

In press 2007. **The Derivation of Clausal Gerunds.** *Syntax: A Journal of Theoretical, Experimental and Interdisciplinary Research* 10.2..

Title: The Derivation of Clausal Gerunds

Author: Acrisio Pires

Affiliation: University of Michigan

Address:

**Acrisio Pires
University of Michigan**

Email: pires@umich.edu

THE DERIVATION OF CLAUSAL GERUNDS*

Abstract. This paper investigates the syntax of clausal gerunds, a class of gerunds that can have either a null subject or an overt DP Case-marked with accusative or nominative. First, it addresses the difficulty to account for gerunds that allow both null and overt subjects in Principles & Parameters/ Minimalist approaches to Case and to Control. Second, the paper explores the existence of a common structure for the two clausal gerunds, supported by the absence of empirical distinctions in their feature specification, especially regarding tense. Third, the paper introduces new observations about the distribution of clausal gerunds and argues that the complex alternations and restrictions on their distribution results from the interaction between Case and Agreement valuation, the limited possibility of A-movement out of a clausal gerund, and convergence considerations resulting from the existence of distinct numerations.

Keywords: *gerunds, control, syntax, minimalism, subjects, Case.*

1 Introduction

This paper proposes an analysis of the syntax of a class of gerunds,¹ referred to here as clausal gerunds, in which the subject can be either a PRO or an overt DP Case-marked with accusative Case (*acc-ing*) (1).

- (1) a. Susan preferred PRO being late for dinner.
 b. Susan preferred John/him being late for dinner.

These structures yield problems for different approaches to Case and to Control, given that the position in which overt DPs can occur is taken to be ruled out as a position where control null subjects (PRO in standard Principles & Parameters approaches) are allowed. Conversely, the position where control null subjects occur is ruled out by different theories as a possible position for overt subjects. First, it is argued here that the alternation between overt and null subjects exactly in the same context in Clausal gerunds (henceforth CGs), can be successfully analyzed without the adoption of entirely independent structures for the two types of CGs, which is supported by the fact that there are no relevant distinctions in their feature specification that can be taken to be responsible for the distinction between CGs with control null subjects vs. CGs with overt subjects.

Furthermore, this analysis will account for various important facts about the distribution of CGs, which have not been entirely addressed in the literature, including for instance the fact that CGs cannot be licensed in the complement position of passive (2a) and raising matrix verbs (see section 4). CGs can be licensed in these clauses only if the embedded CG head moves to the matrix subject position pied-piping the whole clause (2b):

- (2) a. * John was preferred [reading a book].
 b. [John reading a book] was preferred.

Three main environments in which clausal gerunds can occur will be used for different tests here (see e.g. also Reuland 1983, Milsark 1988, Kaiser 1999, Pires 1999, in press, and references therein): complements to verbs (3a), complements to prepositions (3b, c); in subject position (3d).

- (3) a. Mary favored [Bill taking care of her land].
 b. Susan worried about [Mark being late for dinner].
 c. Sylvia wants to find a new house without [Anna helping her].
 d. [Sue showing up at the game] was a surprise to everybody.

Parallel to the CG cases considered here (1)-(3), gerunds also occur as poss(essive)-*ing* constructions such as (4), where the lexical subject is marked with genitive Case (see Chomsky 1970, Abney 1987 and references therein):

- (4) Mary's winning the contest was a big surprise.

Although I compare CGs and poss-*ing* constructions below, the analysis of poss-*ing* constructions is beyond the scope of this paper, which considers specially the CGs in (1)-(3) as empirical evidence, given the fact that they are the only gerunds that allow an alternation between overt or null subjects.

Section 2 discusses the structure of CGs, analyzed here as projecting a TP/IP. This section discusses the way in which CGs can be distinct from *poss-ing* constructions (*V-ing* constructions whose subject is marked with genitive Case), from regular DPs and from finite clauses. Section 3 analyzes the behavior of null subjects in CGs (standardly analyzed as PRO). It presents new problems for Tense/Null Case approaches to Control and then proposes that the null subject with obligatory Control properties (OC) of CGs can be effectively analyzed as the result of A-movement. As for CGs in subject position, they show non-OC properties and their null subjects do not result from movement. Section 4 lays out the detailed analysis that explains old and new facts about the distribution of different CGs, showing among other cases how their common features can account for the occurrence of null or overt subjects in their derivation.

2 The clause structure of clausal gerunds

CGs behave in most respects like clauses (see e.g. Horn 1975, Williams 1975, Reuland 1983, Kaiser 1999). Several properties distinguish CGs both from regular DPs and from *poss-ing* constructions. *Poss-ing* constructions have been shown to pattern with DPs in different respects (Chomsky 1970, Williams 1975, Abney 1987; Milsark 1988 among others). Abney (1987) argues that gerunds should be re-categorized as DPs at some point in the derivation. For the purposes of the analysis developed in sections 3 and 4, it will suffice to treat CGs as projecting a TP in the derivation.

At least two properties apply to both CGs and *poss-ing* but distinguish them from DPs, making them pattern with clauses: (i) Both *V-ing* constructions, but not DPs, can be modified by (VP)-adverbs (5); (ii) both *V-ing* constructions can directly select for a complement (6a), which can satisfy a Case requirement without need for *of*, which is instead required in an NP-complement (6b-c):

- (5) a. John's/John quickly leaving surprised everybody.
 b. *John's quickly departure surprised everybody.
- (6) a. Mary's/Mary revising the book.
 b. Mary's revision *of* the book.
 c. *Mary's revision the book.²

Despite these two properties common to both CGs and *poss-ing*, CGs are different from *poss-ing* (and from regular DPs) but like regular clauses in various other respects:³ (i) CGs accept certain sentential adverbs which *poss-ing* and DPs do not (7); (ii) CG complements allow (long-distance) *wh*-extraction, whereas *poss-ing* complements do not (8); (iii) The subject position of CGs, but not that of *poss-ing*, can be filled with an expletive, as indicated by the distribution of the pure expletive *there* (9). If CGs are TPs, an appropriate expression needs to occur in [Spec, TP] (in order to satisfy the EPP requirement of T⁰ — see e.g. Chomsky's 1995, 2000 analysis).⁴ (iv) *Poss-ing* are like DPs with respect to Case assignment to their subjects, which are also assigned genitive Case, a possibility that does not hold for CGs (10):⁵

- (7) Mary(*'s) probably being responsible for the accident, the attorney did not want to defend her.
- (8) a. What did everyone imagine Fred(*'s) singing?
 b. Who did you defend Bill(*'s) inviting?
- (9) You may count on *there*(*'s) being a lot of trouble tonight.
- (10) a. Mary worries about him being tired of the trip.

In sum, CGs share several properties with regular clauses that distinguish them from DPs and, in most cases, from *poss-ing* as well. This indicates that they cannot be analyzed as DPs. Given the evidence that CGs behave partially as clauses, they will be analyzed here as projecting up to a TP.

However, there is one general fact about the distribution of CGs that raises a problem for a characterization of CGs as other types of clauses. Different from regular clauses but similarly to DPs and *poss-ing*, the CGs under consideration here have to appear in a Case position in the course of the derivation. This is supported by three kinds of evidence. First, the CGs considered in this paper (3) all appear in case positions, as indicated by (2) repeated below:

- (11) a. Mary favored [Bill taking care of her land].
 b. Susan worried about [Mark being late for dinner].
 c. Sylvia wants to find a new house without [*Anna helping her*].⁶
 d. [*Sue showing up at the game*] was a surprise to everybody.

Second, these CGs are ungrammatical in positions that are standardly treated as Case-less positions, such as passive clause complements (12), a position in which only other clause types are possible (13).⁷ CGs are otherwise allowed in the subject position of passive clauses and in the corresponding active sentence complement (14):

- (12) a. *It was expected [Frank reading this novel].
 b. *Frank was expected [reading this novel].
- (13) a. It was expected [that Frank would read this novel].
 b. John was expected [to read this novel].
- (14) a. [Frank reading the book] was expected.
 b. I prefer [Frank reading the book].

Taking CGs to carry a Case requirement can also account for why CGs can occur as complements to prepositions (15b), unlike finite and infinitive clauses (15a). In this respect, CGs behave like *poss-ing* and DPs (15c):

- (15) a. *Mary talked about [(that) John moved out/ John to move out].
 b. Mary talked about [John moving out].
 c. Mary talked about [John's moving out/ John's move] .

There is one other context in which a distinction between CGs and finite clauses may ultimately be the result of the Case dependency found in CGs. This is the case of *it*-extraposition. Reuland (1983) suggests that expressions that require Case marking cannot undergo *it*-extraposition in English. If this is so, no expression that needs to be assigned Case should be allowed in such contexts. Crucially, this is born out by the fact that besides CGs, neither DPs nor *poss-ing* can be *it*-extraposed, arguably because they are obligatorily Case marked. Along these lines, *it*-extraposition is possible with finite clauses (16a), but impossible with CGs (16b), with *poss-ing* (16c) and with regular DPs (16d):⁸

- (16) a. It was tragic that Paul lost the elections.
 b. *? It was tragic Paul losing the elections.
 c. * It surprised me Mary's leaving town. (Williams 1975 (76)).
 d. * It surprised me Paul's loss.

A similar pattern holds for the contrast between CGs and infinitive clauses, since the latter are allowed in the *it*-extraposition context (17a), presumably also due to the fact they do not need to satisfy a Case requirement, contrary to CGs (17b):

- (17) a. It is impossible [to read this book.]
 b. *It is impossible [reading this book.]⁹

The case dependency of CGs will be accounted for here by assigning to them a Case feature that needs to be checked in the course of the derivation (or valued, under Chomsky's 2000, 2001 approach). However, given the empirical difficulties summarized before for collapsing CGs with regular DPs, this paper will adopt the position that CGs with null and with overt subjects should not be analyzed as projecting a root DP projection.

Different generative accounts of CGs have assumed that they share with finite clauses at least part of their structure. On the one hand, Abney (1987), Milsark (1988) and Kaiser (1999) argue that these constructions project as VPs or IPs and are (re)categorized into an NP or a DP in the syntax, by a process of syntactic affixation, in which *-ing* is affixed or adjoined to the IP/VP projection *converting* it into a DP. Abney and Milsark argue that the *-ing* projection is underspecified for morphological or syntactic features, and that allows the re-categorization to take place. However, the various differences between CGs on the one hand and DPs and poss-

ing on the other hand raise complications for a re-categorization of CGs as DPs. Furthermore, it is unclear how a process of conversion of a syntactic projection from category X to category Y can take place in narrow syntax, if one wants to distinguish derivational processes of the kind found in derivational morphology (as opposed to inflectional morphology) from operations that apply in narrow syntax (see e.g. Chomsky 1970 for seminal discussion).¹⁰ Reuland (1983) and Johnson (1988) each present an account that dispenses with the process of syntactic re-categorization/affixation of CGs into DPs (although these two analyses rely heavily on the properties of Government in order to account for the behavior of CGs, an aspect that does not come into play in the analysis I lay out in this paper). What is common to different analyses is the empirical evidence showing that CGs always project as clauses (TPs here). In addition, given standard assumptions about projection, especially considering the Bare Phrase structure approach to syntactic structure (Chomsky 1995, 2000) one can in fact argue that it is simply not possible to have a DP projecting from the TP. Abney (1987:173) argues that the only property *acc-ing* (the CGs with overt subjects considered here) and noun phrases have in common is their external distribution. I take this distributional property of CGs to be dependent specifically on their Case requirement (see also Reuland 1983), and instead of postulating a DP projection as the root of each CG, I take their Case requirement to be associated directly with a Case feature carried by the top projection of the CG.

One aspect of the analysis developed in this paper is that CGs allow both overt and null-subjects, and both structures display common internal syntactic properties (for instance, they cannot be distinguished on the basis of their tense properties, as will be seen in section 3.1). Notice this paper does not deny the existence of gerunds that may allow only overt subjects or of gerunds that allow only null subjects. This is addressed in section 4.3..

In order to proceed with the derivation of CGs, we need to consider a possible hypothesis regarding the feature specification of their heads, argued here to be at least T(ense).¹¹ Some relevant background is in order here. Here are some properties of an account of certain CGs proposed by Reuland (1983):

- (18)
- i. There are 2 kinds of clauses: tensed or infinitival, with a distinctive marking for each realized on an inflectional head.
 - ii. Tensedness and finiteness are separate parameters.¹²
 - iii. Finiteness is realized by AGR – the agreement marker – which is syntactically present in the Infl of a finite clause and instrumental in the assignment of Case to the subject.
 - iv. AGR is nominal (its feature matrix is simply [N]).
 - v. CGs are taken by Reuland to be tenseless finite clauses, thus [-TENSE, +AGR].
 - vi. *-Ing* in CGs is a nominal agreement marker appearing in Infl that realizes the properties usually associated with AGR.

For Reuland, AGR assigns Case, by transmission. In tensed clauses the presence of a [+tense] feature causes AGR to be assigned Case (given AGR's nominal behavior); this case can then be transmitted to the subject.¹³ Reuland then argues that clausal gerunds are tenseless, and that there is no [+tense] feature in them to assign Case to AGR.

However, there is one crucial problem specifically with the approach to tense in Reuland's analysis. Throughout this paper, the distinction between [+tense] and [-tense] is shown to map directly into a distinction in tense interpretation of the embedded clause. If the embedded clause allows a temporal interpretation that is distinct from the matrix clause, then it is

taken to carry a [+tense] feature specification. In case its tense interpretation can only be identical to the matrix clause, the embedded clause is specified as [–tense]. Given this approach, contrary to what Reuland argues, CGs are in fact [+tense], as clearly shown by their ability to license a temporal interpretation that is independent from the matrix clause (19) (see also Pires 2001b, in press). To avoid this problem, I argue that the mechanism by which clausal gerunds depend on an outside functional head to license an overt subject is not determined by their being [–tense], which they are not.

- (19) a. Mary worried *yesterday* about [Paul coming to dinner *tonight*].
 b. Mary worried *yesterday* about [coming to dinner *tonight*].¹⁴

Given that a [±tense] distinction cannot account for the distinct properties of finite clauses and CGs (they are both [+tense]), it is necessary to account for the special properties and distribution of clausal gerunds in a different way, dispensing with the need to rely on a tense distinction to account for Case checking/valuation in these cases.¹⁵ This is in line with recent minimalist proposals regarding Case checking/valuation of overt DPs (Chomsky 1995, 2000, 2001) which are primarily dependent on the ϕ -feature specification of the Case checking/valuing head. Therefore, it is not possible to adopt the [–tense] approach to CGs proposed by Reuland, given that CGs are in fact [+tense]. The main aspect of Reuland’s analysis that will be directly retained in the analysis implemented here is the idea that the head of a CG (I^0 for Reuland, T^0 here) is instrumental in the assignment of Case to the CG’s subject DP, as it is assumed in most approaches to structural Case assignment to non-ECM subjects (see section 4.1 for evidence that CGs are not ECM complements).¹⁶

One final point is that given the analysis of CGs as projecting a [+tense] Tense Phrase, one might expect them to project a CP as well. However, two important pieces of evidence indicate that there is no CP projection available in CGs. First, they do not allow the occurrence of complementizers, differently from finite and *to*-infinitive clauses (see also Stowell 1982):

- (20) a. Ann wants very much [for Mike to work at home].
 b. Ann wants very much [(**for*) Mike working at home].
- (21) a. Mark prefers [that Mary travel with him].
 b. Mark prefers [(**that*) Mary/PRO traveling with him].

Second, CGs can never occur as indirect questions; that is, partial *wh*-movement, possible with infinitives (22b), is always excluded with CGs (23b, d). If the argument is right that CGs never project an independent CP, there is no intermediate CP position in (23b, d) in which a *wh*-feature can be checked, explaining why partial *wh*-movement is impossible with CG complements.

- (22) a. Bill didn't remember [to buy groceries].
 b. Bill didn't remember [what to buy t].
- (23) a. Sue didn't remember [buying groceries].
 b.* Sue didn't remember what [buying t].
 c. Joan didn't remember [visiting her relatives two years ago].
 d.* Joan didn't remember [when visiting her relatives t].

However, as seen before, long wh-movement applies freely out of CGs, either from subject or object position (24), see also (8):

- (24) a. Linguistics is what we'd favor [him studying *t*]. ((a-b) from Kayne 1981).
 b. The only one who we'd favor [*t* studying linguistics is John].
 c. The winter is when Sue prefers [staying at home *t*].
 d. Where do you defend [Bill moving to *t*]?

If there were a need for an intermediate [Spec, CP] for successive cyclic wh-movement to take place, there would be no way to derive the long-distance wh-movement cases out of CGs in (8) and (24), given the analysis of CGs as bare TPs. Nevertheless, it is possible to argue that the moved wh-phrase in these cases raises directly from its base position inside the CG to the [Spec, CP] of the higher clause. That is consistent with the conception of phases in Chomsky 2000, 2001. If the root TP of CGs were a phase, the wh-element would need to move to the edge of that phase in order to be able to move later to [Spec, CP] of the higher clause. However, TPs are not considered to be phases (see Chomsky 2000, 2001 and references therein), thus a wh-phrase internal to a CG is accessible for movement without needing to move to the edge of the CG.¹⁷ This correctly account for why long-distance wh-movement can apply freely in (8) and (24), whereas partial wh-movement out of CGs is blocked for independent reasons (that is, the impossibility of checking a wh-feature in the absence of an intermediate CP position in the embedded clause) in (23b-d).

3 Null subjects in Clausal Gerunds

In this section I address the licensing of null subjects in the subject position of clausal gerunds. First, I show that null subject CGs split into constructions that only license obligatory control/OC

null subjects and constructions that license non-obligatory control/NOC null subjects. I provide empirical evidence supporting an A-movement analysis of control null subjects in CGs. That analysis is then developed in section 4, as part of a unified analysis that accounts for both new and previously known facts about the complex distribution of CGs with null and overt subjects.

3.1 Questions for a Null Case/Tense Approach to Control in Gerunds

In different minimalist accounts of control (e.g. Chomsky & Lasnik 1993, Martin 1996, 2001, Bošković 1997), PRO is an independent element in the lexicon and in the numeration, and it has also been argued that PRO can only be licensed by being assigned null Case in a [+tense] non-finite domain (Martin 1996, 2001; Bošković 1995, 1997). However, I present below empirical evidence from gerunds against a [+Tense]/Null Case theory of Control. Chomsky and Lasnik (1993) proposed that null Case licenses PRO, stipulating that both are absent in raising and ECM. Martin (1996, 2001) attempted to provide a more principled approach for this distinction, by arguing instead that Control infinitives are [+tense] (Stowell 1982), which licenses null Case PRO (25a), whereas raising (25b) and ECM infinitives (25c) are [–tense], so preclude null Case PRO.

- (25) a. John decided [PRO to leave].
 b. Ann seemed [*t*/*PRO to be interested in the new job].
 c. Bill believed Mary [*t* to be a good friend]./*Bill believed [PRO to be a good friend].

However, contrary to what is argued by Martin, there is no one-to-one correspondence between null Case/PRO and a [+tense] interpretation. First, there is a class of gerunds that is related to CGs but can be distinguished by several properties, including the fact that they only license a Control null subject (PRO). Crucially, different from CGs, these gerunds only allow a

[–tense] interpretation (26a-b) (see Stowell 1982, who first discusses this [–tense] type of gerunds).¹⁸ This is shown by the fact that these gerunds cannot have a tense specification distinct from the matrix clause (26a-b), as shown by the ungrammatical use with these gerunds of time adverbials indicating a distinct time from the matrix clause (the examples are otherwise grammatical without the temporal adverbials), different for example from certain infinitives (27). However, these gerunds still license PRO, which argues against Martin's hypothesis that only [+tense] heads license PRO/null Case.

- (26) a. * Bill tried *today* [talking to his boss *tomorrow*].
 b. * Philip avoided *last night* [driving in the freeway *this morning*].
- (27) Mark decided *today* [to leave to Rome *tomorrow*].

Second, in approaches to Control such as the null Case approach it is required that the Spec, TP position that licenses null Case and PRO be specified so that only PRO can be licensed in this position, and not overt subjects. However, the systematic alternation between overt and null subjects in CGs presents another serious problem for such approaches to Control. More specifically, the CG that licenses an overt subject and its counterpart with a null subject are not distinguishable in terms of properties such as tense, as shown by the alternation in (28), in which both CGs carry a [+tense] interpretation.¹⁹ The [+tense] interpretation in both cases is again shown by the independent temporal modification from the matrix clause. Here there is no direct correlation between the tense specification of the two CGs and the occurrence of either a null or an overt subject in either case. More importantly, aside from the occurrence of an overt DP in (28a) there is no other syntactic property that distinguishes it from (28b).

- (28) a. Sue favored (yesterday) [Anna moving to Chicago (today)].
 b. Sue favored (last week) [moving to Chicago (today)].

Third, in an attempt to provide further empirical motivation for his approach to control, Martin (1996, 2001) assumes that event-denoting predicates contain an event variable that needs to be bound by [+tense] or some other operator - e.g. auxiliary *be/have* (Enç 1990). Given Martin's attempt to argue that control complements are [+tense], his prediction is that they have to allow an individuated event interpretation (as in (25a)), in which the *leaving* event can be identified as needing to occur once at a specific point in time. This is contrary to [-tense] clauses, which for Martin can only allow a stative or habitual interpretation in the absence of tense or some other operator (as he argues for [-tense] raising/ECM predicates such as (25b-c)).

However, both correlations fail regarding gerunds that allow control null subjects. First, the [-tense] control gerunds of (26) are expected by Martin to block an individuated event interpretation, but they allow that interpretation without any problems, as shown in (29). Second, given Martin's prediction that control complements have to allow an individuated-event interpretation, it would be unexpected to find a subset of gerunds that do not allow that interpretation. But contrary to what is predicted by Martin's analysis, control complements such as (30) can only allow a habitual, generic interpretation, when they are embedded under a present tense.²⁰

- (29) Sue tried [leaving at 10a].
 (30) Mark hates/loves [talking to Mary (in the morning)].

These three sets of facts indicate that the distinctive properties of Control and non-Control structures postulated in null Case/Tense approaches to Control face significant empirical difficulties in the domain of gerunds.²¹

The analysis I propose in section 4.1 provides an approach to CGs different from the one reviewed in this section, and dispenses with the need to appeal directly to tense distinctions (absent for the CGs analyzed in this paper, which are [+tense]) to explain why CGs license either an overt or a null subject.

3.2 Obligatory Control in Clausal Gerunds

Concerning null-subject CGs such as (28b), an alternative account to the Null Case theory of Control is to take obligatory control (OC) PRO to be a copy resulting from the movement of the controller DP (an account proposed for infinitives by Hornstein 1999, 2001, and also instantiated in independent approaches by O'Neil 1995 and by Manzini & Roussou 2000). Considering the empirical problems for the application of the null Case theory to gerunds, as I showed above, I instead apply the movement approach to PRO in CGs, arguing that the null subject of certain types of CGs is the residue of A-movement. Crucially, relevant advantages of adopting this approach to null subjects in CGs include: (i) It will provide a way out of the problem resulting from the attempt to link to a tense distinction the alternation between PRO and overt subjects in CGs such as (28), given that the movement analysis postulated in this paper hinges on ϕ -feature and Case-feature specification, and not simply on tense distinctions. (ii) It will pave the way for a related treatment of overt-subject and null-subject CGs (28), especially given that they do not display syntactic distinctions aside from the possible occurrence of the embedded overt DP in (28a).²²

Besides allowing overt subjects, the CGs I consider in detail here also license null subjects, which have been standardly analyzed as PRO. By applying different diagnostics to

CGs, I show in the sets of examples in (31) to (38) that three of the four instances of CGs considered in detail in this paper (3) consistently show properties of obligatory control/OC, when they occur (i) in the complement position of a verb ((b) examples in (31) to (38)), (ii) in the complement of a subcategorized preposition ((31c) to (38c)) ; (iii) in adjuncts introduced by a preposition (cases in (31d) to (38d)). The examples in (31a) to (38a) show similar instances of OC PRO with *to*-infinitives.²³ The various structural and interpretive restrictions applying in (31)-(38) can be effectively accounted for as resulting from the application of A-movement, as it will be formalized as part of the analysis of these CGs in Section 4. For instance, in (31) to (33), PRO requires a local, c-commanding antecedent because these are restrictions imposed on the application of A-movement to the controller.

i. OC PRO must have an antecedent:

- (31) a. *It was expected PRO to shave himself.
 b. *It was never liked PRO staying up late.
 c. *It was aimed at PRO hurting himself.
 d. *It wasn't expected PRO to start the play without PRO turning the lights off.

ii. The antecedent of OC PRO must be local:

- (32) a. * John_j thinks that it was expected PRO_j to shave himself.
 b. *Paul_j thinks that Mary enjoyed/preferred PRO_j shaving himself.
 c. *Peter_j thinks that Mary counted on PRO_j shaving himself.
 d. *Peter_j thought that Mary would leave without PRO_j shaving himself.

iii. The antecedent must c-command PRO:

- (33) a. *John_j's campaign expects PRO_j to shave himself.
 b. *John_j's mother favored PRO_j shaving himself.
 c. *Peter_j's girlfriend worried about PRO_j hurting himself.
 d. *Bill_j's sister left without PRO_j having shaved himself.

iv. OC PRO only permits a sloppy interpretation under ellipsis: In (34), the DP that is interpreted as part of ellipsis gap has to be interpreted as being coreferent with the closest antecedent c-commanding the site of the gap, and this is the same structural relation found between the antecedent and the gap in cases of A-movement, as shown in (35):

- (34) a. John expects PRO to win and Bill does too (= Bill win).
 b. John hates PRO losing and Bill does too. (= Bill lose/*= Bill hates John losing).
 c. John worried about losing and Bill did too. (= Bill lose/*= Bill hates John losing)²⁴
 d. Peter left the party after kissing Mary and Bill did too. (= Bill kissed Mary)

(35) Peter is likely to kiss Mary and Bill is too (=Bill is likely to kiss Mary).

v. OC PRO cannot have split antecedents: If OC PRO results from A-movement, the impossibility of split antecedents can be explained by the fact that two different DPs in the clause cannot have both moved from the same base position:

- (36) a. *John_i told Mary_j PRO_{i+j} to wash themselves/each other.
 b. *Bill_i knew that Mary_j hated PRO_{i+j} hurting themselves/each other.
 c. *Bill_i believed that Mary_j worried about PRO_{i+j} hurting themselves/ each other.
 d. *Peter_i expected Susan_j to break up with him without PRO_{i+j} hurting themselves/ each other.
- vi. OC PRO only has a *de se* interpretation (cf. e.g. Higginbotham 1992: 86-90). For instance, given the specific scenario proposed by Higginbotham, (37b) is false in a situation in which the unfortunate is someone who lost his memory of getting a medal but then reads about the medal award, does not recognize himself as the medal recipient, and is pleased because that person (himself, who he does not recognize!) received the medal.^{25, 26}
- (37) a. The unfortunate expects PRO to get a medal.
 b. The unfortunate liked PRO getting a medal.
 c. Mary worried about PRO getting a medal.
- vii. *Only NP* constructions (Fodor 1975). In (38), the binder of PRO must be the expression formed by *Only + NP*. Take for instance (38c), which cannot be paraphrased with the interpretation in (38c''). It can only be interpreted as in (38c'), under which *only Columbus* could be proud that he himself discovered the West Indies (i.e. the expression 'only Columbus' and not just 'Columbus' is in fact the controller of PRO). This follows in a straightforward way from an A-movement approach if the whole phrase *only Columbus* is base-generated in the embedded clause and moves to the matrix clause. The same interpretive restriction applies to (38a-b, d):

- (38) a. Only Mary remembered to return the book to the library.
 b. Only Churchill remembers giving the BST speech. (Hornstein 1999, (4g)).
 c. Only Columbus_j was proud of PRO_{j, *k} discovering the West Indies.
 Cf. c'. Only Columbus was proud of himself discovering the West Indies.
 c''. #Only Columbus_j was proud that he_{j, k} discovered the West Indies.
 d. Only Churchill_j was congratulated after PRO_{j, *k} giving the BST speech.
 Cf. d'. Only Churchill was congratulated after himself giving the BST speech.
 d''. # Only Churchill_j was congratulated after he_{j, k} gave the BST speech.

These different tests show that CGs display a number of core properties that are better subsumed under an A-movement approach. In the analysis developed in section 4 I show how the control null subject of these CGs can be analyzed as the residue of movement of an overt DP to a higher domain in order to check its Case feature. Crucially, I will show there how this analysis can be compatible with the alternative licensing of overt DPs in the subject position of CGs, a possibility that cannot be easily accounted for in approaches to Control such as a Null Case/Tense approach, as I pointed out in section 3.1.

3.3 Non-Obligatory Control in Clausal Gerunds

CGs in subject position are the only instances of CGs where non-obligatory Control/NOC PRO properties systematically hold: NOC PRO does not require an antecedent ((a) below); if it has an antecedent, the antecedent does not need to be local (b); the antecedent does not need to c-command NOC PRO (c); a strict reading of the elided VP is possible in (d); NOC PRO allows split antecedents (e); a non-*de se* interpretation is available for NOC PRO with respect to the matrix subject (f); and finally, the binder of NOC PRO does not need to be interpreted as the

Only+NP expression available in the matrix sentence in (g), since George Bush didn't win the Second World War:

- (39) a. It was believed that graduating from college was important.
 b. John thinks that it is believed that PRO shaving himself is important.
 c. Clinton's campaign believes that PRO keeping his sex life under control is necessary for electoral success.
 d. John thinks that PRO getting his resume in order is crucial and Bill does too.
 e. John_j told Mary_k that PRO_{j+k} washing themselves/each other would be fun.
 f. The unfortunate believes that PRO getting a medal would be boring.
 g. Only George Bush remembers that PRO winning the Second World War was crucial.

Hornstein (1999) argues that instances of NOC PRO are distinct from OC PRO in that they cannot be analyzed as the residue of A-movement. He assumes that they are instances of *pro*, in order to eliminate PRO as an element of the grammar. Standard instances of PRO are then reanalyzed as either copies left behind by A-movement (OC PRO) or instances of *pro* (NOC PRO). If that is the case, what distinguishes pro-drop from non-pro-drop languages is not the lack of *pro* in the latter, but rather the fact that although *pro* is present across languages in general, its occurrence is much more widespread in standard pro-drop languages. I put aside here further consideration of gerunds with NOC subjects, and in the next section focus especially on CGs with control null subjects and with overt subjects.²⁷

4 The syntax of Clausal Gerunds

In this section I propose an analysis of the syntax of clausal gerunds that attempts to account for five core syntactic properties of CGs, regarding especially their distribution and

licensing of subjects. The current analysis explores an approach to Case checking/valuation that is related to agreement (ϕ -feature) checking/valuation in the Minimalist program (Chomsky 2000, 2001 and references therein). The current analysis attempts to explain, in a unified approach, the following empirical facts about the behavior of CGs:

i. The subject of a CG may be an empty category (standardly analyzed as a PRO) or an overt DP:

(40) a. The manager preferred [PRO being considered for the position in the downtown office].

b. The manager preferred [Mary being considered for the position in the downtown office].

ii. CGs need to satisfy a Case requirement (as argued in detail in section 2) :

(41) a.* It is expected [John reading the book]. (cf. (41d-e))

b.* John is preferred reading the book.

c.* John is impossible reading the book.

d. [John reading the book] was preferred.

e. I prefer [John reading the book].

iii. CGs do not behave as Exceptional Case Marking/ECM complements, which is indicated by the fact that they do not occur as complements of standard ECM verbs like *believe* (42b-c) nor allow raising of their subjects to the subject position of any passive structure (in ECM (42c) and non-ECM cases (43b)):

- (42) a. Paul is believed to be smart.
 b. *Mary believes [John being smart].
 c.. *John is believed [being smart.]
- (43) a. Mary prefers [Paul swimming in the morning.]
 b. *Paul is preferred [swimming in the morning.]

iv. CGs can never occur as complements of subject raising verbs (44a-b), although they can occur as a single constituent in the subject position of raising predicates (44c):

- (44) a. *There seems [being a man in the room].
 b. *John appears [liking Mary].
 c. *It appears [John liking Mary].
 d. [(John) talking to Mary] seems impossible.

v. The subject position of a CG must be filled out in the course of derivation, either by a lexical DP that may further move (43a), as I propose here, or by a pure expletive (45).²⁸ In standard terms, this means that the subject of a CG needs to check an EPP feature in [Spec, TP] of the CG:

- (45) a. Bill enjoys [there being many people in the party.]

4.1 The derivation of Clausal Gerunds

In order to explain the properties of CG above, my analysis explores the hypothesis that three properties apply to the CGs that allow alternation between overt and null subjects:

- (46) a. the inflectional head corresponding to *-ing* in CGs carries a feature specification that forces the occurrence of CGs in positions accessible to Case valuation (section 2);
- b. in the derivation of a CG, the Case feature of its external argument DP can be valued within the CG itself, in examples such as (47a), or
- c. the external argument DP can move out of the CG before the CG can value the Case feature of this DP. This yields a null-subject CG (a CG with a control PRO subject, in standard terms), as in (47b).

- (47) a. Sue prefers [John/him swimming.]
- b. John prefers [swimming.]

One possibility is that the hypotheses (46b) and (46c) should be treated as resulting from two completely different structures, and one would have to find empirical motivation showing that the syntax of the CGs in (47a) and (47b) is clearly distinct. However, there are no syntactic feature distinctions between the CGs with an overt vs. a null subject (as in (47)); see for instance arguments in section 2 for lack of tense distinctions between these two CGs). The lack of feature distinctions between both types of CGs raises problems for different P&P approaches to overt and null subjects, for two major reasons. First, the syntactic properties that license an overt subject (e.g. Case Theory in GB) block the occurrence of PRO. Conversely, the syntactic conditions that determine where PRO is licensed block the occurrence of an overt subject. So, most P&P approaches induce complementary distribution of PRO vs. overt subjects, but CGs indicate that complementarity is factually incorrect. The analysis below attempts to reconcile the treatment of these two problems regarding CGs, together with the treatment of the other complex facts about the distribution of CGs presented in 4 above.

In order to account for (46a) it is proposed here that not only the external argument DP of a CG, but also its root node (a Tense head in the analysis below), carry an uninterpretable Case feature that needs to be valued. Under this approach, the head T of the CG itself will be a goal for Case valuation. As I explain in detail below, if the DP subject can move out of the CG before the Case feature of the T head of the CG is valued (and the latter must happen by the end of the derivation), a standard OC PRO construction is licensed, yielding (47b). The Case requirement on CGs is formalized in (48i). Furthermore, I propose that a condition which blocks feature valuation applies during Agree (48ii), preventing valuation from happening for as long as both probe and goal carry an uninterpretable instance of the same feature:

- (48) i. The Tense (T^0) head of a CG carries an uninterpretable Case feature that needs to be valued.
- ii. A probe cannot value an uninterpretable/unvalued feature of its goal while the probe itself has an uninterpretable/unvalued feature of the same kind.

Property (48i) is a direct formalization of (46a), discussed in detail in section 2. Property (48ii) is instantiated in this analysis by the fact that the T of a CG can value the Case feature of a DP only after its own (unvalued) Case feature has been valued by an appropriate probe (e.g. a matrix v or T). In what follows it is shown how the complex distribution of CGs (as presented in section 4 above) follows from the interaction between these two properties and the general architecture of feature checking/valuation in recent approaches to the Minimalist Program (e.g. Chomsky 2000, 2001).²⁹ Given this architecture, the head of a CG -- its Tense head -- is taken to display three other properties that are equivalent to the ones that apply to feature checking/valuation in *to*-infinitives:

- (49) i It has an EPP feature that needs to be checked.³⁰
- ii. It enters the numeration as ϕ -defective.
- iii. When the Tense (T) head of the CG (a probe) Matches/Agrees with the embedded subject DP (a goal), the DP merges in [Spec, TP] of the CG to check EPP and values the ϕ -features of T.

In the approach to feature valuation adopted here (which follows in most respects Chomsky 2001), the functional heads ν and T in finite and non-finite clauses carry uninterpretable/unvalued ϕ -features. Both ν and T can probe the derivation for matching active goals that can check the uninterpretable/unvalued ϕ -features of ν and T. A DP has interpretable ϕ -features and enters the derivation with an uninterpretable/unvalued Case feature that makes it active to induce Agree with a probe. The case feature of the DP is valued after that DP enters into Match/Agree with a functional head that is ϕ -complete. Both the head (probe) and the matching DP (goal) have to be active in order to enter into a Match/Agree relation. Only uninterpretable/unvalued features (e.g. uninterpretable, unvalued ϕ -features and Case) activate a probe and a goal, thus inducing Match/Agree. Once the functional head T^0 (the probe) of a finite clause and the subject DP (the goal) enter into Match/Agree, the uninterpretable Case feature of the goal DP and the uninterpretable ϕ -features of T are valued.³¹ Overt movement to [Spec, TP] is necessary to satisfy an EPP-requirement. In the case of clausal gerunds, T^0 is ϕ -defective, but it can still attract its subject DP in order to satisfy its EPP requirement, as in instances of T_{def} (ϕ -defective T) in infinitives.³²

Consider how this architecture can account for the core properties of clausal gerunds in the derivations that follow. *I represent the head (T, as referred above) of the embedded CG as AGR, to make its status clear in the different derivations.*

First, consider in detail the core case of a null-subject CG as in (50). It is proposed here that the null subject in such cases results from A-movement of the embedded CG subject to the matrix clause, as supported by the different A-movement diagnostics presented in detail in section 3.2. Crucially, since I analyze instances of OC PRO in CGs (47b) as the result of A-movement, I assume that θ -roles can be assigned through movement and not only by first merge (cf. Bošković 1994; Lasnik 1995; Bošković and Takahashi 1998; Hornstein 1999, 2001). The idea is that θ -roles can also be assigned in the course of the derivation, and are satisfied not in a configuration, but in a set of configurations (i.e. transformationally).

In (50) The head AGR of the CG starts ϕ -defective and with an uninterpretable Case feature C_{AGR} . Remember that the uninterpretable Case feature on AGR corresponds to the Case requirement of the different CGs analyzed here, as stated in (48i), and is responsible for their restriction to Case positions (section 2). As *John* enters into Match/Agree with AGR in (50b), *John* moves to Spec TP1 for EPP checking and values the ϕ -set of AGR by Agree.³³ Crucially, valuation of the ϕ -set of AGR by the DP eliminates its defectiveness.³⁴ However, since AGR still has an uninterpretable Case feature at the point in (50b), Case valuation of the embedded subject DP cannot yet take place. This is the instantiation of (48ii), by which a probe that carries an uninterpretable Case feature cannot value the Case of its goal until its own Case feature has been valued:

(50) John prefers swimming. =(47b)

a. [T'AGR [vP John swimming]]...

C_{AGR} θ/C

b. [TP1 John [T'AGR [vP John swimming]]]...

EPP/ϕ C_{AGR} θ/C

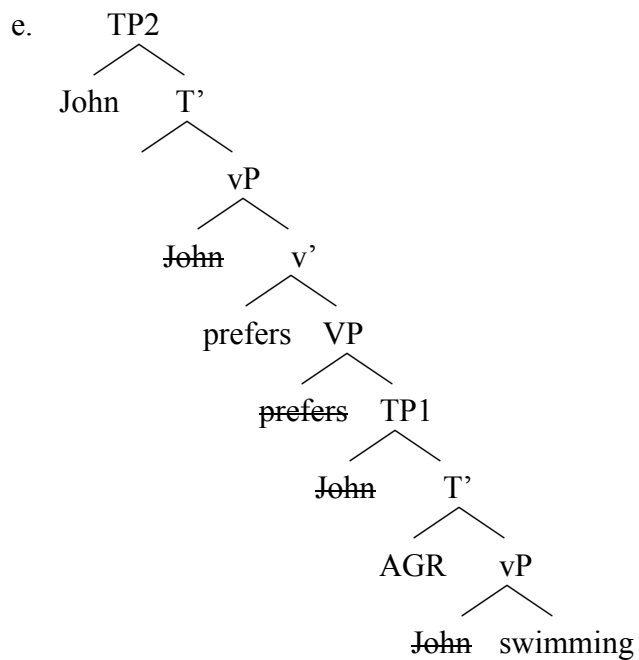
c. [vP John [v' prefers [VP prefers [TP1 John [T'AGR [vP John swimming]]]]]...

$2\theta/C$ C_{AGR} EPP/ϕ θ

d. [TP2 John [T' [vP John [v' prefers [VP prefers [TP1 John [T'AGR [vP John

$\phi/C/EPP$ 2θ C_{AGR} EPP/ϕ θ

swimming]]]...



As the matrix V is inserted in the derivation, the embedded CG is assigned the propositional internal θ -role of the matrix verb (50c). When matrix v enters the derivation, it attracts the embedded DP *John* and assigns an experiencer θ -role to it. The matrix v then Matches/Agrees in ϕ -features with the embedded AGR in CG and values the uninterpretable Case feature (C_{AGR}) that AGR still carries (50c). Finally, *John* moves from matrix [Spec, vP] to [Spec, TP2] to check its own uninterpretable Case feature and the EPP feature on T2 (50d). The complete derivation is represented in (50e).³⁵

The analysis outlined above gives us the means to explain the grammaticality and ungrammaticality of a host of other cases involving the distribution of CGs. Before we proceed to them, there are a couple of important considerations I would like to discuss regarding the steps in the derivation above. First, when the matrix v is inserted in the derivation (50c), it carries an external θ -role³⁶ and the uninterpretable ϕ -features which allow it to enter into the Match/Agree operation that will value the case of the embedded CG. The sequence in which θ -role assignment and Case valuation take place is entirely restricted by the fact that this is the only possible convergent derivation of (50) (this will be shown in further detail in the discussion of how an ungrammatical derivation of the same example crashes in (50') below). Crucially, the ordering of steps in the derivation (50) is also fully compatible with cyclicity (as defined for instance in Chomsky 1995:233),³⁷ which is satisfied at all points. More specifically, at the point matrix v is inserted in the derivation (50c), it assigns a thematic role to the embedded DP, which moves to the matrix vP external argument position. Before the derivation leaves the matrix- vP cycle, the matrix v values its uninterpretable ϕ -features and the Case feature of the embedded CG. Furthermore, the embedded DP and the embedded AGR are equidistant from the probing matrix v , if one adopts the idea that they are in the same minimal domain in the embedded clause (the domain of the embedded T itself; see Chomsky 1995), being both accessible to the operations

that take place at the point matrix ν is inserted in (50c).³⁸ Also, AGR, the embedded DP and the matrix ν are all available within the same strong phase represented by the matrix ν P.

Second, the existence of multiple features that need to be checked in the derivation of (50) raises questions as to whether the derivation could proceed otherwise. However, Minimalist approaches determine that different restrictions block the possibility of alternative derivations, and such restrictions also come into play in the derivation of different CGs, including (50). One important restriction that blocks alternative derivations is the need for convergence at the interfaces. Consider one alternative derivation of (50) (also compatible with cyclicity) that is in fact blocked because it would yield a syntactic object that cannot be interpreted at the interface. If the matrix ν (*prefer* + ν) in (50c) valued the Case of AGR before *John* moved out of [Spec, TP1], this would also allow AGR to value the Case feature of *John* in [Spec, TP1] (given that AGR would now no longer have an uninterpretable Case feature (see (48ii)), as shown in (50c') below. After having its Case valued *John* would effectively be prevented from raising to the matrix clause. However, the matrix subject θ -role and the EPP feature in [Spec, TP2] would end up not being satisfied (50d'), because *John* would now be inactive for further Match/Agree relations, and would no longer be able to move out of the embedded clause. Given the absence of other DPs in the numeration of (50) that could satisfy the EPP requirement in [Spec, TP2], this yields a crash in the derivation ending in (50d') below, further motivating (50) as the only possible derivation of the null subject CG case in (47b).

(50') *prefers [John swimming]

c'. [vP [v' prefers [VP prefers [TP1 John [T' AGR [vP John swimming]]]]]...

€_{AGR} €/EPP/φ θ

d'. [TP2 [T' [vP [v' prefers [VP ~~prefers~~ [TP1 John [T' AGR [vP ~~John~~

φ[?]/EPP[?] θ[?] €_{AGR} EPP/€/φ θ

swimming]]]]]...

In sum, despite the impression that the derivation up to the insertion of the matrix *v* (50c) could allow an alternative path starting in (50c'), this path in fact leads to a non-convergent derivation. The question of optionality of derivations will come up again in connection with the derivation of the overt subject case in (47a), as analyzed in (51). I will show that there is no free optionality applying there either (consistent with the minimalist hypothesis that optionality is restricted in the grammar), because apparently alternative derivations are in fact not alternative, due to their distinct numerations (as shown by the contrast between (50) and (51)), or due to the fact that only one of them is convergent (as illustrated with (50) vs. (50')).³⁹

Consider now the core case involving an overt subject within the CG (51). The derivation of (51) proceeds in the same way as (50), but this similarity stops after the step in (51b). When the matrix *vP* is generated in (51c), the derivation proceeds differently from (50), clearly because at this point (step c), *Sue* is still available in the numeration, whereas the numeration of (50) did not have this additional DP. Crucially, the existence of two different numerations rules out any consideration of (50) and (51) as competing derivations, given that they do not share the same numeration.⁴⁰ Besides this difference, there is in fact only one path that can lead to convergence in the derivation of (51), as I will show in the next steps. Matrix *v* matches/agrees with AGR (the Tense head of the CG), valuing the Case feature of AGR in (51c).

At this point, the Case feature of the embedded DP *John* can also be checked, and this is done by AGR of the CG, with which the DP *John* has already established an Agree relation. Case valuation of the DP *John* is possible only at this point because this is when the restriction in (48ii) is finally eliminated (that is, the Case of the CG itself has been valued).

(51) Sue prefers [John swimming]. =(47a).

a. [T'AGR [ν P John swimming]]...

C_{AGR} θ/C

b. [*TP1 John* [T'AGR [ν P John swimming]]]...

~~EPP~~/ ϕ C_{AGR} θ/C

c. [ν' prefers [*VP prefers* [*TP1 John* [T'AGR [ν P John swimming]]]]]...

C_{AGR} ~~EPP~~/ C θ

d. [*TP2 Sue* [T' [ν P ~~Sue~~ [ν' prefers [*VP prefers* [*TP1 John* [T' AGR [ν P ~~John~~

$\theta/C/\del{EPP}$ θ C_{AGR} ~~EPP~~/ C θ

[ν' swimming]]]...

Notice furthermore that there is no violation of cyclicity regarding Case valuation: it is only when matrix ν is merged that both the Case feature of the CG and the Case feature of the embedded subject *John* can be valued, and both instances of valuation take place in this order at this point in (51c), as determined by the features at play here⁴¹ (compare this later with (52), in which a different path for Case valuation here would yield a non-convergent derivation).

Finally, in (51d) the matrix subject *Sue* merges in [Spec, vP] where it checks the matrix external θ -role, and raises to [Spec, TP2] in order to check the EPP feature and its own Case feature.

Again, the need for the derivation to converge imposes a restriction on the possibility of movement of the embedded subject, as I show through this section. What determines how the Case feature of the embedded subject DP is going to be valued are the steps that can yield a convergent derivation, as shown in the contrast between (50) and (51). In different cases, although it might appear that optionality is at play, the sequence of steps in each derivation is in fact determined by the need for convergence at the interfaces. In each case, given a particular numeration, only one choice results in convergence at the end of the derivation. In examples such as (50) and (51) this is further supported by the fact that there are in fact different numerations under consideration. In one case, the embedded subject moves out in case the host CG has not yet checked its own features (yielding a CG with a null subject and OC properties, as in (50)). In the other case, after the CG has already checked its own Case feature, the embedded subject (in case it has not moved out) can only check its features internally to the CG, and freezes in its embedded position, yielding a CG with a lexical subject, as in (51). But for the latter to yield a convergent derivation, an additional argument DP needs to be available in the numeration, as it happens with *Sue* in (51). In sum, there is no issue of optionality involving a comparison of the derivations of (50) and (51) because the two cases correspond to different numerations.

Furthermore, *within* the individual derivations of (50) and (51) there is no issue of optionality of the sort that comes into play for instance in the analysis of empirical phenomena requiring appeal to economy of derivation, by which a more economical derivation is chosen over another, also convergent one.⁴² (50) and (51) are different in this respect because each one is the only path that will ultimately converge in each case, and there is no other alternative path

with the same numeration in each case that would also lead to a convergent derivation. The latter has clearly been shown in the comparison between (50) and (50'). The fact that, despite appearances to the contrary, there is no global or local optionality involved in the path taken by each individual derivation (i.e. each derivation proceeds exactly the way it does due to the need for convergence), supports a specific approach to minimalism, by which no optionality should be freely allowed in narrow syntax.

The lack of free optionality can be further explored regarding why (51) does not allow the alternative derivation shown in (52), within which the embedded *John* subject would value its Case by ECM, a possibility that does not hold in general for CGs, as I showed in section 4, property (iii). Here is an analysis showing why ECM is unavailable with CGs, ruling out a derivation of (47a) as ECM (52):

(52) Sue prefers John-*ECM swimming.

[_{TP} Sue [_{T'} [_{VP} Sue [_{VP} John [_{v'} [_{VP} prefer [_{TP} John [_{T'} AGR [_{VP} John [_{v'} swimming]]]].

€/EPP θ € EPP *C_{AGR} θ

At the point at which matrix *v* enters the derivation, one possibility would be for the subject *John* of the CG to enter into Match/Agree directly with matrix *v* to value its own Case (different from (51) and as in ECM constructions), but this derivation is ruled out straightforwardly because the Case feature on AGR would end up unvalued, yielding crash at the interfaces.^{43, 44}

4.1.1 CG pied-piping from the complement position of passive and raising verbs

Consider now a derivation in which the whole CG moves to the subject position of the matrix clause, when it is generated as the complement of a verb that cannot value (accusative)

Case of an internal argument. As predicted by the analysis of CGs developed here, this is possible only in certain cases in which there is *not* an accusative Case available in the matrix clause. The first case is when a CG is generated as the complement of a passive verb, as shown in the derivation of (53). According to the most widely adopted analysis of passives, the passive morphology eliminates the matrix external θ -role and the accusative Case position in the matrix clause [Spec, v P].

(53) Bill swimming was preferred.

[TP2 [TP1 Bill swimming] [T' [v P [v' θ [v' was preferred [TP1 ~~Bill~~ [T' AGR

\in ~~EPP~~/ \in _{AGR}

EPP

[v P ~~Bill~~ [v' swimming]]]...

θ

T2 (matrix T) matches/agrees with AGR, and AGR raises, pied-piping the whole clause (TP1) to [Spec, TP2], in order to check the EPP feature in the [Spec, TP2] of the matrix clause.⁴⁵ In the [Spec, TP2] AGR (the T head of the CG) has its own Case valued and is now able to value the Case of its subject *Bill*, with which it has already established a Match/Agree relation in the course of the derivation. Other cases of CGs in subject position can be handled along the same lines, including CGs as complements of raising verbs:

(54) [Bill swimming] seems/is [~~Bill swimming~~] impossible.

One important point shown by this and other remaining cases is that raising the whole CG to the matrix [Spec, TP] is the only way in which both the Case feature of the CG subject

and the Case feature of the CG itself can be valued in examples in which the CG is generated as complement of a passive verb or a raising verb (cf. later e.g. (56)-(58)). The fact that a *to*-infinitive does not carry a Case feature, as opposed to a CG, explains why the same pied piping process is not possible with infinitives, because it does not allow subsequent Case valuation in the matrix [Spec, TP2] of an overt subject DP (*Bill* in (55)) with *to*-infinitives. Despite its pied piping to the matrix [Spec, TP] the infinitival clause does not have the necessary feature specification to value Case of its own overt subject in (55):

- (55) a. *[TP2 [TP1 *Bill* to swim] is preferred].
 b. * [TP2 [TP1 *Bill* to swim] is/seems impossible].

Since a *to*-infinitive cannot be assigned Case, it cannot further value the Case of its embedded subject (*John* in (55)), and the derivation crashes.⁴⁶

Consider now other derivations in which each type of CG (with a null or an overt subject) is base generated as the complement of a matrix passive verb. Given the absence of an accusative Case feature on the matrix VP, the CG fails to have its Case valued, and cases such as (56) and (57) are always ungrammatical. Consider the explanation for why (56) is ungrammatical despite the fact that the DP subject *Bill* can have its Case valued in [Spec, TP2] (matrix clause). The DP *Bill* is base generated as the external argument of the CG, and moves to the matrix clause where it has its Case valued as nominative Case. Given this, why should the sentence then be ungrammatical? Given the analysis proposed here, there is only one Case checking head in the matrix clause (T), as it is standardly assumed for the passive of transitive verbs, and if this T head values the Case of *Bill*, this prevents the Case feature (C_{AGR}) of the CG from being checked.

(56) *Bill was preferred [swimming].

[TP2 Bill [T' [vP [v' θ [v' was preferred [TP1 Bill [T' AGR [vP Bill

ϕ /C/EPP

EPP *C_{AGR} θ

[v 'swimming]]]...

(57) *It was preferred [Bill swimming].⁴⁷

The existence of these ungrammatical cases provides further support for why T⁰ of CG itself needs to enter into Agree with the T head of the matrix clause, pied-piping the whole clausal gerund as in (53) to satisfy the EPP. Since there is only one Case position available in the matrix clause, it has to value the Case feature of the CG, which can in turn value the Case feature of its subject DP.⁴⁸ The required pied-piping of the CG for satisfaction of the matrix clause EPP, as in (53) and (54), allows the valuation of both Case features in the clausal gerund (i.e. the one in the head T⁰ of CG and the one in its DP subject).

A derivation along the same lines of (56) can account for the ungrammaticality of raising constructions with CGs, if it is assumed that the DP Bill in (58a) is also base generated as an argument of the CG and raises to the matrix clause for Case valuation, using up the only available Case feature in the matrix clause. This in turn prevents the CG itself from satisfying its Case requirement (for the alternative grammatical derivation in which the whole CG raises to matrix [Spec, TP] see (53)-(54)). As is well-known, ungrammaticality does not arise with standard raising-infinitive constructions when the embedded subject raises out of the embedded clause for Case valuation, as in (58b). This is due to the fact that infinitives, different from CGs, do not carry a Case feature, under the analysis developed here, and only the embedded DP Bill

needs to have a Case feature valued in (58b). In this respect, compare examples in which the whole infinitive would raise, yielding an ungrammatical derivation (55), for the same reasons.

- (58) a. *Bill seems [~~Bill~~-swimming well].⁴⁹
 b. Bill seems [~~Bill~~ to swim well].

In sum, the restrictive, unified analysis of CGs proposed for the different cases above has the advantage of accounting for a large number of apparently complex restrictions on the distribution of CGs in different syntactic contexts, representing a significant improvement over previous approaches, which did not consider the whole range of distinct cases analyzed here. Second, the complex phenomena analyzed here are reduced to individual feature properties of lexical heads in the derivation. Finally, this approach avoids appeal to unmotivated tense distinctions in the analysis of CGs, as discussed in section 3.

4.2 Absolute Gerunds

In this section I address briefly the properties of gerunds that occur as adjunct absolute clauses (59). Given that they also display an alternation between overt and null subject, we should in principle consider an analysis for them that is similar to the CGs I investigated so far:

- (59) a. Mike expected to win the game, [he/him/PRO being the best athlete in the school].
 b. [He/him/PRO being the best athlete in the school], Mike expected to win the game.

One might take gerunds as absolute clauses to have different properties from CGs, regarding how they satisfy Case requirements: If absolute gerunds need to be analyzed as CGs, the Case marking on them is not clearly dependent on subcategorization by an overt Case

checking head, different from CGs. However, a similar possibility also exists for certain topic DPs, which are not realized in a standard structural Case position (59), and which have their Case feature valued/checked in the course of the derivation, presumably as the result of the structural position in which they are realized.⁵⁰

- (60) a. This book, John told me it is interesting.
 b. Mike/him, I never met.

In other respects that I considered, Absolute gerunds share properties with the CGs I investigated. More specifically, when they display a null subject they consistently display properties of obligatory control. As I showed in section 3.2 for other CGs, in Absolute CGs the null subject is OC PRO. As evidence for this, it must have an antecedent (61a). The antecedent of OC PRO must be local (61b) and must c-command PRO (61c). OC PRO only permits a sloppy interpretation under ellipsis (61d) and cannot have split antecedents (61e). Finally, in cases involving *only NP* constructions (Fodor 1975), the binder of PRO must be the expression formed by *only + NP* (61f).⁵¹

- (61) a. *It was expected to start the concert soon, PRO having turned the lights off.
- b. John_j told Peter_k that Mary_m would arrive on time, PRO_{?j/*k/m} being responsible for starting the conference.⁵²
- c. Peter_k's daughter_j went on to college, PRO_{j/*k} being the best student in the class.
- d. Having kissed Mary at the door, Peter left the party with some friends, and Bill did too.(= Bill kissed Mary and left).
- e. PRO_{*j+k} understanding the importance of a good education, Peter_j expected his son_k to go to a good college.
- f. PRO_{j,*k} Having given the BST speech, only Churchill_j was congratulated by everybody.

4.3 Gerunds without the overt vs. null subject alternation

The analysis in this paper focused on a class of clausal gerunds (CGs) that allow an alternation between null subjects and overt (accusative/nominative) subjects in the same structure, and which I treated as clausal gerunds (e.g. (1)). However, there are other gerunds that are distinct both from *poss-ing* and from the CGs I considered here in that they allow either only an overt subject or only a null subject. I mention some of these cases below, but I do not go into details regarding their analysis here, due to space limitations. First, one example of gerunds that may seem to allow only overt subjects corresponds to gerunds as complements of perception verbs (62a). Second, an example of gerunds that allow only null (obligatory control), as in (62b) was used as additional evidence for the evaluation of the null Case theory in section 3.1, and I have referred to them elsewhere as TP-defective gerunds.

- (62) a. Bill saw Jane/**ec* leaving the house.
 b. Frank tried *ec*/*Jane working at home.

In Pires (2001a,b, forthcoming) I showed that these two types of gerunds – gerunds as perception verb complements (62a) and TP-defective gerunds (62b) – are clearly distinct from the clausal gerunds I analyze in this paper, as shown first by the fact that only CGs display the overt vs. null subject alternation. Second, in addition to what I discussed in section 3.1, and different from clausal gerunds, neither type of gerund in (62) carries a [+tense] feature, since neither one allows a temporal specification distinct from the matrix clause:

- (63) a. **Tomorrow* Bill will see [Jane leaving town *today*]
 b. * Bill tried *today* [talking to his boss *tomorrow*].

For these and other reasons, in Pires (2001a, forthcoming) I proposed for the two of gerunds in (62)-(63) an analysis that is distinct from that of clausal gerunds, and which explores the empirical fact that they do not display an alternation between overt and null subjects. Crucially, in both cases in (62) I take the head of the gerund itself not to be involved in Case checking/valuation of an embedded subject. Even with the PVC-gerund, I present evidence that what appears to be an embedded overt subject (*Jane* in (62a)) values/checks its Case directly in the matrix clause. Evidence for this comes from the fact that the embedded subject of the PVC gerund is affected by passivization of the matrix verb (this contrasts with clausal gerunds, which under passivization of the main verb block movement of the embedded subject alone to the matrix clause, as I showed in (56) and (57)).

- (64) a. *Bill/It was seen Jane leaving the house.
 b. Jane was seen leaving the house.

In addition, I argue that, like the gerunds in (62), other gerunds that also do not display an alternation between overt and null subjects should not be collapsed with CGs, which do display this alternation. This includes certain cases of gerunds that behave as adjuncts and which are introduced by *while* (65), as discussed originally by Reuland (1983). Crucially, the gerunds in (65) only allow a null subject, different from clausal gerunds in adjunct position (66), which display the overt/null subject alternation (and are arguably introduced by a preposition, e.g. *without*):

- (65) a. *John kept walking slowly, [while the plane drenching the road with insecticides]
 b. John kept walking slowly, [while PRO drenching the road with insecticides].⁵³
- (66) a. Bill_j left [without him*_j having finished the report].
 b. Bill_j left [without PRO_j having finished the report].

Given the contrast above, I take cases such as (65) not to display the same properties of clausal gerunds I proposed in this paper. However, given that the null subject PRO in (65) seems to display obligatory control PRO, it can in principle be analyzed under a movement analysis of control, similar to the CG in (66b). Given that the extension of the movement analysis to these cases involves proposing that movement out of non-finite adjuncts is possible, I put aside here the details of such an analysis (which is proposed in detail in Hornstein 1999, 2001; see also Pires 2001, forthcoming, regarding the case of gerunds).^{54, 55}

5 Conclusion

This paper analyzed the special properties of clausal gerunds (CGs) in English and proposed an analysis that attempted to explain the restrictions in their distribution, accounting at the same time for their ability to license both overt and null subjects, despite the lack of syntactic distinctions between CGs in both instances. The existence of structures such as CGs which license overt subjects or control null subjects exactly in the same context raises significant problems for theories of Case and of Control. This paper addressed these problems by proposing an analysis in which the possibility of the two types of subjects in CGs results from the interaction of the same grammatical mechanisms, the ones involved in Case and agreement valuation. The paper also points out empirical problems raised by CGs for null Case/tense-based approaches to Control, and adopts instead an A-movement analysis of control CGs, without appealing to distinctions between the feature specification of CGs with null (control) subjects and CGs with overt subjects.

References

- Abney, S. 1987. The English noun phrase in its sentential aspect. Ph.D. dissertation, MIT.
- Aoun, J, L. Choueiri and N. Hornstein. 2001. Resumption, Movement, and Derivational Economy. *Linguistic Inquiry* 32.3:371–403.
- Barrie, M. 2004 Moving towards partial control. In K. Moulton and M. Wolf. *Proceedings of the 34th Meeting of the Northeast Linguistics Society*, 133-46.
- Boeckx, C. and N. Honstein. 2004. Movement under Control. *Linguistic Inquiry* 35: 431-52.
- Belleti, A. 1990. *Generalized Verb Movement*. Torino: Rosenberg & Sellier.
- Bošković, Ž. 1994. D-Structure, Theta Criterion, and movement into theta positions. *Linguistic Analysis* 24:247-286.

- Bošković, Ž. 1997. *The Syntax of Nonfinite Complementation: an Economy Approach*. Cambridge, Mass.: MIT Press.
- Bošković, Ž, and D. Takahashi. 1998. Scrambling and Last Resort. *Linguistic Inquiry* 29:347-366.
- Bošković, Ž. 2002. A-movement and the EPP. *Syntax* 5: 167-218.
- Castillo, J. C., J. Drury and K. K. Grohmann. 1999. Merge over Move and the Extended Projection Principle. In *University of Maryland Working Papers in Linguistics* 8:66-103.
- Chomsky, N. 1970. Remarks on Nominalization. In *Readings in English Transformational Grammar*, ed. Jacobs and Rosenbaum, 184-221. Ginn Press.
- Chomsky, N. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.
- Chomsky, N. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.
- Chomsky, N. 2000. Minimalist inquiries: The framework. In *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, ed. R. Martin, D. Michaels, and J. Uriagereka. Cambridge, MA: MIT Press.
- Chomsky, N. 2001. *Derivation by phase*. In Kenstowicz, ed. 1-52.
- Chomsky, N. and H. Lasnik. 1993. Principles and Parameters theory. In Chomsky 1995.
- Culicover, P. W. and R. Jackendoff. 2001. Control is not movement. *Linguistic Inquiry* 32: 493-512.
- Davies, W. D., and S. Dubinsky. 2004. *The grammar of raising and control: A course in syntactic argumentation*. Malden, MA: Blackwell.
- Enç, M. 1990. On the absence of the present tense morpheme in English. Ms., University of Wisconsin, Madison.
- Epstein, S. D., E. Groat, R. Kawashima, and H. Kitahara. 1998. *A derivational approach to syntactic relations*. New York: Oxford University Press.

- Epstein, S. D. and N. Hornstein. eds. 1999. *Working Minimalism*. Cambridge, MA: MIT Press.
- Epstein, S. D. and D. Seely. 1999. SPEC-ifying the GF “subject”: Eliminating A-chains and the EPP within a derivational model, ms., University of Michigan and Eastern Michigan University.
- Epstein, S. D., A. Pires and T. D. Seely. 2005. EPP in T: More controversial subjects. *Syntax* 8: 165-80.
- Epstein, S. D. and D. Seely. 2006. *Derivations in Minimalism: Exploring the Elimination of A-Chains and the EPP*. Cambridge, Cambridge University Press.
- Fodor, J. 1975. *The Language of Thought*. New York: Thomas Y. Crowell.
- George, L. and J. Kornfilt. 1981. Finiteness and boundedness in Turkish. In *Binding and filtering*, ed. F. Heny. London: Croom Helm.
- Grimshaw, J. 2000. Locality and extended projection. In Coopmans, P., M. Everaert and J. Grimshaw, eds. *Lexical specification and insertion*. Amsterdam: John Benjamins.
- Higginbotham, J. 1992. Reference and control. In Larson et alii, eds.
- Horn, L. 1975. On the nonsentential nature of the POSS-ING construction. *Linguistic Analysis* 1.4: 333-387.
- Hornstein, N. 1999. Movement and control. *Linguistic Inquiry* 30:69-96.
- Hornstein, N. 2001. *Move! A Minimalist Theory of Construal*. Oxford: Blackwell.
- Jackendoff, R. & P. W. Culicover. 2003. The Semantic Basis of Control in English. *Language* 79:517-56.
- Johnson, K. 1988. Clausal gerunds, the ECP, and Government. *Linguistic Inquiry* 19.4:583-609.
- Kaiser, L. 1999. *The morphosyntax of clausal nominalization constructions*. Ph.D. dissertation. University of Yale, New Haven.

- Kapetangianni, K. and T. D. Seely. 2003. Subjunctives in Modern Greek: explanation by deduction; agreement by degrees. Paper presented at *NASC 1 – North American Syntax Conference*. Concordia University, Montreal.
- Kayne, R. 1981. ECP extensions. *Linguistic Inquiry* 12:93-133.
- Kenstowicz, M., ed. 2001. *Ken Hale: A life in language*. Cambridge, MA: MIT Press.
- Kiguchi, H. 2002. PRO Gate and Movement. in *Proceedings of the 25th Pennsylvania Linguistics Colloquium*, University of Pennsylvania, Philadelphia, PA.
- Landau, I. 2003. Movement out of control. *Linguistic Inquiry* 34: 471-98.
- Landau, I. 2004. The scale of finiteness and the calculus of control. *Natural Language and Linguistic Theory* 22: 811-77.
- Larson, R., S. Iatridou, U. Lahiri and J. Higginbotham. eds. 1992. *Control and grammar*. Dordrecht: Kluwer.
- Lasnik, H. 1995. Last Resort and Attract F. In *Proceedings of FLSM 6*, 62-81. Indiana University Linguistics Club, Bloomington, Indiana.
- Lasnik, H. 1999. *Minimalist analysis*. Malden, MA: Blackwell.
- Lasnik, H. 2001a. A Note on the EPP. *Linguistic Inquiry* 32: 356-361.
- Lasnik, H. 2001b. Subjects, Objects, and the EPP. In W. Davies and S. Dubinsky (eds) *Objects and other subjects*. Dordrecht: Kluwer.
- Lebeaux, D. 1985. Locality and anaphoric binding. *The Linguistic Review* 4:343-363.
- Manzini, M. R. and A. Roussou. 2000. A minimalist theory of A-movement and control. *Lingua* 110:409-447.
- Martin, R. 1996. A Minimalist theory of PRO and control. Ph.D. dissertation, University of Connecticut, Storrs.

- Martin, R. 1999. Case, The Extended Projection Principle, and Minimalism. In Epstein and Hornstein.
- Martin, R. 2001. Null Case and the distribution of PRO. *Linguistic Inquiry* 32.1:141-66.
- Milsark, G. L. 1988. Singl-ing. *Linguistic Inquiry* 19.4: 611-634.
- Monahan, P. 2003. Backward object control in Korean. *Proceedings of the West Coast Formal Conference in Linguistics* 22:101-14.
- O'Neil, J. 1995. Out of control. In *Proceedings of NELS 25*: 361-371. GLSA, University of Massachusetts, Amherst.
- Pesetsky, D. and E. Torrego. 2001. T-to-C Movement: Causes and Consequences. In Kenstowicz, ed. 355-426.
- Pires, A. 1999. A minimalist approach to clausal gerunds. Ms. University of Maryland, College Park.
- Pires, A. 2001a. The Syntax of Gerunds and Infinitives: Subjects, Case and Control. Ph.D. dissertation. University of Maryland, College Park.
- Pires, A. 2001b. Clausal and TP-Defective Gerunds: Control without Tense. In *Proceedings of NELS 31*, GLSA, University of Massachusetts, Amherst.
- Pires, A. in press. *The Minimalist Syntax of Defective Domains: Gerunds and Infinitives*. Amsterdam: John Benjamins.
- Pires, A. and C. Rodrigues. 2002. Null Subjects of Non-finite Adjuncts: A Case of Remnant Movement. Paper presented at the 76th Annual Meeting of the Linguistic Society of America, San Francisco, CA.
- Polinsky, M. & E. Potsdam. 2002. Backward Control. *Linguistic Inquiry* 33: 245-82.
- Reuland, E. 1983. Governing -ing. *Linguistic Inquiry* 14.1: 101-136.

- Reuland, E. and S. Avrutin. 2005. Binding and Beyond: Issues in Backward Anaphora. In A. Branco et al. (eds.) *Anaphora Processing: Linguistic, cognitive and computational modelling*. Amsterdam: John Benjamins.
- Rizzi, L. 1990. On the anaphor-agreement effect. *Rivista di Linguistica* 2:27-42.
- Rodrigues, C. 2004. Impoverished morphology and A-movement out of Case domains. PhD dissertation, University of Maryland, College Park.
- Rouveret, A. 1979. Sur la notion de proposition finie. *Langages* 60.
- Safir, K. 2004a. *The Syntax of (In)dependence*. Cambridge, MA: MIT Press.
- Safir, K. 2004b. *The Syntax of Anaphora*. New York: Oxford University Press
- San Martin, I. 2004. On subordination and the distribution of PRO. Ph.D. dissertation, University of Maryland, College Park.
- Schueler, D. 2004. Extended-projection, categorization and the English morpheme –ing. *Paper presented at WECOL 2004*.
- Schütze, C. 2001. On the nature of default Case. *Syntax* 4:205-38.
- Stowell, T. 1981. *Origins of Phrase Structure*. Ph.D. dissertation, MIT.
- Stowell, T. 1982. The Tense of Infinitives. *Linguistic Inquiry* 13:561-570.
- Uriagereka, J. 1995. Aspects of the Syntax of Clitic Placement in Western Romance. *Linguistic Inquiry* 26.1: 79-123.
- Williams, E. 1975. Small clauses in English. In *Syntax and Semantics*, Vol. 4, ed. J. Kimball. New York: Academic Press.
- Zubizarreta, M. L. 1980. Remarks on inflected infinitives in Portuguese. Ms. MIT.

Acrisio Pires

University of Michigan, Ann Arbor - Dept. of Linguistics

1177 Undergraduate Science Bldg. - 204 Washtenaw Av.

Ann Arbor, MI 48109-2215

pires@umich.edu

*Core aspects of the analysis in this paper were presented at different venues. I thank especially audiences at *GLOW*, *NELS*, the *Germanic Linguistics Annual Conference*, University of Maryland, Michigan State University, University of Michigan and Wayne State University for relevant feedback. Thanks, for various useful comments and suggestions, especially to Norbert Hornstein, David Lightfoot, Juan Uriagereka, Zeljko Boskovic, Jan-Wouter Zwart, Sam Epstein, Daniel Seely and also to Howard Lasnik, Esther Torrego, Amy Weinberg, Cilene Rodrigues, Kleanthes Grohmann, Carson Schütze, Itziar San Martin, Ljiljana Progovac, and three anonymous reviewers of *Syntax*. Many thanks to the editors of *Syntax*, especially David Adger and Elizabeth Laurençot, for all their help since the paper was submitted for publication. Thanks for judgments or comments on the data to Peggy Antonisse, Laura Benua, Christopher Becker, Stephen Crain, Charles Crissman, Cati Fortin, Stephan Greene, Jeff Lilly, Paul Pietroski, Daniel Shannon, Heather Taylor and Mark Wilson. This paper develops specific aspects of a broader analysis of English gerunds in Pires 2006, parts of which are presented here with kind permission from John Benjamins.

¹ Gerunds themselves are part of a large set of structures with the morphological structure *V-ing* (where *V* stands for a verb root) that include among others the progressive form ('John is sleeping'), adjectival modifiers ('an interesting person'). These are outside the scope of this paper.

² Notice that in this respect, the *V-ing* constructions in (i) pattern with regular DPs, and cannot be collapsed with CGs or with *poss-ing* constructions (see Chomsky 1970, Abney 1987). Contrary to CGs and *poss-ing*, a VP-adverbial is not possible in (ic):

- (i) a. Mary's revising of the book.
- b. Mary's quick revising of the book.
- c. *Mary's quickly revising of the book.

³ The tests in this section are illustrated with the distinction between CGs and *poss-ing*, but DPs and *poss-ing* pattern together here, and differently from CGs.

⁴ Consider also cases with expletive *it*, in which a contrast similar to (9) holds:

- (i) I wouldn't count on *it*(*s) raining tomorrow.

In different approaches that appeal to EPP requirements, besides lexical DPs and overt expletives, empty categories such as control PRO and arbitrary control PRO can also satisfy the EPP requirement (ii):

- (ii) a. You_i may count on PRO_i winning the game tonight.

b. PRO_{arb} arriving on time is what matters now.

See however, Epstein & Seely 1999, 2006, Martin 1999, Castillo, Drury & Grohmann 1999, Bošković 2002, Epstein, Pires & Seely 2005 and references therein, for attempts to derive the EPP effects from other properties of the grammar.

⁵ Abney (1987) proposes an external DP treatment for *acc-ing* (CGs with an accusative overt-subject), but he points out a host of other properties that distinguish *acc-ing* from *poss-ing* and from regular DPs, including the impossibility with CGs (but not with *poss-ing*/DP) of long-distance binding of their subjects (i) (Abney 1987:175-6). These and other differences further support the treatment of CGs as distinct from *poss-ing* and from regular DPs (see section 4):

(i) a. *They thought that each other's paying the bill was acceptable.

b. They thought that each other paying the bill was acceptable.

⁶ I assume that PP adjunct CGs behave in the relevant respects as complement CGs. See section 4.2 for some discussion about absolute CGs.

⁷ Section 4 will analyze in detail the derivation of different cases that are ungrammatical due to absence of a source of Case for the clausal gerunds.

⁸ However, it is not clear why *it*-extraposition should be impossible for Case-marked expressions. Reuland (1983) proposes a possible GB account which awaits detailed re-evaluation in the context of the Minimalist Program. One possibility is that the expletive *it* also needs to have a Case feature valued (see fn. 47). Notice furthermore that the restriction on *it*-extraposition of DPs does not extend to heavy-NP shift, raising a question whether the two are closely related phenomena.

⁹ Pires (in press, p. 22-3) considers two other cases where the occurrence of CGs and *poss-ing* (together with DPs) is accepted by most speakers, but where finite and infinitive clauses are not accepted: subject position of interrogative questions and cleft focus position.

¹⁰ In Abney's (1987:223, 226) analysis, *-ing* in *acc-ing* constructions is affixed to an IP to project a DP, but crucially for Abney, there is no D-head in *acc-ing* cases. Abney does not clarify how this can be compatible with the X-bar Theory approach he adopts, based on endocentricity of X-bar projections (Stowell 1981). Recently, Schueler (2004) proposed a DP-analysis of gerunds that is similar to Abney's, but which avoids the endocentricity violation by adopting the view that gerunds are nominalizations at different {F}(shell) levels (Grimshaw 2000). A

DP-analysis of CGs is not necessarily incompatible with the main tenets of the analysis of CGs with null subjects and with overt subjects that I focus on in this paper, provided a revised DP-analysis of overt-subject CGs can be compatible with the analysis of null-subject CGs as well (1a).

¹¹ T⁰ will be taken to carry any tense or inflectional features that the CG may have, given that no separate AgrP is necessary below or above TP to account for the empirical facts dealt with in this paper (see also Chomsky 1995, 2000).

¹² The idea of tensedness and finiteness as separate parameters goes back to George & Kornfilt 1981, Rouveret 1979 and Zubizarreta 1980.

¹³ Reuland's 1983 structural relations were also grounded primarily on different properties of Government, whose consideration I put aside in the analysis developed in this paper.

¹⁴ But see section 3.1, example (26) for a type of gerund (TP-defective gerunds) that behaves differently from CG in this respect, by showing a [-tense] specification (see Stowell 1982, Pires 2001b, in press).

¹⁵ There is one other important empirical advantage of this account, in the treatment of control clauses with gerunds, since tense distinctions are also not sufficient there, in the way they are used in the Null Case approach to Control (see section 3.1 for detailed discussion of some relevant problems for this approach).

¹⁶ In addition, according to my first reviewer, the present proposal improves on Reuland's approach at the conceptual level. She states that although the current analysis appeals to the mechanisms that will be later specified in (48), it is more principled than the approach developed in Reuland (1983), which according to the reviewer "links the PRO-ing option to affix-hopping in an obscure way."

¹⁷ Although intermediate wh-movement to the edge of the first available (strong) phase (in this case, matrix vP) may be required in the course of long-distance wh-movement.

¹⁸ In Pires (2001b, in press) I analyze the gerunds in (26a-b) as TP-defective gerunds and treat them differently from clausal gerunds (for instance, CGs are different regarding their tense specification, in that they are [+tense], see examples (19) and (28)). See section 4.3 for relevant discussion.

¹⁹ Furthermore, they cannot be distinguished regarding the possible projection of a CP (taken to be a distinguishing element between control and non-control infinitive clauses in early P&P approaches), given that a CP is absent in CGs in general (see arguments in section 2).

²⁰ One might take the latter to be case because of the use of the present tense in the matrix clause in (30). However, this explanation cannot be adopted under Martin's analysis, because for him the feature specification of a control clause has to be sufficient to force the individuated event reading, and should not be dependent on the tense specification of the matrix clause, contrary to what is in fact shown in (30). In addition, other examples with present tense matrix verbs that take control CGs as complements such as (i) do not block the individuated-event interpretation, even in the absence of overt event binding operators.

- (i) a. I remember [dreaming of sirens (soon after midnight yesterday)].
 b. I count on [leaving tomorrow at 10a].

²¹ My third reviewer states that "this paper only argues against a particular implementation [of null case], not against the approach in principle." To my knowledge, there is no other approach to null Case that avoids the empirical problems I raise here. It is clearly possible to expect that other logically possible approaches to null Case could be proposed, but in the absence of such approaches, I take the criticism presented here to be sufficient, within the scope of this paper.

²² Alternative approaches to control have more recently been the object of significant debate in the literature. Additional references representing different perspectives include Culicover & Jackendoff 2001, Polinsky & Potsdam 2002, Jackendoff & Culicover 2003, Landau 2003, 2004, Kapetangianni & Seely 2003, Monahan 2003, Barrie 2004, Boeckx and Hornstein 2004, Davies & Dubinsky 2004, San Martin 2004, Rodrigues 2004 and references therein. Safir 2004a,b (mentioned by my third reviewer) presents arguments against the treatment of anaphora as resulting from movement, but to my knowledge does not discuss control in detail.

²³ Tests (i-v) were originally presented in Lebeaux (1985). Infinitive examples (a) are from Hornstein (1999).

²⁴ Notice that this and other tests in this section are relevant to establish the properties of CGs that display null subjects. In the presence of an overt subject in the embedded CG, as in (i), the VP-ellipsis site (marked by *did too*) is interpreted as having the same overt subject (*Mary*) as in the antecedent CG [TP1 in (i)], instead of establishing identity with a c-commanding antecedent (either *John* or *Bill*), which would only be possible in case the antecedent CG itself had a null subject that needed to be interpreted in the ellipsis site. This contrast results from constraints on the properties of VP-ellipsis, and not from distinctive properties of null vs. overt subject CGs, as can be seen by the similar behavior of an infinitival complement of *want* in the presence of an overt subject:

(i) John worried [TP1 about Mary losing] and Bill did too (=Bill worried about Mary (not John, not Bill) losing).

(ii) John wanted [TP1 Mary to leave] and Bill did too. (= Bill wanted Mary to leave (not, John, not Bill)).

²⁵ Adjuncts introduced by a preposition cannot usually be tested for the *de se* interpretation. Consider, however, a pair that may show the relevant distinction:

(ii) a. The unfortunate_k was pleased after [*e*_k getting a medal].

b. The unfortunate_k was pleased after [*he*_k got a medal].

In an interpretation in which there is a causal relation between the embedded clause event (to get a medal) and the matrix clause event (being pleased), the *de se* interpretation is obligatory in (iia), but not in (iib), in the same scenario considered in the text.

²⁶ My first reviewer suggests that the fact that “OC PRO has a *de se* reading is not particularly good evidence that movement is involved.” According to her, “*de se* interpretations are found in many languages even with overt pronouns, and conversely, a semantic analysis of *de se* deriving from movement is not obviously more natural than one for a separate element.” The reviewer does not indicate the languages in which this is the case. In any case, she tacitly assumes that only null elements can be derived by movement, and does not take into consideration two different points: (i) in the absence of a principled analysis that clearly distinguishes obligatory control PRO from non-obligatory control PRO (such as the one that I further argue for in this paper, by means of a movement analysis for OC PRO), it remains mysterious why only domains that display OC PRO should allow *de se* readings, contrary to NOC PRO (as shown in section 3.3); (ii) one cannot rule out the possibility that even overt elements can be shown to be related to an antecedent by means of a movement analysis; this has been proposed for different kinds of overt elements, including for instance resumptive pronouns (see e.g. Aoun et al. 2001) and clitics in connection to their doubles in clitic doubling (e.g. Uriagereka 1995 and references therein).

²⁷ Kiguchi (2002) analyzes a subset of gerunds in subject position that actually display obligatory control properties. He also proposes an analysis of OC PRO in these constructions as the result of A-movement.

²⁸ See also (9) and fn. 4.

²⁹ One might raise the question whether there is independent evidence for the properties in (48). First, to my knowledge, there is no analysis of CGs that has been able to dispense entirely with the need to instantiate these two properties by means of some special mechanism (for instance, Reuland’s very insightful analysis instantiates

both properties by other mechanisms, but proposes a significantly complex architecture especially to deal with certain CGs, with additional empirical complications such as the ones regarding the treatment of tense, as I discussed in section 2). I discussed an alternative to (48i) in section 2, by reviewing briefly the alternative of adding a DP projection above TP, although I pointed out different problems in treating CGs as DPs. Given those complications, I won't consider this alternative in the derivations that follow. Pesetsky & Torrego (2001) propose a widespread connection between Case and Tense features, by suggesting that nominative Case is in fact an uninterpretable T feature on D. This connection between Case and an uninterpretable T feature may bear on the approach I adopt, although I don't explore the possible connections here. Regarding (48ii), it may not be restricted to CGs and may have some cross-linguistic correlates. Pires (in press, p. 65-8, ch. 3) discusses examples of non-finite constructions in other languages (e.g. Basque, Portuguese and Quechua) which may instantiate similar mechanisms, in which a certain inflectional head can value the Case of a DP only if this inflectional head itself appears in the domain of a Case-valuing probe.

The advantage of the approach I propose in section 4 is that, by appealing to these two properties of CGs in interaction with a general architecture of feature valuation, this approach may effectively account for a host of facts in the complex distribution of CGs that so far have not been entirely analyzed.

³⁰ Chomsky (1995:232) claims that the EPP (Extended Projection Principle) corresponds to a strong D-feature of I. In Chomsky (2001:7) EPP is treated as an uninterpretable selectional feature of a syntactic object (of T and v , in the alternative he explores in detail). I do not focus here on evaluating different alternatives to deal with EPP requirements; see also in this respect fn. 4.

³¹ This interaction arises in a system where ϕ -features are the attractors (Chomsky 2000) or what makes the probe active for Match/Agree (Chomsky 2001).

³² The fact that T is ϕ -defective does not prevent Match with the full ϕ -set of DP from applying. Although Chomsky (2001:4) considers identity to be the optimal candidate for Match, he argues that Match is not strictly speaking identity, but non-distinctness. For Match to occur, probe and goal must share the same features, independent of value. In this case, the ϕ -features of T on a clausal gerund are simply unvalued when DP-movement takes place to satisfy the EPP. That is, they are non-distinct from the ones on the DP, as in other instances of T_{def} , thus Match between the ϕ -defective T and the DP can apply.

³³ However, given the phase approach of Chomsky 2001, I assume that the uninterpretable ϕ -features of AGR are not deleted immediately after being valued. This is still consistent with the idea that uninterpretable features must be deleted. This will be more especially relevant in the derivation of (51), in which the uninterpretable ϕ -features and Case feature of AGR as well as the Case feature of the CG subject will be deleted only at the end of the first available strong phase (the matrix ν P) . I assume this is compatible with the phase approach in Chomsky 2001: valued uninterpretable features do not need to be deleted before the end of their strong phase. Alternatively, uninterpretable features that have been valued remain active until the phase is no longer available for further computation.

³⁴ A question arises why this elimination of the defectiveness of ϕ -set of a CG does not occur in other instances of Match/Agree, such as with infinitives. Crucially, infinitive T in English does not allow overt subjects unless their Case is valued by a higher inflectional head (e.g. under ECM). This follows from the fact that the infinitives are consistently ϕ -defective. Contrary to this, it is plausible that AGR (the T^0 of a CG) gets its ϕ -set to become non-defective after Match/Agree with the DP in its Spec because of the nominal character of AGR, formalized here by the fact that it also carries an uninterpretable Case feature that needs to be checked (48i) (see discussion in section 2). The precise distinction in ϕ -feature specification is hard to pin down here, given that CGs share properties with both infinitives (the possibility of null subjects) and with finite clauses (the possibility of overt subjects) at the same time, and for this reason cannot be identical to either one regarding Case valuation.

³⁵ The analysis represents movement by generating additional copies which are deleted in the derivation, but no significant aspect of the analysis hinges on this approach to movement.

³⁶ Presumably it has an internal θ -role as well, which is assigned to the embedded CG, although this is not shown in the derivation.

³⁷ As Chomsky (1995:233) proposes “[A strong feature] induces cyclicity: [it] cannot be “passed” by α that would satisfy it, and later checked by β ; that would permit Relativized Minimality violations (wh-island, super-raising).” Putting aside complications that should not come into play here regarding overt and covert syntax, a strong feature corresponds in the present analysis to an uninterpretable, unvalued feature.

Specifically at the point in (50c), the Case feature of the embedded subject *John* has to “wait” to be valued only after it moves to the matrix clause because at the point *John* raises to the matrix ν P there is no more local probe that can value its Case in the derivation (given that the CG head T cannot value Case while it still carries an

uninterpretable Case feature of its own — see also (50c)' for why an alternative derivation with different order of steps is ungrammatical).

³⁸ However, see in fn. 45 an alternative approach to the issue of locality and equidistance in such cases.

³⁹ Notice that a derivation along the lines of (50'), but in which an expletive is merged in the matrix [Spec, TP] (i), would still crash, possibly because expletive *there* cannot be inserted in a thematic position, or it is blocked from taking the CG as an associate, which is explained in the current analysis because the CG is not a DP, and does not carry the properties restricted to indefinite DPs required in existential-*there* structures. Crucially, the matrix T cannot directly probe the indefinite DP *a man* as its goal, because this would prevent the CG itself from having its own Case feature valued, yielding a non-convergent derivation.

(i) *There prefers [a man swimming]

The restriction on *there*-expletives with CGs as in (i) lends support to the view that *there* expletives are only allowed in a subset of structural Case checking configurations and are not allowed in all contexts in which EPP and structural Case checking come into play. This is consistent with the fact that besides (i), neither case in (ii) is possible, despite the fact that expletive *there* can satisfy any EPP requirement whereas the associate-DP (*a man, two kids*) could be generated in its base thematic position and have its Case valued by Match/Agree (Chomsky 2000, 2001, Lasnik 1999, Lasnik 2001a,b and references therein; see Chomsky 1995 for earlier analyses involving a *there*-associate relation):

(ii) a. *[TP There [vP a man danced]].

b. *[TP I [vP expected [TP there to [vP two kids enjoy the movie]]]].

⁴⁰ This is similar to the rationale adopted in the derivation of existential clauses such as (i), which have been the focus of significant attention in the P&P literature. Both cases have a common derivation only up to the point at which the embedded T is inserted. Given the existence of *there* in the numeration of (ia), it does not compete with (ib), which takes place independently (see e.g. Chomsky 1995, 2000, 2001). This is like the contrast between (50) and (51): there is no issue of optionality in the steps of each derivation in either pair of cases.

(i) a. there seems [to be a man outside].

b. a man seems [to be ~~a man~~ outside].

⁴¹ see fn. 33 regarding deletion of valued features (in this case ϕ -features) only at the (strong) phase-level. See fn. 45 regarding why no locality violations arise in the choice of either the embedded subject DP or the embedded T head as accessible goals for Match/Agree.

⁴² My first reviewer suggests that at certain steps in the derivation there is local optionality at play, but again this is only apparent, given that true optionality here would involve the availability of two equally convergent paths in a derivation with a single numeration, which is not the case (as I also showed with (50) and its *non-convergent* counterpart (50') with the same numeration and a different derivational path).

⁴³ An analysis of clausal gerunds that distinguishes their derivation from ECM has other potential advantages, especially if coupled with recent approaches to accusative Case checking/valuation under ECM. Different minimalist analyses have proposed that under ECM the subject of the embedded clause in fact checks/values its Accusative case in the Spec, vP of the matrix clause (see e.g. Chomsky 1995, Lasnik 1999 and references therein). This is distinct from what is proposed here for CGs, in which the subject of the CG does not undergo ECM nor raises (overtly or covertly) to the matrix clause in cases such in (51). This may provide a straightforward way to account for the observation that a quantified NP in the subject of a CG complement strongly favors a narrow scope interpretation, as in (ii)/(iiib) different from object DPs (i)/(iiia) and subjects of ECM complements (as observed by Reuland 1983:111; thanks to my third reviewer for raising this issue). The distinctive behavior regarding quantifier scope can be established by tracking down different steps of (A-)movement in the syntax. Assuming that object DPs and subjects of ECM-complements would raise to spec, vP of the matrix clause, contrary to the overt subjects of CG complements, DP objects and ECM-subjects can in principle occur in a higher position from which wide scope in relation to the matrix verb can be established:

(i) I hated everyone.

(ii) I hated [everyone I liked being hanged.]

(iii) a. (every x) I hated x.

b. I hated [(every x) [x is hanged]. (adapted from Reuland 1983:(23))

In addition, the approach suggested here departs significantly from the approach that was proposed in Reuland (1983), who proposed a Government approach to these phenomena. However, contrary to the standard Government approach regarding Case assignment, by which a lexical head could assign Case only to an XP that it governed, Reuland instead argued that a head could assign Case to an XP that it did not Govern: “in the exceptional

Case-marking constructions, the matrix verb is understood to directly govern the subject of the complement across S. In NP-ing constructions [CGs], the matrix verb does not govern the subject of the complement at all; rather, only its Case ends up there.” (Reuland 1983:120).

⁴⁴ Pires (in press, p. 69-71) discusses the question why the morphological realization of Case of the embedded subject in (51) is accusative (in the case of a pronoun), despite the fact that it does not involve ECM (see also Schütze 2001 and references therein). In this respect, my first reviewer states that “Reuland (1983) proposed that the Case assigned to the gerund clause was *transferred* to the subject. [Reuland’s 1983] proposal makes incorrect predictions for gerunds in subject positions (in English at least), which it must rescue through a [...] stipulation:

- (i) Him/*He winning the race upset me.”

As the reviewer further points out: “While the current paper’s analysis does not *explain* the existence of accusative subjects even in gerunds in subject position, it at least does not directly predict the opposite, since it merely says that the Case-valuing ability of the T head is dependent on its own Case features being valued; not that the values themselves are transferred.”

⁴⁵ One issue arises in the case of pied-piping of the whole CG to the subject position of the matrix clause. What prevents the matrix T from probing the embedded subject DP *Bill* as its goal? Under a representational view, Bill c-commands AGR (the head T of CG), and AGR does not c-command *Bill*, so in terms of c-command Bill should be closest for Match/Agree with matrix T. So, the Match/Agree between matrix T and the embedded DP *Bill* would be expected, but in fact it would yield a non-convergent derivation corresponding to (56), which will be analyzed in detail later. But the question remains why both *Bill* and the embedded AGR of CG are accessible goals for the matrix probe T, so that the convergent derivation in (53) is in fact possible, under which it is the embedded AGR that works as the goal for matrix T. There are two alternative solutions for this problem. First, it is possible to adopt a derivational c-command approach (Epstein et al. 1998) to the relation between Bill and AGR of CG, by which they mutually c-command each other, given that at an earlier point in the derivation T c-commanded *Bill*. Second, it is possible to adopt an approach by which the matrix probe T in fact attracts a feature of the whole embedded TP of CG, and not just of its head independently. This is the approach to attract adopted for instance by Pesetsky & Torrego (2001:363 (13)), as their *Head Movement Generalization*. Under this approach, given that matrix T is in fact targeting either the embedded DP *Bill* or the embedded TP of CG, the two goals are equidistant to

matrix T, if closeness is still subsumed under c-command, and not under domination (that is, neither *Bill* nor the embedded TP c-command each other in (53), because the TP dominates *Bill*; therefore, both are equidistant to the matrix probe). Either the *Derivational C-command* approach or the *Head Movement Generalization* approach can help explain why the derivation of (51) does not induce a locality violation in the application of Match/Agree.

⁴⁶ However, English has alternative cases (i) in which *for* can value the Case of the embedded subject of infinitives in the same position as (55):

- (i) [*For* Bill to swim] is/seems impossible.

Assuming that *for* is generated as a complementizer, it is possible to explain why this alternative is not available for CGs in positions in which they would not otherwise have their Case valued, given the fact that they do not project a CP (see section 2):

- (ii) a. *It is impossible/preferred [*for* Bill swimming]. (see also (57)).

- b. *Bill was preferred [*for* swimming]. (see also (56)).

⁴⁷ In (57), it is also assumed that Agree and Case valuation cannot occur in instances in which the category that needs to value its Case does not move overtly, even when an expletive *it* could be inserted in Spec, TP of the higher clause. One possible solution to this problem is to assume that, under an Agree approach to expletive-associate structures, expletive *it* is assigned Case and the extraposed clause remains caseless. This provides an explanation for why *it*-extraposition is impossible with CGs (see section 2). The same kind of approach may be necessary to rule out cases such as (i), in which one possibility is that the DP *a man* would not have its Case valued, given that the matrix T can have its feature requirements satisfied and become inactive by insertion of the expletive *it*, which additionally values its Case in the matrix [Spec, TP]

- (i) *It seems [to be a man in the yard].

⁴⁸ The approach to deletion of valued features adopted here is similar to the one in Chomsky (2001:12). Features valued during a cycle remain active for further computation at least until the strong phase level. That is, even though the ϕ -features of AGR (the T of CG) are valued, they can remain active for further computation until the end of their strong phase, corresponding either to matrix v (as in the case of (51)) or to matrix T (as in the case of matrix passive and raising verbs ((53) and (54)).

⁴⁹ The restrictions on expletive constructions with CGs (see fns. 39 and 47) also extend to raising constructions with CGs, presumably for the same reasons associated with the impossibility of expletives in other

cases: Expletive *it* needs to have its Case valued (i), and once that happens in matrix [Spec, TP] of raising constructions, there is no other head that can value the Case feature of the embedded CG. In the case of *there* expletives, If the Case valuation mechanism underlying the *there*-associate DP interaction involves the need for an indefinite DP as the *there*-associate, (cf. (iii) below), this explains why the whole CG clause cannot occur in *there*-existentials (ii). That is, the CG itself does not have the semantic properties that indefinite DPs have which allow them to be licensed in *there*-existential constructions. This is captured straightforwardly in the current analysis, under which CGs are not even analyzed as DPs:

- (i) *It seems [a man swimming well].
- (ii) *There seems [a man swimming well].
- (iii) *There seems [Sue to be outside].

⁵⁰ However, my third reviewer suggests that “Whatever licenses DP-ing in absolute position (DP with NOM or ACC) is not sufficient to license POSS-ing.” I take the reviewer to use DP-ing to refer to absolute gerunds such as (59). I take this restriction to be an indication that the choice of morphological Case associated with the subject of the gerund is not random. The same can be observed with topicalized DPs, in a more restricted way:

- (i) *He/his, Mike never met. (cf. (60b).

See also fn. 43.

⁵¹ In the case of the different CGs analyzed in this paper (as it is the case with Control infinitivals, see e.g. Chomsky 1981, 1986, O’Neil 1995, Martin 1996, 2001, Hornstein 1999, 2001 among others), the different properties of obligatory control are taken to be partly dependent on the fact that the controller c-commands the obligatory control PRO. Given that absolute CGs display the standard properties of obligatory control, it is arguably the case that they are low enough in the structure at the point they are generated in order to be c-commanded by the subject of the matrix clause. In this respect they should behave similarly to adjunct gerunds that are introduced by an overt preposition, which were considered in section 3.2. (see Pires & Rodrigues 2002 for an analysis of left-peripheral non-finite adjuncts).

⁵² The fact that *John* in the higher matrix clause is marginally licensed as antecedent of the embedded null subject is related to the fact that the adjunct Absolute CG can attach either to the higher finite clause or to the intermediate one.

⁵³ My second reviewer points out that similar effects can be found in other non-finite clausal constituents.

She cites examples of participial small clauses (see also Belletti 1990 for data from Italian). I agree that the corresponding examples show a paradigm similar to gerunds as the complement of *while*, and not to the CGs I focus on here. For instance, participial small clauses introduced by *once* only allow a null subject (ia), but never an overt subject (ib), similar to *while* gerunds, but contrary to the CGs I analyzed in section 4.1 (see also fn. 54 for an attempt to explain this behavior in the case of *while* gerunds):

- (i) a. *Once* PRO written, the letter was carefully hidden.
- b. **Once* the letter written, it was carefully hidden.

⁵⁴ In addition, one can consider why a gerund introduced by *while* does not allow an overt subject, different from the CGs I analyzed here... My second reviewer suggests that *while* in the gerunds in (65) corresponds to a complementizer (see also Reuland 1983) (this would also distinguish them from the clausal gerunds I investigated, which I have argued do not project a CP). It is less clear whether these ‘*while*-gerunds’ indeed project a CP domain, given that there seems to be no direct supporting empirical evidence for this of the kind I pointed out for instance, for embedded infinitival clauses as in (22b), which allow A’-movement (overt *wh*-movement) to an embedded CP). However, if it is indeed the case that *while* in the gerunds in (65) corresponds to a complementizer, this can help explain why an overt subject cannot be licensed in the embedded gerunds, contrary to CGs. First, the embedded gerund in (65) does not carry the feature specification that allows CGs to check/value the Case of an embedded subjects under specific derivational conditions (as analyzed in detail in section 4.1), and *while* as a complementizer does not have Case checking/valuing properties that would be sufficient to license an embedded overt subject in the absence of a finite TP, as shown by the contrast in (i) (similar to what one observes complementizer *that*, as seen in (ii):

- (i) a. *Bill kept walking, while the rain drenching the road.
- b. Bill kept walking, while the rain was drenching the road.
- (ii) a. *Bill said that Mary calling.
- b. Bill said that Mary was calling.

⁵⁵ Left-peripheral non-finite adverbial clauses also exhibit obligatory control (ii below), which would not be problematic if one assumes reconstruction. However, my third reviewer cites Reuland & Avrutin (2005) and suggests that reconstruction is problematic for the finite counterparts of these adverbial clauses, since the presence of reconstruction would obviate the contrast between (ia) and (ib). First, I found speakers that do not accept *that*

student as a bound variable, and I exclude these speakers from (i). The judgments by the remaining informants still vary significantly, and the contrast in (i) is weaker regarding negative quantifiers:

- (i) a. Every/ ok/?/?/*no student_i turned in an assignment before that student_i went home.
 b. *Before that student_i went home, every/no student_i turned in an assignment

My reviewer goes on to connect the data to left-peripheral adjunct gerunds. Crucially for her, the adjunct gerunds in (ii) show obligatory control, but the controller does not c-command into the adverbial clauses in their surface, left-peripheral position in (iib to iid). Reconstruction of the left peripheral adverbial clauses does allow c-command of the OC PRO by its antecedent, but it would eliminate the grammaticality contrast between (iia) (with *no student*) and its left peripheral counterpart in (iic). However, notice first that speakers vary widely regarding the acceptability of (iic) with a negative quantifier, and there are speakers for which the contrast does not hold. In addition, no speaker notices any contrast between (iia) and (iib), and other quantifiers and referential controllers can bind a gap corresponding to OC PRO in left peripheral gerunds as well (iid). These facts indicate that the reconstruction necessary for c-command of the obligatory control PRO in left-peripheral adjunct gerunds is indeed possible, in different cases.

- (ii) a. Every/no student_i turned in an assignment before PRO_i going home.
 b. Before PRO_i going home, every student_i turned in an assignment.
 c. ok/?/?/* Before PRO_i going home, no student_i turned in an assignment
 d. Before PRO_i going home, Mary_i/ two students_i turned in an assignment.

Regarding the speakers who do not allow a negative quantifier to bind a variable corresponding to OC PRO in left peripheral adverbial clauses (iic), this may have to do not with the general properties of left peripheral adjunct gerunds (as shown by the grammatical (iib, d), but rather with a distinction in the scope properties of negative quantifiers, in contrast with other variable-binding quantifiers. As evidence of this contrast, we find scope ambiguity between *someone* and *every* (iii), but this scope ambiguity is not possible by having *no* either in the object position, instead of *every*, as in (iv), or in the subject position, instead of *someone*, as in (v). This contrast indicates that the quantifier *no* seems to have more limited (scope) possibilities in interaction with other quantifiers. I leave detailed investigation of the connection to variable binding above for future work.

- (iii) a. Someone danced with every woman.
 b. Someone called every student. (someone>every; every>someone)

(iv) a. Someone danced with no woman.

b. Someone called no student.

(someone>no; *no>someone)

iv) a. No one danced with every woman.

b. No one called every student.

(no>every; *every>no)