- Partee, Barbara. 1991. Topic, focus and quantification. In Steven K. Moore & Adam Zachary Wyner (eds.), *Proceedings of Semantics and Linguistic Theory (SALT)*, 1, 159–188. Ithaca: Cornell University.
- Richards, Norvin. 1997. *What Moves Where When in Which Language?* Cambridge: MIT dissertation.
- de Rijk, Rudolph P. G. 1978. Topic fronting, focus pospositioning and the nature of the Verb Phrase in Basque. In Frank Jansen (ed.), *Studies on Fronting*, 81–112. Lisse: Peter de Ridder Press.
- Rizzi, Luigi. 1990. *Relativized Minimality*. Cambridge: MIT Press.
- Rizzi, Luigi. 1996. Residual verb second and the *wh*-criterion. In Adriana Belletti & Luigi Rizzi (eds.), *Parameters and Functional Heads: Essays in Comparative Syntax*, 63–90. Oxford: Oxford University Press.
- Rizzi, Luigi. 2001. On the position “*int(errogative)*” in the left periphery of the clause. In Guglielmo Cinque & Giampaolo Salvi (eds.), *Current Studies in Italian Syntax: Essays Offered to Lorenzo Renzi*, 267–296. Amsterdam: Elsevier BV.
- Shlonsky, Ur & Gabriela Soare. 2011. Where’s ‘*why*’? *Linguistic Inquiry* 42(4). 651–669.
- Stepanov, Arthur & Wei-Tien Dylan Tsai. 2008. Cartography and licensing of *wh*-adjuncts: a cross-linguistic perspective. *Natural Language & Linguistic Theory* 26(3). 589–638. doi:[10.1007/s11049-008-9047-z](https://doi.org/10.1007/s11049-008-9047-z).
- Stromswold, Karin J. 1990. *Learnability and the Acquisition of Auxiliaries*. Cambridge: MIT dissertation.
- Sulaiman, Mais. 2017. Verb second not verb second in Syrian Arabic. In Laura R. Bailey & Michelle Sheehan (eds.), *Order and structure in syntax I: Word order and syntactic structure*, 325–331. Berlin: Language Science Press. doi:[10.5281/zenodo.1117720](https://doi.org/10.5281/zenodo.1117720).

- Takita, Kensuke & Barry Chung-yu Yang. 2014. On multiple wh-questions with ‘Why?’ in Japanese and Chinese. In Mamoru Saito (ed.), *Japanese Syntax in Comparative Perspective*, 206–227. Oxford: Oxford University Press.
- Thornton, Rosalind. 2004. Why continuity. In Alejna Brugos, Linnea Micciulla & Christine E. Smith (eds.), *Proceedings of the 28th Boston University Conference on Language Development (BUCLD 28)*, 620–632. Sommerville: Cascadia Press.
- Thornton, Rosalind. 2008. Why continuity. *Natural Language & Linguistic Theory* 26(1). 107–146. doi:[10.1007/s11049-007-9031-z](https://doi.org/10.1007/s11049-007-9031-z).
- Torrego, Esther. 1984. On inversion in Spanish and some of its effects. *Linguistic Inquiry* 15(1). 103–129.
- Tsai, Wei-Tien Dylan. 2008. Left periphery and *how-why* alternations. *Journal of East Asian Linguistics* 17(2). 83–115. doi:[10.1007/s10831-008-9021-0](https://doi.org/10.1007/s10831-008-9021-0).
- Uriagereka, Juan. 1988. Different strategies for eliminating barriers. In James Blevins & Juli Carter (eds.), *Proceedings of NELS 18, 1987*, 509–522. Amherst: GLSA.
- Uriagereka, Juan. 1999. Minimal restrictions on Basque movements. *Natural Language & Linguistic Theory* 17. 403–444. doi:[10.1023/A:1006146705483](https://doi.org/10.1023/A:1006146705483).
- Yeo, Weichiang Norman. 2010. *Unifying Optional Wh-movement*. York: University of York dissertation.