

On complementizers and embedded gapping in English, Spanish and Polish

JAVIER FERNÁNDEZ-SÁNCHEZ

*Received 15.02.2021,
received in revised form 03.09.2021,
accepted 03.09.2021.*

Abstract

This paper examines two sequences which display gapping under two different embedding configurations in English, Spanish and Polish. I claim that the different distribution of the finite complementizer in these configurations and across these three languages provides further evidence for the idea that gapping is not a uniform phenomenon, and that different structures may correlate with different heights at which coordination can take place in gapping.

Keywords

syntax, ellipsis, gapping, complementizers, coordination

O spójnikach podrzędnych i podrzędnym *gapping* w języku hiszpańskim, angielskim i polskim

Abstrakt

W tym artykule przeanalizowano dwie sekwencje, które wykazują *gapping* w dwóch różnych strukturach podrzędnych w języku angielskim, hiszpańskim i polskim. Twierdzą, że różny rozkład skończone-

go spójnika w tych konfiguracjach, i w tych trzech językach, dostarcza dalszych dowodów na to, że *gapping* nie jest zjawiskiem jednorodnym i że różne struktury mogą korelować z różnymi wysokościami, na których może mieć miejsce koordynacja w *gapping*.

Słowa kluczowe

składnia, elipsa, *gapping*, spójniki podrzędne, koordynacja

1. Introduction

Gapping is a phenomenon in which the verb in the rightmost conjunct of a sentence coordination structure is elided under identity with the verb in the leftmost conjunct (1a), which I will refer to as the *antecedent clause*. Examples (1b) and (1c) show that ellipsis may target elements other than the main verb, like complements or adjuncts, even if these elements do not appear to conform a constituent (1c):

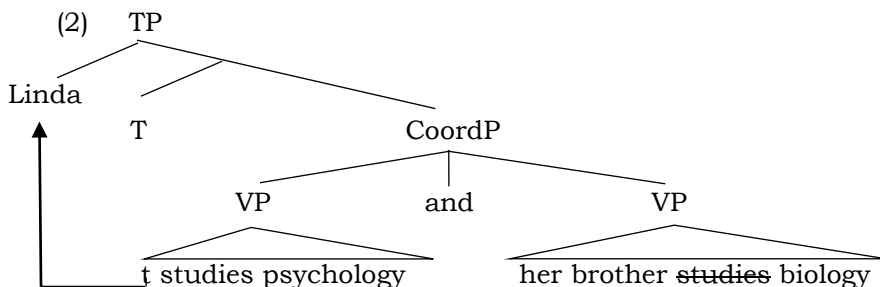
- (1) a. *Linda studies psychology, and her brother ~~studies~~ biology.*
 b. *I will travel to Sri Lanka in the summer, and my neighbour ~~will travel to Sri Lanka~~ in autumn.*
 c. *I will travel to Sri Lanka in the summer, and my neighbour ~~will travel to Israel in the summer~~.*

For the purposes of this paper, the examples in (1) will be referred to as canonical gapping, which can be defined as gapping occurring in matrix clauses.¹ Very broadly speaking, the various existing analysis of canonical gapping differ along two main questions: (i) what formal mechanism is responsible for the gap in the second conjunct?; and (ii) at what height does coordination take place in gapping? With respect to the first

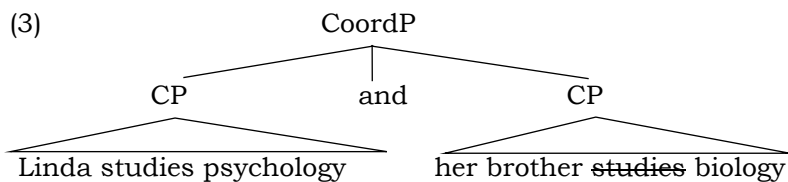
¹ Most of the literature on gapping has indeed focused on canonical gapping. The term is not supposed to have any theoretical relevance, I use it simply to distinguish it from the two gapping structures that I examine in this paper.

question, I will assume that ellipsis involves deletion of syntactic material at PF (i.a. Sag 1976). Following standard practice, I represent elided material in strikethrough text, as illustrated in (1). With respect to the second question, two main analyses have been put forth, which are typically referred to as *low* and *high coordination* accounts.

Low coordination analyses (Coppock 2001, Lin 2002, Johnson 2009, i.a.) posit that coordination in gapping holds at the level of the VP. Under these accounts, the example in (1) would receive the structure in (2). For simplicity reasons, I will represent coordination using non-binary branching, see Zhang (2010) for discussion.²

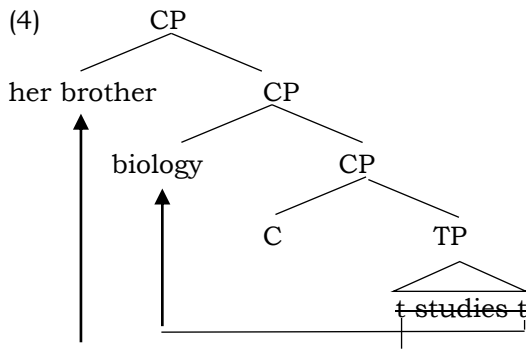


Alternatively, under *high coordination* accounts (Neijt 1979, Hartmann 2000, Reich 2006, *inter alia*), canonical gapping involves coordination of two CPs. Compare (2) to (3):



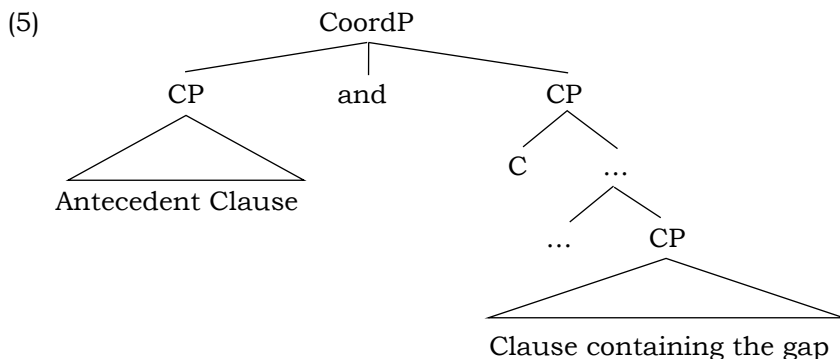
² The representation in (2) is not without its problems. For example, it is unclear why extraction of the preverbal subject from the leftmost VP does not violate the Coordinate Structure Constraint, or how the subject is licensed in the second conjunct; see Johnson (2009) for discussion.

One issue with the representations in (2) and (3) is that the PF deletion operation they display appears to target non-constituents (see the discussion on Fernández-Sánchez 2020: chap.4). To avoid this, it is customary to assume that remnants, i.e. the elements that survive ellipsis – in (1) those would be the DP *her brother* and the NP *biology* – undergo movement to the left edge of the ellipsis domain. Therefore, as an illustration, the rightmost CP in (3) would actually look like (4):³



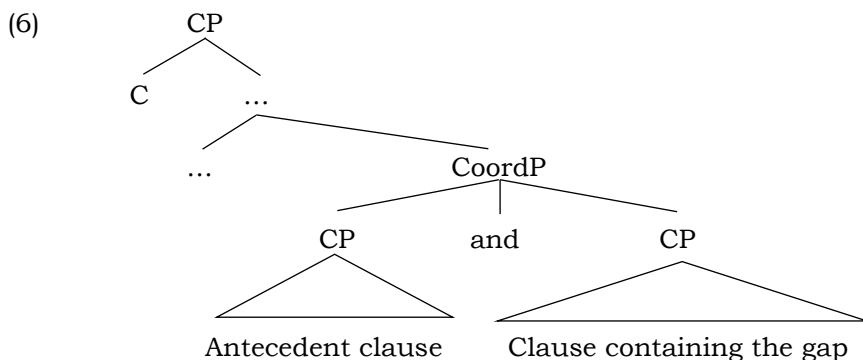
It is interesting to see that canonical gapping is *a priori* compatible with both low and high coordination structures. In this short paper, I focus on two non-canonical gapping configurations in three languages, namely English, Spanish and Polish. These configurations involve gapping in subordination contexts: Non-Canonical Gapping 1 (NCG1) displays an asymmetric coordination structure, where the clause containing the gap is not directly coordinated with its antecedent:

³ For low coordination accounts, remnants would move to the left edge of the CP. For the purposes of this paper the exact position and motivation for such movements are not relevant. We can assume, following the contrastive nature of the remnants in gapping (Kuno 1976), that they move to the specifier of a Focus Phrase.



Such cases have been argued to be ungrammatical in English (Hankamer 1979), but they have been reported to be fine in Spanish and Polish (Fernández-Sánchez 2016), as well as in English (Wurmbrand 2017) and in other languages like Farsi (Farudi 2013), Georgian and Russian (Erschler 2016). I address NCG1 in section 2. Note that the structure in (5), as opposed to cases of canonical gapping, is absolutely incompatible with a low coordination structure, and must be given a high/clausal coordination analysis.

In turn, Non-Canonical Gapping 2 (NCG2) involves cases where the clause containing the gap is directly coordinated with its antecedent clause, just like in canonical gapping (cf. 1); however, in this case, the entire coordination is embedded under one main verb. NCG2 is illustrated in (6):



While (6) is indeed compatible theoretically with both a high and a low coordination account, we will see in section 3 that there are reasons to believe that NCG2 involves a low coordination structure, which means that the representation in (6) will not be entirely accurate. The claims in this paper suggest that gapping is therefore not a unified phenomenon, a conclusion which goes in line with previous research (Repp 2009, Centeno 2011, Jung 2016...).

Before concluding the paper, in section 4 I will tentatively address the syntax of an understudied gapping string which I will take to be a run-of-the-mill case of NCG1 where the embedding predicate is in turn gapped.

2. Non-Canonical Gapping 1

2.1. The No Embedded Constraint

Hankamer (1979) proposed that gapping was subject to the *No Embedded Constraint* (NEC hereafter), which essentially states that neither gaps (7a) nor their antecedents (7b) can be embedded (examples from Hankamer):

- (7) a. * [*Alfonse stole the emeralds*] and [*I think [that Mugsy ~~stole~~ the pearls]*].
 b. * [*I think [that Alfonse stole the emeralds] and [Mugsy ~~stole~~ the pearls]*].⁴

In this paper I have nothing to say about (7b), see Toosarvandani (2016). With respect to (7a), low coordination accounts to gapping, *cf.* (2), can easily explain this restriction: one single T head cannot be shared by two VPs if one of them is embedded in another T head. However, equivalent sentences

⁴ It is important to note that (7b) is grammatical under the reading where coordination holds at the level of the embedded clause. This configuration, illustrated in (i), corresponds to what I call NCG2 in this paper; see section 3 for details:

(i) I think that [*Alfonse stole the emeralds*] and [*Mugsy ~~stole~~ the pearls*].

to (7a) have been argued to exist in other languages like Polish (8) or Spanish (9):

- (8) *Lukasz pojechał do Tajlandii, a zgaduję, że jego brat*
 Łukasz travelled to Thailand and guess that his brother
do Berlina.
 to Berlin
 'Łukasz travelled to Thailand and I guess that his brother
 travelled to Berlin.'
- (9) *Susana compró una casa en el centro de Madrid y*
 Susana bought a house in the centre of Madrid and
diría que Martina un apartamento en la playa.
 would say that Martina an apartment in the beach
 'Susana bought a house in the centre of Madrid, and I'd say
 that Martina bought an apartment by the beach.'

Note that data like (8) or (9) can only be accounted for under a high coordination analysis with clausal ellipsis applying in the embedded clause.⁵ The question is: why would English be different from these languages? Is this a typological split? It is important to mention, however, that English is not that different from Spanish or Polish, despite Hankamer's initial observation: structures like (7a) are possible provided that, as observed by Wurmbrand (2017), no complementizer precedes the remnants:

- (10) Alfonso stole the emeralds and I think Mugsy the pearls.⁶

⁵ Of course this does not mean that gapping in these languages must *always* involve high coordination structures. As an anonymous reviewer mentioned, various authors have developed eclectic accounts of gapping where both high and low coordinations are involved in different gapping strings within the same language; see Repp (2009), Centeno (2011) or Wong (2016); The main claim in this paper is, precisely, that the two configurations under scrutiny here must involve different coordination heights.

⁶ An anonymous reviewer wonders whether this is truly a case of embedding, or whether (10) involves a run-of-the-mill gapping structure where the antecedent clause and the clause containing the gap are directly coordinated and the sequence "I think" is a parenthetical comment clause (Schneider 2007, Griffiths 2013) which provides an epistemic/evidential qualification over a proposition. First, the equivalent structures in Polish and Spanish

In order to capture the data, Wurmbrand proposes the following condition:

(11) *The Embedded Gapping Constraint*

Gapping of embedded clauses is only possible when the embedded clause lacks a CP.

To explain the ungrammaticality of (7a) and the grammaticality of (10), she makes the following assumptions: first, she argues – in line with others (Gallego 2009, Bošković 2014, Aelbrecht 2016) that ellipsis is licensed by phasal heads. Second, she contends that, while there are two phasal domains – thematic and propositional, which roughly correspond to vP/VP and CP respectively – phases should be defined contextually or configurationally. In particular, she defends that phases are the highest head in a phasal domain. Third, she assumes that remnants move to a functional projection (FP) above TP prior to clausal ellipsis, along the lines of (4). Finally, and crucially, she follows Bošković (1997) in claiming that *that*-less embedded clauses are TPs.

After having established the main features of Wurmbrand's analysis, let us see how she derives the facts. Take the example in (10): the verb *think* selects for a TP (following her last premise), as illustrated in (12a). In order for clausal ellipsis to apply, remnants move to a FP above the TP to escape the domain of ellipsis. Ellipsis is then licensed by the highest head in the embedded, propositional phase, which in this case is the head of FP, which triggers ellipsis of its complement, i.e. the TP:

(12) a. *Alfonse stole the emeralds and I think* [_{TP} *Mugsy stole the pearls*].

b. ... *and I think* [_{FP} *Mugsy*_i [_{FP} *the pearls*_{S_k} [_{TP} ~~*t_i stole t_k*~~]]]].

are *bona fide* cases of embedding, as evidenced by the overt complementizer, so one would expect embedding to be possible in English as well. Second, regular fragment answers, which display a very similar syntax to gapping (Reich 2006), can be truly embedded (see Weir 2014), see section 2.

If the complementizer is present, as in (7a), then the verb *think* selects for a CP complement. In this scenario, it is C and not F that is the highest phase in the propositional domain. Consequently, C ought to trigger ellipsis of its complement, which encompasses FP. Under this configuration, remnants would stay trapped within the ellipsis spell-out domain.

Although it is an interesting proposal, Wurmbrand's analysis falls short of empirical coverage as it cannot explain why in languages like Spanish or Polish, the complementizer *must* be present; compare (13) to (8) and (9):

- (13) a. * *Lukasz pojechał do Tajlandii, a zgaduje, jego brat do Berlina.*
 b. * *Susana compró una casa en el centro de Madrid, y diría que Martina, un piso en la playa.*

In what follows I claim that NCG1 should be viewed as cases of (embedded) fragment answers, in the sense of Merchant (2004).

2.2. Embedded fragments

A question like (14) can be answered, at least, in two ways: one involves repetition of the presupposed content (14a), and the other one involves pronouncing only the focus of the sentence (14b). The latter is what is commonly referred to as a fragment answer:

- (14) Who did you see yesterday?
 a. Yesterday I saw Mary.
 b. Mary.

We follow Merchant (2004)'s standard analysis that (14b) is derived from (14a) via clausal ellipsis.⁷ In particular, this au-

⁷ That fragments have an underlying clausal structure can be easily shown in languages with case marking on nominal categories. The equivalent example to (14b) in Polish would be *Marię* (Mary.ACC). The case marking

thor claims that the fragment undergoes movement to a functional projection above the TP prior to ellipsis:

(15) [FP [NP Mary]_i F [TP ~~yesterday I saw t_i~~]].

Importantly for the purposes of this paper, fragments can be embedded, as in (16), from Weir (2014: 221); see fn.6:

(16) A: Who is responsible for the 9/11 attacks?
B: Well, Michael Moore believes Bush.

What I defend here is that NCG1 can be derived by means of the same mechanism that derives (embedded) fragment answers (16). The difference would be that in NCG1 two remnants undergo movement to FP. This analysis is defended on the basis of two parallelisms between embedded fragments and NCG1: (i) the types of predicates under which the remnants can be embedded, section 2.2.1; and (ii) the presence/absence of the complementizer in various languages, section 2.2.2.

2.2.1. Embedding predicates

While fragment answers can be embedded, it has noted that not all predicates can embed them (de Cuba and MacDonald 2013, Weir 2014). This is illustrated in these minimal pairs:

(17) A: Who stole the jewels?
B: I {guess/suppose/think} your son.
(18) A: Who stole the jewels?
B: * I {know/regret/hate} my son.

Let us assume that the key component here is factivity:⁸ factive predicates disallow embedded fragments. One possible

follows naturally from the fact that (14b) contains an elided verb that assigns accusative to the object.

⁸ De Cuba and MacDonald (2013) actually claim that it is not factivity that is at stake, but rather the related – yet independently motivated – notion

explanation is that this is due to the fact that these predicates select for a truncated clausal structure (Vikner 1995, Haegeman 2006) which crucially lacks structural space for remnants to move to prior to ellipsis at PF. The explanation is indeed reminiscent, and correlates nicely, with the classic findings in Hooper and Thompson (1973), who noted that certain syntactic operations like topicalization cannot target the left periphery of clausal complements to factive predicates:

- (19) a. *The inspector explained that each part he had examined carefully.* (Hooper and Thompson 1973: 474, their (50))
 b. * *I resent the fact that each part he had to examine carefully.* (ibid.: 479, their (109)).

If NCG1 involves the same structure as embedded fragment answers, we should expect the same restrictions observed in (17) and (18). The following examples show that this prediction is borne out: (20) illustrates that non-factives (a) are compatible with NCG1 in Spanish, and factives (b) are ungrammatical. (21) showcases the same contrast in Polish:

- (20) a. *Alfonso robó las esmeraldas* { *creo/ imagino/...* }
 Alfonso stole the emeralds and think imagine
que Mugsy las perlas.
 that Mugsy the pearls
 b. * *Alfonso robó las esmeraldas* { *lamento/ odio/...* }
 Alfonso stole the emeralds and regret hate
que Mugsy las perlas
 that Mugsy the pearls
 (21) a. *Alfons ukradł szmaragdów a* { *myślę/ zgduję/...* }, *że*
 Alfons stole emeralds and think suppose that
Mugsy perły.
 Mugsy pearls

of *referentiality*. For the purposes of this paper factivity is enough, as we are interested simply in the descriptive parallelisms between embedded fragments and NCG1; but see de Cuba and Macdonald (2013) for discussion and references.

- b. * *Alfons ukradł szmaragdów a {zaluje/ wiem/ ...}, że*
 Alfons stole emeralds and regret know that
Mugsy perły.
 Mugsy pearls

2.2.2. Presence/absence of the complementizer

Languages differ with respect to whether embedded fragments are preceded by an overt complementizer. English fragment answers cannot be headed by a complementizer (see Fernández-Sánchez and Llinàs-Grau 2017 for discussion), but in Spanish or Polish the complementizer is compulsory:⁹

- (22) A: *What exotic fruit did John buy?*
 B: *I {guess/think/suppose...} (*that) a kiwano.*
- (23) A: *¿Qué fruta exótica compró Juan?*
 B: *{Creo/pienso/supongo...} *(que) un kiwano.*
- (24) A: *Które owoce egzotyczne kupił Janek?*
 B: *{Myślę/przypuszczam/zgaduję...}, *(że) kiwano.*

The distribution of the complementizer in embedded fragments corresponds crosslinguistically with the distribution of the finite complementizer in NCG1 which, taken along with the facts about embedding, strongly suggest that we are indeed dealing with the same phenomenon.

⁹ In fact, crosslinguistically speaking, languages appear to choose one or the other option, i.e. either obligatory presence of C (Spanish, Catalan, Polish, Czech...) or obligatory absence of C (English, Greek, Dutch...). Trying to relate the obligatory absence of C in English to a *that*-trace effect – which is an environment in which English forces an empty complementizer – is not a fruitful line of research (see Weir 2014: 221-233). Furthermore, Greek behaves like English in forcing an empty complementizer but there is no *that*-trace effect in this language (I am indebted to Anna Roussou for discussion on the Greek data).

3. Non-Canonical Gapping 2

The second embedded gapping string I would like to examine involves cases where both the antecedent and clause containing the gap are coordinated at the same level, and coordination appears embedded under a matrix verb. One example is provided in (25):

- (25) I think [Alfonse stole the emeralds] and [Mugsy the pearls].

Contrary to what happens in NCG1, where only a high coordination account is able to explain the data, NCG2 is in principle compatible with both a high and a low coordination analysis (just like any other case of canonical gapping). However, closer scrutiny reveals that a low coordination account fares better with the data.

3.1. Embedding predicates

Suppose that the predicate under which coordination is embedded is a factive one. If NCG2 involved clausal ellipsis like NCG1, then we would expect gapping to be unavailable, given that the coordinated clausal complement would lack the relevant projections for remnants to move to. However, gapping in such cases is possible even with factive predicates, as shown in (26) through (28) for English, Spanish and Polish:

- (26) *I {dislike/regret...} that John goes out with Sonja and Jason
~~goes out~~ with Lilly.*
- (27) *Me desagrada que Pedro me haya servido la sopa
to displeases that Pedro to me has served the soup
fría y su mujer ~~me haya servido~~ el helado derretido.
cold and his wife the ice-cream melted
'It displeases me that Pedro has served me the soup cold
and his wife the ice-cream melted.'*
- (28) *Co za nudne lato tu w Warszawie bez moich
what for boring summer here in Warsaw without my*

przyjaciół. Nienawidzę, że Łukasz pojechał do Tajlandii
 friends hate that Łukasz travelled to Thailand
i Krzys̄ pojechał do Berlina.
 and Krzys̄ to Berlin

‘What a boring summer here in Warsaw without my friends. I hate it that Łukasz has gone to Thailand and Krzys̄ to Berlin.’

3.2. Absence/presence of the complementizer

Hartmann (2001:157) pointed out that in sequences like the one we are dealing with, i.e. NCG2, *that* must be absent in English, an observation she attributes to Fiengo (1974):

- (29) Jim said that Alan went to the ballgame and (*that) Betsy ~~went~~ to the movies.

In NCG1, the lack of an overt complementizer in English was associated with whatever mechanism disallowed complementizers in embedded fragment answers. The lack of the complementizer in sequences like (29), however, cannot be attributed to that same mechanism, for the simple reason that if a unified account was to be pursued, we would expect the complementizer in Spanish and Polish to be mandatorily overt. This prediction, however, is not borne out: NCG2 must involve a null complementizer in these languages as well:

- (30) a. *Przypuszczam, że Łukasz kupił stary samochód, a*
 suppose that Łukasz boughtold car and
 (* że) *Maciek rower.*
 thatMaciek bicycle
 ‘I guess that Łukasz bought an old car and (*that) Maciek a bike.’
- b. *Co za nudne lato tu w Warszawie bez moich przyjaciół!*
*Nienawidzę, że Łukasz pojechał do Tajlandii, a (*że)*
Krzys̄ pojechał do Berlina. (cf.28)
- (31) a. *Supongo que María traerá las bebidas y (* que)*
 suppose thatMaría will bring the drinks and that

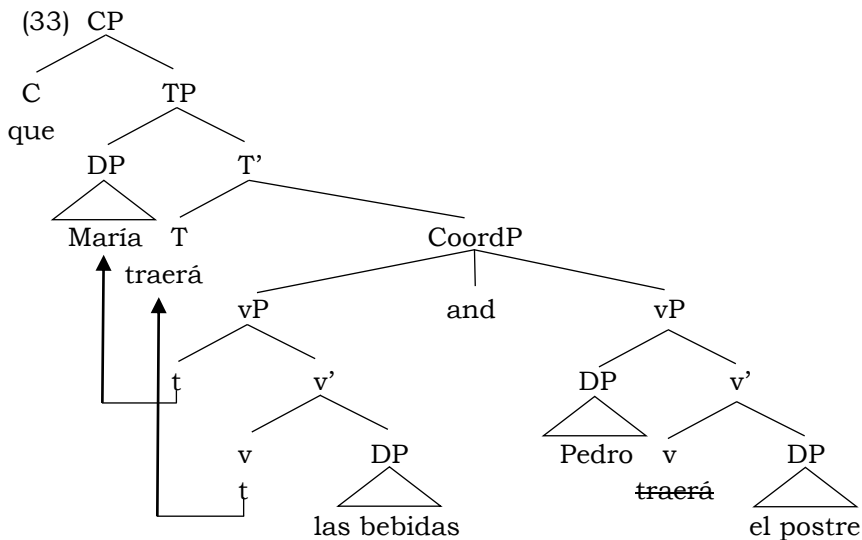
Pedro el postre.
 Pedro the dessert
 'I suppose that Mary will bring the drinks and (* that) Pedro the dessert.'

b. *Me desagrada que Pedro me haya servido la sopa fría y (*que) su mujer me haya servido el helado derretido.* (cf.27)

The fact that in NCG2 is incompatible with the complementizer appears to hold for many languages. Hartmann (2001: 158) observes that the same is true in German:

(32) *Ich glaube, dass Peter mit seiner Frau nach Indien reist und (*dass) Martin mit seinen Kollegen in die Schweiz.*
 I think that Peter withhis wife to India travels
 and that Martin withhis colleagues in the Switzerland
 'I think that Peter travels to India with his wife and Martin ~~travels~~ with his colleagues to Switzerland.'

Taken together, the facts presented in sections 3.1 and 3.2 naturally follow if we assume a low coordination to gapping: take (31a) as an example. According to my proposal, it would involve a structure along the following lines (I only represent the embedded sentence for the sake of simplicity):



In essence, the lack of a complementizer follows obviously from the fact that the second conjunct is not clausal, but rather a vP (but see below). The insensitivity to the factivity of the embedding predicate is expected: under a low coordination account, it is irrelevant whether the left periphery of the embedded predicate is truncated or not. This is so because, again, coordination takes place at a lower level, so no C-domain is involved.

As we saw before, low coordination accounts of gapping assume that coordination holds at the level of the vP. However, as correctly pointed out by a reviewer, the facts presented in this section could still follow from IP-coordination, a solution indeed entertained, but ultimately rejected, by Hartmann (2001) for German. Determining the actual syntactic node at which coordination takes place in NCG2 deserves a more careful examination of the data, a task I leave for further research.

The question that remains is, of course, what is it that bans coordination of two CPs in NCG2. The same reviewer argues that coordination of CPs must be allowed in NCG2 in languages like Spanish at least, because these strings are compatible with gapping involving left dislocated remnants (underlined for expository purposes):

- (34) *Juan aseguró que el dinero lo había guardado en el
 Juan claimed that the money it had saved in the
 banco y las joyas en la caja fuerte.
 bank and the jewels in the strongbox
 ‘Juan claimed that the money, he had saved it in the bank,
 and the jewels in the strongbox.’*

Note that under the assumption that left dislocated phrases are in the left periphery of the clause, the DP *las joyas* (‘the jewels’) must be in a CP-position. Data like (34), however, should be handled with care. To start with, note that the second conjunct is not – and in fact it cannot – be headed by

a complementizer, contrary to what would happen if ellipsis had not applied:

- (35) a. *Juan aseguró que el dinero lo había guardado en el banco y (*que) las joyas en la caja fuerte.*
 b. *Juan aseguró que el dinero lo había guardado en el banco y *(que) las joyas las había guardado en la caja fuerte.*

Testing structure with clitic left dislocation is complicated by the fact that, as shown in Fernández-Sánchez (2017), clitic left dislocated phrases often appear in syntactic contexts where it can be shown independently that there is no structural space, a fact that some authors have taken to mean that dislocated phrases should be viewed as parenthetical elements (Fernández-Sánchez 2017, 2020, Fernández-Sánchez and Ott 2020).¹⁰

But leaving these issues aside, note that there is an important asymmetry between NCG1 and NCG2, in that while it is true that the latter may in theory be compatible with two different structures (whatever they are exactly), the former is not: such cases must necessarily involve a clausal coordination. Given this, we could hypothesize that in cases where two potential derivations would yield the same output, the simplest/most economic one is preferred. Such an economy constraint would be similar to Bošković (1997)'s Minimal Structure Principle:¹¹

(36) *The Minimal Structure Principle*

Provided that lexical requirements of relevant elements are satisfied, if two representations have the same lexical structure and serve the same function, then the representation

¹⁰ One could still adopt a less radical view and claim, along the lines of Jiménez-Fernández and Miyagawa (2014), that clitic left dislocation involves IP-adjunction. This way, (34) would still be compatible with a lower-than-C coordination. Again, I leave this for further research. What is important is to stress that NCG2 cannot involve CP-coordination.

¹¹ A very similar conclusion was reached in Fernández-Sánchez (2020), where I looked at the interplay between coordination and ellipsis in dislocation structures.

that has fewer projections is to be chosen as the syntactic representation serving that function.

Similar claims have been made in the generative literature (see Collins (2001) and Dalrymple *et al.* (2015) for discussion. Unfortunately, exploring this falls outside the scope and goals of this paper, so I leave this issue for further research.

4. Double gaps

4.1. Canonical gapping + NCG1

Before concluding the paper, I would like to bring to the fore a construction which, to my knowledge, was firstly noted in Brucart (1987)'s seminal work on ellipsis in Spanish and which involves two verbal gaps being separated by the finite complementizer *que* ('that'):

- (37) *Pedro aseguró que nevaría en los Alpes, y Juan*
 Pedro claimed that would snow in the Alps and Juan
 ___ *que* ___ *en los Pirineos.*
 that in the Pyrenees
 'Pedro claimed that it would snow in the Alps, and Juan
 claimed that it would snow in the Pyrenees.'
- (38) *Juan confirmó que Susana llegará en avión y Pedro*
 Juan confirmed that Susana will arrive in plane and Pedro
 ___ *que* ___ *en coche.*
 that in car
 'Juan confirmed that Susana will arrive by plane and Pedro
 confirmed that she would arrive by car.'

Polish allows this construction as well, but English does not:

- (39) * *John claimed that Susan would travel by plane and Peter*
 ___ *that* ___ *by car.*
- (40) *Janek powiedział, że Andrzej studiował matematykę a*
 Janek said that Andrzej studied maths and

Wojciech ___ że ___ inżynierię.
 Wojciech that engineering
 'Janek said that Andrzej studied maths, and Wojciech said
 that he studied engineering.'

Brucart (1987) contends that the two gaps are the result of the same operation, i.e. gapping. He attributes the unavailability of this construction in English to the fact that the rightmost gap is actually a complex object formed by the unpronounced verb preceded by a null *pro*. Given that English lacks *pro*, the ungrammaticality of (39) follows. Brucart's explanation would also account for the grammaticality of (40), given that *pro* is available in the grammar of Polish.

The reason to postulate the existence of *pro* comes from Jackendoff (1971)'s suggestion that gaps must contain remnant material at their left and right edges. However, it is well known that remnants of gapping must be focused constituents (Kuno 1976, i.a.) and it is unclear how *pro* can be a focused element. Further, note that under the assumption that the structure under scrutiny is unavailable in English because of the lack of *pro* in this language, we expect this construction to be possible if an overt subject is placed. The prediction, however, is not borne out:

(41) * *John claimed that Susan would arrive by plane and Pedro
 ___ that Laura ___ by car.*

I would like to suggest an alternative account of these facts. Descriptively, these examples featuring a double gap can be explained in the following way: the leftmost gap is an instance of canonical gapping – the matrix verb is deleted under identity with the matrix verb in the antecedent clause. The rightmost gap is embedded under the gapped main verb so, in other words, the rightmost gap is an instance of NCG1. There are reasons to believe this. For example, if we try to use a factive

verb as an embedding predicate, the sentence becomes ungrammatical:

- (42) **Juan lamenta que el gobierno haya subido el IVA*
 Juan regrets thatthe government has raised the VAT
y Pedro lamenta que el gobierno haya subido el
 and Pedro that the
impuesto de sucesiones.
 tax of succession
 ‘Juan regrets that the government has raised VAT and Pedro
 (regrets) that (the government has raised) the estate tax.’

The ungrammaticality of (42) must be attributed to the rightmost gap. We can see this because the two gaps are independent of each other. (43b) is thus ungrammatical for the same reason that (20b) is:

- (43) a. *Juan lamenta que el gobierno haya subido el IVA y Pedro*
~~*lament*~~*a que el gobierno haya subido el impuesto de*
sucesiones.
 b. **Juan lamenta que el gobierno haya subido el IVA y Pedro*
~~*lament*~~*a que el gobierno haya subido el impuesto de*
sucesiones.

The question that remains to be addressed is how is it that English disallows this double gap construction. I discuss this in the next section, where I argue that it is the lack of an overt complementizer heading NCG1 in this language that explains the unavailability of double gaps.

4.2. The clause-mate condition on gapping

To fully understand why English does not allow this construction, it is important that we introduce one locality condition to which gapping is subject: the clause-mate condition on remnants. Empirically, this condition captures the fact that the

gap in (44), here indicated with *e*, can only be interpreted as (45a) and not as (45b):¹²

- (44) *Julia said that Rose speaks Russian and Matthew [e] Polish.*
 (45) a. [*e*] = speaks *Embedded reading*
 b. **[e]* = said that Rose speaks *Matrix reading*

The explanation for the clause-mate condition cannot simply be to assume that the gap is restricted to only one instance of lexical verb. Ross (1970) already noted that the gap can contain more than one verb (46). In light of data like this one, the relevant generalization is that the gap cannot contain a finite clause boundary:

- (46) a. *I want to try to begin to write a novel, and you a play.*
 b. *...and you want to try to begin to write a play.*

The clause-mate condition appears to hold crosslinguistically. (47) shows that gapping in Spanish cannot contain a finite clause boundary, whereas (48) illustrates that it may contain a non-finite clausal node. Examples (49) and (50) illustrate the same point with Polish data:

- (47) *Juan aseguró que Susana llegaría en avión y Pedro*
 Juan claimed that Susana would arrive in plane and Pedro
 [*e*] *en coche*
 in car
 ‘Juan claimed that Susan would arrive by plane and Pedro by car.’
 a. [*e*] = ... y Pedro ~~llegaría~~ en coche.
 b. **[e]* = ... y Pedro ~~aseguró que Susana llegaría~~ en coche.
 (48) a. *Luis prometió casarse en Barcelona y Ana [e] en*
 Luis promised to marry in Barcelona and Ana in in

¹² The labels *matrix* and *embedded* reading capture the height at which coordination must take place in order to derive the corresponding meanings. Therefore, the embedded reading is obtained by coordination at the level of the embedded clause, and the matrix reading by coordination at the root.

Bilbao.

Bilbao

‘Luis promised to get married in Barcelona and Ana in Bilbao.’

b. [e] = (prometió) casarse.

(49) *Janek powiedział, że Andrzej studiował matematykę a*

Janek said that Andrzej studied maths and
Wojciech [e] inżynierię.

Wojciech engineering

‘Janek said that Andrzej studied maths and Wojciech engineering.’

a. [e] = ... a Wojciech studiował inżynierię.

b. *[e] = ... a Wojciech ~~powiedział, że Andrzej studiował~~
inżynierię.

(50) a. *Janek chce studiować matematykę, a Andrzej [e]*

Janek wants to study maths and Andrzej
inżynierię.

engineering

b. [e] = (chce) studiować.

The clause-mate condition poses a challenging theoretical question, given that aside from gapping, it has been argued to hold in many phenomena which involve ellipsis to the exception of more than one remnant like pseudogapping (Jayaseelan 1990), multiple sluicing (Lasnik 2014) or wh-stripping (Ortega-Santos, Yoshida and Nakao 2014), which strongly suggests that there must be a general, across-construction explanation.

Suppose now that we want to derive a double gap structure in English (51). The matrix verb can undergo ellipsis via canonical gapping (51a). This operation leaves the subject DP remnant and the clausal remnant. Now to derive the embedded gap (NCG1), the remnants-to-be need to undergo movement to the left edge of that embedded clause (Merchant 2014), as shown in (51b). Crucially, as we have seen before, embedded fragments are never preceded by the complementizer in English. In the absence of a complementizer (51c), the two remnants must be interpreted as clause mates, *as per* the clause-mate condition on remnants. In other words: the se-

quence *and Rose [e] Hebrew*, can only be interpreted as *and Rose speaks Hebrew*, and not as *and Rose claimed that Susan speaks Hebrew*.

(51) *John claimed that Susan speaks Arabic and Rose claimed that Susan speaks Hebrew.*

a. Matrix coordination, canonical gapping:

John claimed that Susan speaks Arabic, and Rose ~~claimed~~ that Susan speaks Hebrew.

b. TP ellipsis in the clausal remnant in (51b):

[Hebrew]_i; that Susan speaks ~~t_i~~.

c. Resulting string

John claimed that Susan speaks Arabic and Rose Hebrew.

I would like to suggest, thus, that the availability of the double gap construction depends on whether in a particular language embedded fragments (and by extension NCG1) are preceded by an overt complementizer. If they are not, the clause-mate condition on remnants will disallow the intended meaning.

5. Conclusions

In this paper I have looked at two structures which involve non-canonical gapping. NCG1, once (wrongly) thought to be ungrammatical (Hankamer 1979) at least in English, must involve clausal coordination, so it is incompatible with low coordination accounts to gapping. In this configuration, the remnants must be headed by an overt complementizer in Spanish and Polish, but in English this complementizer must be empty. Focusing on the English data, Wurmbrand (2017) proposes an account based on a flexible theory of phases, but her analysis is incompatible with the Spanish and Polish data. I have argued instead that the distribution of the finite complementizer in these three languages can be explained if we posit that the mechanism deriving NCG1 is the same one that yields embedded fragment answers. This allows, in turn, to explain why NCG1 is sensitive to the type of embedding predicate.

With respect to NCG2, I have suggested that coordination must be lower than in NCG1. This conclusion is based on the fact that remnants in this configuration are never introduced by a complementizer, even in languages where the complementizer obligatorily heads remnants, as well as by the insensitivity of NCG2 to the type of embedding predicate.

Consequently, this paper shows, in line with others (Repp 2009, Centeno 2011, Jung 2016...) that gapping is not a unified phenomenon, and that this phenomenon can result from the interplay between ellipsis and coordination at different points in the structure.

Finally, I have briefly addressed the syntax of a construction which features two gaps, which are separated by the finite complementizer, in Spanish and in Polish. I have defended that while the rightmost gap is the result of canonical gapping, the embedded gap is an instance of NCG1. This construction does not exist in English for the simple reason that in this language NCG1 cannot be headed by a complementizer, and therefore the two gaps will end up creating a complex string that is not possible to interpret.

References

- Aelbrecht, Lobke (2016). "What ellipsis can do for phases and what it can't, but not how". *The Linguistic Review* 33: 453-482.
- Bošković, Željko (1997). *The Syntax of Nonfinite Complementation: An Economy Approach*. Cambridge, MA: MIT Press.
- Bošković, Željko (2004). "Now I'm a phase, now I'm not a phase: On the variability of phases with extraction and ellipsis". *Linguistic Inquiry* 45: 27-89.
- Brucart, Josep M. (1987). *La Elisión Sintáctica en Español*. Bellaterra: Publicacions de la UAB.
- Centeno, Naiara (2012). *Gapping and Determiner Sharing in Spanish*. PhD dissertation, UPV/EHU.

- Collins, Chris (2001). "Economy conditions in syntax". In: Mark Baltin, Chris Collins (eds.). *The Handbook of Contemporary Syntactic Theory*. London: Blackwell, 45-61.
- Coppock, Elizabeth (2001). "Gapping: in defense of deletion". In: Mary Andronis, Christopher Ball, Heidi Elston, Sylvain Neuvel (eds.). *Proceedings of Chicago Linguistics Society CLS 37*. Chicago: University of Chicago, CLS, pp.133-148.
- de Cuba, Carlos, Jonathan E. MacDonald (2013). "Referentiality in Spanish CPs". In: Victoria Camacho-Taboada, Ángel J. Jiménez-Fernández, Javier Martín-González, Mariano Reyes-Tejedor (eds.). *Information Structure and Agreement*. Amsterdam: John Benjamins. 117-140.
- Dalrymple, Mary, Ronald M. Kaplan, Tracy Holloway King (2015). "Economy of expression as a principle of syntax". *Journal of Language Modelling* 3: 377-412.
- Erschler, David (2016). "On high and low licensing of gapping". Paper presented at the *Ellipsis Across Borders* conference. University of Sarajevo, Bosnia and Herzegovina.
- Fernández-Sánchez, Javier (2016). "Embedded gaps as embedded fragment answers". Paper presented at the *Ellipsis Across Borders* conference. University of Sarajevo, Bosnia and Herzegovina.
- Fernández-Sánchez, Javier (2017). "Topics at the left edge of infinitive clauses in Spanish and Catalan". *Borealis: an International Journal of Hispanic Linguistics* 5: 111-134.
- Fernández-Sánchez, Javier (2020). *Right Peripheral Fragments*. Amsterdam: John Benjamins.
- Fernández-Sánchez, Javier, Dennis Ott (2020). "Dislocations. Part I and II". *Language and Linguistics Compass* 14: 1-40.
- Farudi, Annahita (2013). *Gapping in Farsi: A Crosslinguistic Investigation*. PhD dissertation, University of Massachusetts at Amherst.
- Gallego, Ángel J. (2009). "Ellipsis by phase". Paper presented at the XIX Colloquium on Generative Grammar, EHU-UPV.
- Haegeman, Liliane (2006). "Conditionals, factives and the left periphery". *Lingua* 116: 1651-1669.
- Hankamer, Jorge (1979). *Deletion in Coordinate Structures*. New York: Garland Pub.
- Hartmann, Katherine (2000). *Right Node Raising and Gapping: Interface Conditions on Prosodic Deletion*. Amsterdam: John Benjamins.
- Hooper, Joan and Sandra Thompson (1973). "On the applicability of root transformations". *Linguistic Inquiry* 4: 465-497.

- Jayaseelan, Karattuparambil (1990). "Incomplete VP deletion and gapping". *Linguistic Analysis* 20: 39-75.
- Johnson, Kyle (2009). "Gapping is not (VP-)ellipsis". *Linguistic Inquiry* 40: 289-328.
- Jung, Wonsuk (2016). *The Non-unity of Gapping*. PhD dissertation. UPV/EHU.
- Kuno, Susumu (1976). "Gapping: A functional analysis". *Linguistic Inquiry* 7: 300-318.
- Lasnik, Howard (2014). "Multiple sluicing in English?". *Syntax* 17: 1-20.
- Lin, Vivian (2002). *Coordination and Sharing at the Interfaces*. PhD dissertation, Massachusetts Institute of Technology.
- Merchant, Jason (2004). "Fragments and ellipsis". *Linguistics and Philosophy* 27: 661-738.
- Neijt, Anneke (1979). *Gapping: A Contribution to Sentence Grammar*. Dordrecht: Foris.
- Ortega-Santos, Iván, Masaya Yoshida, Chizuru Nakao (2014). "On ellipsis structures involving a wh-remnant and a non-wh-remnant simultaneously". *Lingua* 138: 55-85.
- Reich, Ingo (2006). "Toward a uniform analysis of short answers and gapping". In Denis Pascal (ed.) *Proceedings of the 2004 Texas Linguistics Society Conference*. Somerville, MA: Cascadilla Proceedings Project, pp. 69-78.
- Repp, Sophie (2009). *Negation in Gapping*. Oxford: Oxford University Press.
- Sag, Ivan (1976). *Deletion and Logical Form*. PhD dissertation, Massachusetts Institute of Technology.
- Schneider, Stefan (2007). *Reduced Parenthetical Clauses as Mitigators*. Amsterdam: John Benjamins.
- Toosarvandani, Maziar (2016). "Embedding the antecedent in gapping: low coordination and the role of parallelism". *Linguistic Inquiry* 47: 381-390.
- Vikner, Sten (1995). *Verb Movement and Expletive Subjects in Germanic Languages*. Oxford: Oxford University Press.
- Weir, Andrew (2014). *Fragments and Clausal Ellipsis*. PhD dissertation, University of Massachusetts at Amherst.
- Wurmbrand, Susi (2017). "Stripping and topless complements". *Linguistic Inquiry* 48:341-366.
- Zhang, Niina (2009). *The syntax of coordination*. Cambridge: Cambridge University Press.

Acknowledgements

Parts of this paper were originally presented in the Ellipsis Across Borders conference in Sarajevo in 2016, the 37th Tabu-DAG at the University of Groningen, as well as a seminar at the Universitat Autònoma de Barcelona in 2017. I am grateful to the audiences of these venues, as well as two anonymous reviewers who challenged and improved my first draft.

Javier Fernández-Sánchez
ORCID iD: 0000-0002-6060-2332
Uniwersytet Gdański
Instytut Anglistyki i Amerykanistyki
Wita Stwosza 51
80-308 Gdańsk
Poland
javier.fernandez-sanchez@ug.edu.pl