A Wholesale Late Merge theory of control

Suzana Fong*
sznfong@alum.mit.edu

August 15, 2023

Abstract

I propose a version of the Movement Theory of Control (Hornstein, 1999, et seq.) according to which a lone determiner can be merged in a thematic position in the embedded clause, move into an additional thematic position in the matrix clause, upon which point its NP complement is Wholesale Late Merged (Takahashi & Hulsey, 2009). Along with Takahashi (2019), I propose that Fox's (2002) Trace Conversion consists in transforming the residue of movement into a set of features that are formally indistinguishable from a pronoun. Likewise, along with Erlewine & Gould (2016), I propose that this interpretive procedure can apply at the Narrow Syntax, so that its result is visible for Vocabulary Insertion. The result is a theory of control that can account for why obligatory control PRO in a language like English has pronominal properties, a property that can be accounted for by a theory of control such as Chomsky's (1981) PRO-based GB theory, but not by the original version of the MTC. Furthermore, this paper investigates languages where PRO is not phonologically null, but, rather, an overt pronoun, pointing out hitherto unnoticed generalizations about the nature of the clause where pronounced PRO's occur, as well as about its status with respect to other pronouns that occur in languages that display an overt PRO. The theory proposed, coupled with a realizational framework of the grammar (Halle & Marantz, 1993, 1994) and with independently needed assumptions about phasehood, provides an account of such generalizations.

Keywords: obligatory control, Wholesale Late Merge, Movement Theory of Control, PRO, pronounced PRO, overt PRO, phase, minimal pronoun, Weak Phase Impenetrability Condition

1 Introduction

Obligatory control is a phenomenon whereby the subject of an embedded clause is bound by a matrix argument.

(1) a. Sindhu$_1$ tried [PRO$_{1/*2}$ to watch *Grave of the Fireflies*].
b. Faatu convinced Sindhu$_1$ [PRO$_{1/*2}$ to watch *Grave of the Fireflies*].

The matrix binder can be a subject (1a) or an object (1b) and is referred to as 'controller'. The embedded subject that is bound by it is usually phonologically null and is referred to as 'PRO'.

Two major theories can be distinguished in the account of obligatory control:

(2) a. PRO-based theories: PRO is a primitive of the grammar that imposes particular licensing requirements. In e.g. Chomsky (1981), the syntactic behavior of PRO and its interpretive properties are regulated by principles of Binding Theory.

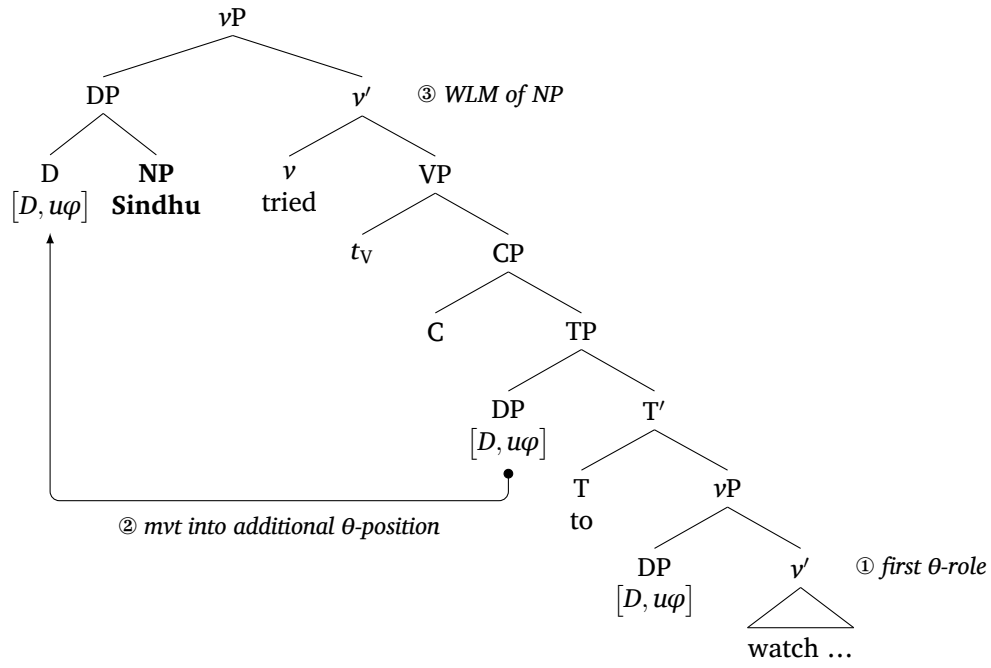
*Thank you to Christos Christopoulos, Lefteris Paparounas, and Anastasia Tsilia for patiently and generously sharing their Modern Greek judgments and for discussion and useful suggestions. Thank you also to Ka Fai Yip and Quihao Charlie Yan for their Mandarin judgments. An additional thank-you to Ka Fai for many insightful suggestions regarding the Mandarin data in section 2.2. I would like to thank the Syntax and Semantics communities at Yale, the participants of a LingLunch presentation at MIT, and the members of the Partial Deletion Reading Group (specially Ka Fai Yip, Qiuhao Charlie Yan, and Tommy Tsz-Ming Lee). Finally, thank you to Zhouyi Sun for reviving my interest on Visser's Generalization, discussed in the Appendix. All errors are my own responsibility.

b. Reductionist theories: PRO is not a primitive of the grammar; rather, it is the byproduct of other, more general principles. In Hornstein's (1999, et seq.) theory, obligatory control reduces to raising into more than one θ-position.

Each one of these theories has it own advantages and drawbacks. While Hornstein's (1999) reductionist theory is empirically supported by control phenomena such as backwards and copy control, it cannot capture the fact that, in a language like English, obligatory control PRO exhibits pronominal behavior. In contrast, this property can be straightforwardly accounted for by Chomsky's (1981) proposal that PRO is simultaneously an anaphor and a pronoun. A PRO-based theory, on the other hand, is hard-pressed to account for control phenomena where PRO is not phonologically null such as backwards control.

In this paper, I propose a modification of the Movement Theory of Control that combines the advantages of both PRO-based and of reductionist theories of control, while doing away with their respective drawbacks. Specifically, assuming that Wholesale Late Merge (Takahashi & Hulsey, 2009) is available in the grammar, obligatory control can be derived by base-generating a determiner alone in the subject position of the embedded control clause. It then moves into an additional θ-position in the matrix clause, where the NP portion of the controller is countercyclically merged into the structure.

(3)



I will propose, furthermore, a modification of Fox's (2002) Trace Conversion, so that the lone determiner that is at the tail of the movement chain resulting in control is converted into a set of features that is formally indistinguishable from a pronoun. In addition, because I propose that this procedure applies at the Narrow Syntax, such a derived pronoun is predicted to be available for exponence.

This is how I model instances of obligatory control PRO that are not phonologically null, as in (1), but rather an overt pronoun. An example from Wolof is previewed below.

(4) *Wolof* (Fong, To appear)

Dimbali-na-a a-b xale [*(**mu**) jàng téere b-i].
help-NA-1SG INDEF-CM.SG child 3SG.SUBJ read book CM.SG-DEF
'I helped a child read the book.'

This paper is structured as follows. Section 2 synthesizes the properties of obligatory control that must be accounted for by any theory. In the same section, we investigate instances of PRO that are not phonologically null, but rather an overt pronoun. Given the evidence available, I formulate generalizations about the overt PRO which, to the best my knowledge, have not been noticed before. More precisely, whenever

a language allows for an alternation between the usual phonologically null PRO and an overt pronominal PRO, the former occurs in a clause that is not a phase, while the latter occurs in a phasal complement. Furthermore, I remark that a pronominal overt PRO is not a sui generis category, but rather is identical to the pronouns that exist independently in a given language. Theories of overt PRO must thus account for how what appears to be the same form occurs as an obligatorily bound variable in some contexts (i.e. as an overt PRO), but as a referential or bound pronoun in other contexts. With the empirical foundation in place, section 3 turns to a summary of two classes of control theory, with a focus on how they account for the properties listed in section 2. In section 4, I lay out the main proposal put forward in this paper, which I claim combines the desirable features of the theories summarized in the previous section, while doing away with the challenges they face. In section 4.5, I provide an account of pronounced PRO's, along with the above-mentioned generalizations. Section 5 concludes.

2 Obligatory control

As previewed in the preceding section, obligatory control is a binding relationship between a matrix argument of verbs like *want*, *try*, *convince*, etc and the subject of the clause that these verbs subcategorize for. For an extensive overview of control phenomena and theories, see Landau (2013).

(5) a. Sindhu$_1$ tried [PRO$_{1/*2}$ to eat natto].
b. Lasha convinced Sindhu$_1$ [PRO$_{1/*2}$ to eat natto].

Such embedded clauses are usually nonfinite (though see finite control in Ferreira 2000, 2009; Lee 2003, a.o.). Their subject is usually phonologically null and is referred to as 'PRO', while the matrix argument that binds it is referred to as the 'controller' (i.e. in (5a) and (5b), the controller is *Sindhu*). PRO's phonological nullness in a language like English is further indicated by the ungrammaticality of replacing PRO with an overt nominal expression like a pronoun:

(6) a. * Sindhu tried [**she** to eat natto].
b. * Lasha convinced Sindhu [**she** to eat natto].

The fact that PRO is a bound variable is taken to be a signature property of obligatory control (Landau, 2013). A bound variable can be diagnosed by a series of properties, for instance, an obligatory *de se* interpretation (7), a sloppy reading under ellipsis (8), and an obligatory bound reading with an *only* antecedent (9). The sentence (7) is only consistent with a context where the matrix subject (*Maryam*) is aware of the identity of the matrix object controller (*Kadeer*).

(7) Maryam told Kadeer [PRO to leave].
a. Maryam is hosting a party. She hears that a certain waiter named Kadeer is being a nuisance. Maryam tells Kadeer "You have to go."
b. # Maryam is hosting a party. She hears that a certain waiter named Kadeer is being a nuisance. Maryam tells the nearest waiter "Kadeer has to go." Unbeknownst to her, she's talking to Kadeer.

While VP-ellipsis can give rise to a sloppy or strict reading, if an obligatory control sentence undergoes ellipsis, as in (8), only a sloppy reading becomes available.

(8) John expects [PRO to win] and Bill does △ too.
a. 'Bill expects that he, Bill, will win.'
b. # 'Bill expects that he, John, will win.'

Finally, (9) indicates that, when the controller is an *only* DP, a referential reading of the embedded subject is eschewed and only a reading where PRO is bound by the matrix *only* DP is possible.

(9) The teacher told only Lasha [PRO not to go to the party].

a. 'Lasha is the only *x* such that *x* the teacher told *x* for *x* not to come to the party.'

b. # 'Lasha is the only *x* such that them (i.e. some person other than Lasha) the teacher told *x* for *x* not to come to the party.'

A perhaps less discussed property of obligatory control PRO is its pronominal character. Postal (1998, 2004) observes that, in English, pronouns are disallowed in certain contexts. One such context is the subject position of expressions like *be wrong with* and *be the matter with*.

(10) (Postal 2004, adapted)

Lots of things$_1$ are wrong with my liver, but fortunately {***they**$_1$/those things} are not wrong with your liver.

Landau (2013) remarks that PRO cannot occur in the subject of the same type of predicate (11b), while the trace of A-movement can (11a). The paradigm in (11), thus, demonstrates both that PRO's behavior patterns with that of a pronoun (10) and also that it is unlike that of the residue of movement.

(11) (Landau, 2013, 16f)

a. * Lots of things can be the matter with your transmission [without PRO being the matter with mine].

b. Lots of things seem [*t* to be the matter with your transmission].

Likewise, Landau (2007, 321) observes that a similar argument is provided by the possible antecedence relationships between certain DPs and their metonyms. More precisely, pronominal antecedence tolerates certain metonymous shifts (12) but not others (13).

(12) (Landau, 2007, (69–70))

a. I am parked on 26th Street (= my car is parked on 26th Street).

b. John$_k$ claimed that he$_k$ was parked on 26th Street.

(13) (Landau, 2007, (69–70))

a. Microsoft went up (= Microsoft's stock's price went up).

b. * Microsoft$_k$ claimed that it$_k$ would go up.

The informal generalization from (12) and (13) is that a pronoun may be metonymous to its antecedent in the case of a car-possessor relation, but not in the case of a company-stock's price relation. Whatever the explanation for this generalization is, the contrast between the (a) vs. (b) examples in (12) and (13) is preserved in control (14), but not in raising (15).

(14) (Landau, 2007, (71) and (73), adapted)

a. John$_k$ plans [PRO$_k$ to be parked on 26th Street].

b. * Microsoft$_k$ plans [PRO$_k$ to go up].

(15) (Landau, 2007, (72), adapted)

a. John$_k$ seems/is likely [t_k to be parked on 26th Street].

b. Microsoft$_k$ seems/is [t_k likely to go up].

This asymmetry indicates again that obligatory control PRO is pronominal and that it contrasts with the trace of a canonical instance of A-movement like raising.

In sum, we briefly surveyed the following properties of obligatory control:

(16) a. PRO is a bound variable.

b. PRO is usually phonologically null.

c. PRO is pronominal.

I would like to add two desiderata to the roster of properties that control theories should be able to account for. In section 2.1, we briefly backwards control and in section 2.2, pronounced PRO's.

2.1 Backwards control

First documented in Tsez and analyzed by Polinsky & Potsdam (2002), backwards control is a phenomenon whereby the controller is pronounced not in the matrix clause, but in the embedded clause, where a phonologically null PRO is otherwise expected. In (17), we see examples from Modern Greek (cf. Alexiadou *et al.* 2010; Alexiadou & Anagnostopoulou 2019). Backwards control is also found in e.g. Ndebele (Pietraszko, 2021).

(17) *Modern Greek* (C. Christopoulos, p.c., modeled after an example from Alexiadou *et al.* 2010)[1]

a. I kathighites iposkhethikame o enas stus mathites tu
promised.1PL the.NOM one.NOM to-the.PL.ACC students.ACC 3SG.GEN
allu [na tus dhosume **i** **kathighites** kalus vathmus].
other.GEN [SBJV 3PL.GEN give.1PL the.NOM professors.NOM good.ACC grades.ACC]
'We professors$_1$ promised each other$_1$'s students to give them good grades.'[2]

b. káthe kathigitís iposkhethike stus mathites tu [na tus dhosi
promised.3SG to-the.PL.ACC student.PL.ACC 3SG.GEN [SBJV 3PL.GEN give.3SG
kathe kathighitis kalus vathmus].
every professor.NOM good.ACC grades.ACC]
'Every professor$_1$ promised their$_{1/2}$ students to give them good grades.'

Even though the controller *kathighites* 'the professors' or *kathe kathighitis* 'every professor' is realized in the embedded clause, that there is a representation of the this DP in the matrix clause can be demonstrated by the fact that it can bind a matrix element such as a reciprocal (17a) or a variable (17b).[3]

It can be demonstrated that backwards control is indeed an instance of obligatory control (cf. Polinsky & Potsdam 2002; Alexiadou *et al.* 2010; Pietraszko 2021). The challenge it poses to theories of control is that the position of the controller and of the phonologically null PRO appear to be switched: the controller is realized in the embedded clause, instead of occupying the expected matrix position, while the opposite holds of PRO.

2.2 Pronominal overt PRO

In this section, we will examine a particular instance of the realization of obligatory control, namely, control clauses where PRO is not phonologically null, but rather an overt pronoun. We will examine two types of languages with pronounced PRO's, namely those where all instances of obligatory control PRO is an overt pronoun and those where an overt PRO alternates with its phonologically null counterpart. Additionally, I lay out empirical generalizations about the distribution of pronounced PRO's that must be accounted for in an analysis of this variety of control realization.

To the best of my knowledge, in the overwhelming majority of languages, obligatory control PRO does not have any phonological matrix. It is no surprise that most existing theories of control seek to provide an account of PRO's phonological nullness. This is the case of the two classes of control theories to be examined in section 3. However, while typologically rare, cases of overt obligatory control PRO have been attested in a few languages. San Lucas Quiaviní Zapotec is well-known for having copy control (Lee, 2003), where the PRO position is occupied by an identical copy of the controller:

(18) *Copy control in San Lucas Quiaviní Zapotec* (Lee, 2003, (62), adapted)

[1] In this paper, uncited Modern Greek data represents the judgments offered by C. Christopoulos. Virtually identical data was also evaluated by A. Tsilia and L. Paparounas, though with slight variations in the degree of grammaticality.

[2] In (17a), the main verb is inflected in the 1st person plural to avoid a parsing where the matrix subject is a dropped 3rd person plural pronoun (i.e. the meaning of (17a) in this parsing would be 'They$_1$ promised each other$_1$'s students that the professors$_2$ would give them good grades'). This parsing is not relevant for a discussion of backwards control. Thank you to C. Christopoulos and L. Paparounas for this insight.

[3] Binding data like (17) first appear in Polinsky & Potsdam's (2002) analysis of Tsez. However, because additional Modern Greek data was more readily available, I opt to represent data from this language. See also section 4.4.

R-cààa'z **Lia Paamm** [g-ahcnèe **Lia Paamm** Gye'eihlly].
HAB-want FEM Pam IRR-help FEM Pam Mike
'Pam wants to help Mike.'

Telugu also displays the same phenomenon (Haddad, 2009). Copy control is similar to backwards control, examined in the previous section, except that, in the latter, the controller is realized only once, in the embedded clause.

In another class of overt obligatory control PRO phenomenon, PRO is not identical to the controller, but is realized as an overt matching pronoun instead.[4] Additionally, as mentioned above, two languages can be distinguished which display overt pronouns where PRO is otherwise expected. In San Martín Peras Mixtec (Ostrove, To Appear), illustrated in (19), and in Gã (Allotey, 2021), it seems that all instances of obligatory control PRO are an overt pronoun.

(19) *San Martín Peras Mixtec* (Ostrove, To Appear, (1a))
Ntùkú Juân$_i$ [ka'ani *(=**rà**$_{i/*j}$) iin ntsìbá'yi].
try:COMP Juan kill:IRR =he one coyote
'Juan tried to kill a coyote.'

Overt PRO languages may also display an alternation between a pronominal PRO and its phonological null counterpart. In Wolof, the alternation occurs across different clauses, selected by difference predicates. In (20a), the controlled subject is obligatorily an overt pronoun, while in (20b), it is obligatorily null.

(20) *Wolof* (Fong, To appear)
a. Dimbali-na-a a-b xale [*(**mu**) jàng téere b-i].
help-NA-1SG INDEF-CM.SG child 3SG.SUBJ read book CM.SG-DEF
'I helped a child read the book.'
b. Xadi jéem-na [(***mu**) togg ginaar].
Xadi try-NA.3SG 3SG.SUBJ cook chicken
'Xadi tried to cook chicken.'

Mandarin (Li, 2021) has a similar pattern, though one type of clause allows for the aforementioned alternation (21a), while the other may require a null subject (21b).

(21) *Mandarin*
a. (Li, 2021, (27), brackets added)
Lisi dasuan [mingnian xiatian (**ta**) qu lüyou].
Lisi plan next.year summer he go travel
'Lisi plans to go travel next summer.'
b. (K.F. Yip, p.c.; Q. Yan, p.c.)
Lisi jiao / bi Zhangsan$_1$ [(%**ta**$_1$) qu shang daxue].
Lisi ask force Zhangsan he go attend college
'Lisi asks/forces Zhangsan to go to college.'

In the alternating languages, the alternation is in fact illusory. The empirical generalization across the different languages that allow for this alternation is that the overt pronominal PRO occurs in a clause that is structurally more complex than the clause where the phonologically null PRO occurs. Clause size can be diagnosed with language-specific properties. For instance, Fong (To appear) shows that in control clauses with an overt pronominal PRO, clitic climbing is prohibited (22a), while it is obligatory in control clauses with a null subject (22b).

[4] That the pronominal pronounced PRO is indeed an obligatorily controlled subject, despite its phonological overtness, is supported by straightforward diagnostics that identify bound variables, such as the ones summarized in (7–9) above. Cf. Livitz (2014, §4.3, European Portuguese, Hungarian, and Italian), Allotey (2021, §3.3.4–§3.3.7, Gã), Li (2021, §3, Mandarin), Sulemana (2021, 4.3, Bùlì), Ganenkov (To Appear, §4, Chirag), Fong (To appear, §3, Wolof), Ostrove (To Appear, §4, San Martín Peras Mixtec).

(22) *Wolof* (Fong, To appear, (38) and (39))

a. Kadeer dimbali-na { *=**ko**} Mareem [mu jënd { =**ko**}].
Kadeer help-NA.3SG =3SG.ACC Mareem 3SG.SUBJ buy =3SG.ACC
'Kadeer helped Mareem buy it.'

b. Maymuna fas-na { =**ko**} [jàng {*=**ko**}].
Maymuna want-NA.3SG =3SG.ACC read =3SG.ACC
'Maymuna wants to read it.'

Clitic climbing is taken to be indicative of restructuring (Wurmbrand, 1998, et seq.). Given the contrast between (22a) and (22b), this means that control clauses in Wolof with an overt PRO are structurally more complex than control clauses with a null subject. Fong (To appear, §4.2) extends this analysis to a further contrast between the clauses in (20) that is based on the occurrence of resumptive pronouns at the tail of *Wh*-movement departing from them.

A similar distinction can be seen in Bùlì (Sulemana, 2023). In this language, the clausal complement of verbs like *sìak* 'agree' can have a null (23a) or overt pronominal subject (23b).

(23) *Bùlì* (Sulemana, 2023, (2c/5b), adapted, with brackets added)

a. Asouk sìak [dà gbăŋ].
Asouk agree buy book
'Asouk agreed to buy book.'

b. Asouk$_i$ sìak [**wà**$_{i/*j}$ dā gbăŋ].
Asouk agree 3SG buy.INF book
'Asouk agreed to buy book.'

The difference in the tone of the embedded verb already suggests that the embedded clauses are not identical. Further indication of the difference is provided by the fact that a clause without an overt PRO cannot license a temporal adverb that is independent from the matrix (24a), though its counterpart with an overt PRO (24b) can.

(24) *Bùlì* (Sulemana, 2023, (15a/16a), adapted, with brackets added)

a. * Dìemwá, Asouk sìak [dà gbăŋ chum].
yesterday Asouk agree buy book tomorrow
Intended: 'Asouk agreed to buy a book tomorrow.'

b. Dìemwá, Asouk$_i$ sìak [**wà**$_i$ dā gbăŋ chum].
yesterday Asouk agree 3SG buy.INF book tomorrow
'Asouk agreed to buy a book tomorrow.'

By the same token, a null PRO clause cannot license negation, while an overt PRO clause can (cf. Sulemana 2023, (17/18)). Sulemana (2023) also analyzes this distinction in terms of clause size. Specifically, control clauses with a null subject are what the author calls 'reduced clauses', which project a subject-less VP that merges directly with the matrix verb's V (cf. restructuring in Wurmbrand 1998). According to this analysis, the fact that the subject of reduced clauses is null is a trivial consequence of there not being a subject to begin with. By the same reasoning, this means that clauses with an overt PRO must be complex enough to host a subject position position.

This contrast in clause size can be generalized in terms of phasehood. More precisely, clauses with a null PRO are smaller and, hence, not phasal, while clauses with an overt pronominal PRO are structurally more complex and, hence, phasal. A distinction anchored on phases, with clause size being a derivative property, has the advantage of being generalizable to a language like Mandarin. Li (2021) argues that both embedded clauses in (21) are CPs. For one, both can host the ex-situ focus phrase *shenme-shi* 'everything', which is taken to be located at CP (Li, 2021, 312)

(25) *Mandarin* (K.F. Yip, p.c.; Q. Yan, p.c., modeled after Li 2021, (32))

a. Lisi jiao / bi Zhangsan [sheme-shi dou ziji jiejue *t*].
Lisi ask force Zhangsan what-matter all self handle
'Lisi asks/forces Zhangsan to handle everything by himself.'

b. Lisi dasuan [sheme-shi dou **ta** ziji jiejue *t*].
Lisi plan what-matter all he self handle
'Lisi plans to handle everything by himself."

Nonetheless, these clauses still differ in the dependencies that can be established across them. According to Huang (2022), the lowering of the aspectual morpheme *guo* is possible across a nonfinite clausal boundary, but prohibited across a finite clause. *Guo*-lowering consists in the possibility of the aspect denoted by this morpheme to be interpreted in the matrix clause and not in the embedded clause where it is realized. This difference can be restated in phasehood terms: *guo*-lowering is not possible across a phase. Against this background, consider the behavior of the lowered *guo* in control clauses: in clauses with a null PRO, *guo*-lowering is possible (26a), but not in clauses with an overt PRO (26b). According to the assumptions made here, the former is not a phase, while the latter is.

(26) *Mandarin* (K.F. Yip, p.c.; Q. Yan, p.c.)

a. Lisi cengjing jiao / bi Zhangsan [$_{CP}$ zuo-guo zhe-dao can].
Lisi previously ask force Zhangsan make-EXP this-CL dish
'Lisi has asked/forced Zhangsan before to cook this dish.'

b. */?? Lisi cengjing dasuan [$_{CP}$ (**ta**) zuo-guo zhe-dao can].
Lisi previously plan [he make-EXP this-CL dish]
'Lisi has planned before to cook this dish.'

Likewise, long distance passivization could also be used to diagnose phasehood: control clauses with a null PRO are not phases, so that passivization across them yields a grammatical sentence (27a), while control clauses with an overt PRO are phases, so that passivization yields an ungrammatical sentence (27b).

(27) *Mandarin* (K.F. Yip, p.c.; Q. Yan, p.c.)

a. Neifeng xin$_1$ bei Lisi jiao / bi Zhangsan [$_{CP}$ shao-diao le t_1].
that.CL letter BEI Lisi ask force Zhangsan burn PRF
'That letter was asked/forced for Zhangsan to burn by Lisi.'

b. * Neifeng xin$_1$ bei Lisi dasuan [$_{CP}$ **ta** shao-diao le t_1].
that.CL letter BEI Lisi plan he burn PRF
Intended: 'That letter was planned to be burnt by Lisi.'

As Li remarks, taken together, these data indicate that, in Mandarin, even though control clauses may have the same grammatical category (viz. CP), their behavior differs. I argue that this difference can be productively stated in terms of phasehood, which allows for a unified view of the distinction between clauses with an overt PRO and clauses with a null PRO, across the typologically distinct languages that allow for this alternation. Phasehood will be incorporated in the analysis to be put forward in section 4.

Besides this this correlation between clause size and the realization of PRO, another empirical generalization regarding pronounced PRO's is that, to the best of my knowledge, they are never a sui generis category. Rather, the overt PRO is a pronoun that occurs independently in the language. The same generalization is well known in the realm of resumptive pronouns (McCloskey, 2017). A successful account of overt pronominal PRO's must thus account for the fact that a pronominal form available in the language can have a bound or deictic reading in some environments, while being obligatorily locally bound in a control environment. This is illustrated in (28) with Mandarin data.

(28) *Mandarin* (K.F. Yip, p.c.)

a. Mei-ge nühai$_1$ dou shuo **ta**$_{1/2}$ xiang zuo taikongren.
every-CL girl DOU say 3SG want do astronaut
'Every girl said ta wants to be an astronaut.'

b. Lisi$_1$ dasuan [**ta**$_{1/*2}$ qu kan-kan].
Lisi plan 3SG to see-see
'Lisi plans to go take a look.'

All in all, a theory of control that accounts for pronounced PRO's must provide an explanation for the following generalizations:

(29) a. When an alternation is available between a phonologically null PRO and a pronominal overt PRO, the former tends to occurs in a clause that is structurally less complex than the clause where the overt PRO tends to occur. This correlation can be profitably restated in terms of phasehood.
b. A pronominal overt PRO is never a separate category in the languages that allow for them. Rather, it is identical to pronouns that occur independently in these languages, albeit one that is necessarily locally bound.

To the best of my knowledge, these generalizations are novel in the scarce literature on pronounced PRO's. In section 4.5, we will discuss to what extent the logic of existing analyses are consistent with (29b). I will offer a speculation as to why (29a) may hold, given the control theory to be put forth in this paper.

2.3 Interim summary

In this section, we examined a few properties that characterize obligatory control in general, as well as two particular instances of control phenomena, namely, backwards control and instances of control where PRO is a pronounced pronoun.

In the next section, I will summarize and discuss two types of control theory and how they fare with respect to the desiderata in (16), updated in (30).

(30) a. PRO is a bound variable.
b. PRO is usually phonologically null.
c. PRO is pronominal.
d. PRO can be controlled backwards.
e. PRO can be realized as an overt pronoun.

We will see that the types of theories examined have their own advantages and shortcomings regarding these desiderata. The control theory to be proposed in this paper combines the virtues of these theories, while avoiding the challenges they face.

3 Two classes of control theory

3.1 PRO-based theories of control

As an exemplar of a control theory that posits a lexical item PRO, I will summarize Chomsky (1981). Chomsky's starting point is the observation that PRO's distribution and interpretive properties share similarities with both pronouns and reflexives. For instance, both reflexives and PRO are variables without an inherent reference. On the other hand, PRO resembles a pronoun in that it can only be bound by an antecedent that belongs to a different clause. Chomsky hypothesizes, thus, that PRO, being both anaphoric and pronominal, must simultaneously obey Conditions A and B.

(31) *Binding Theory* (Chomsky, 1981, 188)
(A) An anaphor is bound in its governing category.
(B) A pronominal is free in its governing category.
(C) An R-expression is free.

Given the definitions in (31), this entails that PRO must at once be bound and free in its governing category.

In order to solve this contradiction, Chomsky proposes that PRO simply lacks a governing category. This has come to be known as 'PRO theorem', since it follows from the principles that define Binding Theory (31).

(32) *PRO theorem* (Chomsky, 1981, 191)

PRO is ungoverned.

As such, Binding Theory accounts for the syntactic distribution and semantic interpretation of control PRO, thus providing an account of the properties (30a) — PRO is a bound variable — and (30c) — PRO is pronominal.

Chomsky (1981) also has an account of its phonological nullness (30b), though in this case, the explanation rests on theory-internal assumptions. Case, as other syntactic relations in Chomsky (1981), must be assigned under government. However, we just concluded that the absence of a governing category for PRO is essential to account for its distributive and semantic properties. The lack of a governing category would then mean that PRO cannot be assigned case. Chomsky relativizes the Case Filter to nominal categories that have phonological matrix:

(33) *Case Filter* (Chomsky, 1981, 49)

*NP if NP has phonetic content and has no Case.

Traces and PRO are exempt from the Case Filter. PRO's phonological nullness is, thus, critical for the account of PRO's Binding Theoretic properties to be reconciled with the Case Filter — if PRO had to be assigned case, this would not be possible, since it is ungoverned.

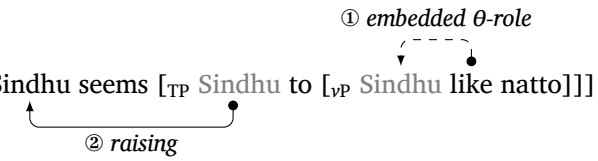
While a PRO-based theory like Chomsky (1981) can provide an account for the properties (30a–30c), the same cannot be said of (30d–30e). Both properties require that PRO be a candidate for phonological exponence, which is incompatible with the need for PRO to be phonologically null in order for the Case Filter to be evaded. The pronominal overt PRO and the overt controller in backwards control are semantically and syntactically identical to their null PRO counterpart. In Chomsky's theory, this would mean that they must be ungoverned. However, the fact that they have phonetic content would subject them to the Case Filter. Because, as mentioned, case assignment requires governing, this principle could not be complied with. As a result, any type of pronounced PRO is predicted to be impossible in Chomsky (1981).

3.2 Reductionist theories of control

Hornstein (1999, et seq.) proposes that PRO is not a sui generis category, nor is there a dedicated module in the grammar that states the conditions that must be met in order for PRO to be licensed. Rather, obligatory control reduces to raising of the controller which passes through more than one thematic position.[5] This proposal renders the derivation of obligatory control very similar to that of raising. The difference between them lies in whether or not the moving DP has passed through more than one thematic position.

(34) *Raising: movement of a DP through one θ-position*

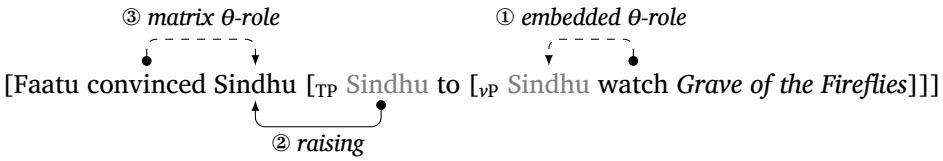
Sindhu$_1$ seems [t_1 to like natto].



(35) *Control: movement of a DP through more than one θ-position*

Faatu covinced Sindhu$_1$ [PRO$_{1/*2}$ to watch *Grave of the Fireflies*].

[5] The Movement Theory of Control has faced extensive criticism. See. among many others, Culicover & Jackendoff (2001); Landau (2003, 2007); Modesto (2007, 2011); Bobaljik & Landau (2009); Ndayiragije (2012); Wood (2012, 2017).



In the MTC, several properties that characterize PRO are a byproduct of the fact that its the trace of A-movement. The fact that PRO in obligatory control is a bound variable reduces to its movement properties, assuming that DP movement creates a λ-abstraction configuration (Kratzer & Heim, 1998), which accounts for property (30a). PRO is, therefore, not a primitive of the grammar under the Movement Theory of Control. Rather, in the embedded subject position of a sentence like (35), we have the residue of A-movement, just like in the raising sentence (34). The reason why PRO is usually phonologically null (30b), thus, reduces to whatever reason the residue of movement is usually phonologically null (e.g. linearization constraints, cf. Nunes 2004, a.o.).

However, the very fact that PRO is a trace of A-movement cannot account for the pronominal nature found in PRO (30c). In other words, the enforcement of anti-pronominality effects in control clauses is not expected from a theory where control is derived by movement. Importantly, the fact that raising and control differ in this regard (i.e. in raising, unlike what happens in control, anti-pronominality effects are not enforced, see the data in section 2) is unexpected in a theory where both control and raising are derived by an instance of movement that leaves a non-pronominal trace behind.

Nonetheless, a major empirical argument in favor of the Movement Theory of Control is backwards control (30d). In fact, Polinsky & Potsdam (2002) argue that backwards control provides unequivocal empirical support for the Movement Theory of Control, if it is coupled with the independently required Copy Theory of Movement (Chomsky, 1995), which states that movement consists in copying and remerging a constituent, with the lower copy of a movement chain usually being unpronounced. This is, however, not always the case. In covert *Wh*-movement, for instance, it is the lower copy of movement that is pronounced. This is illustrated (36b), which showcases data from Mongolian, a language with covert *Wh*-movement.

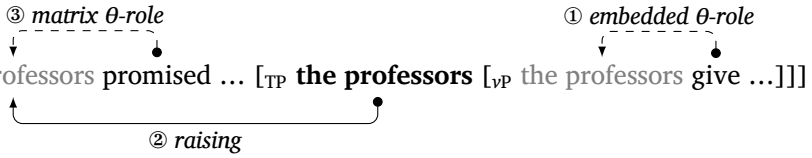
(36) *Covert* Wh*-movement in Mongolian*

a. Bat ene nom-iig unsh-san.
Bat this book-ACC read-PST
'Bat read this book.'

b. yu Bat **yu** id-sen be?
Bat what eat-PST INTERR
'What did Bat eat?'

Under the Movement Theory of Control, combined with the Copy Theory of Movement, a backwards control sentence like (17a), repeated below, is derived as follows:

(37) I kathighites iposkhethikame o enas stus mathites tu allu [
promised.1PL the.NOM one.NOM to-the.PL.ACC students.ACC 3SG.GEN other.GEN [
na tus dhosume **i** **kathighites** kalus vathmus].
SBJV 3PL.GEN give.1PL the.NOM professors.NOM good.ACC grades.ACC]
'We professors$_1$ promised each other$_1$'s students to give them good grades.'



The controller *i kathighites* 'the professors' is base-generated in the embedded clause, where it receives one θ-role. It then moves into the matrix clause and receives an additional θ-role there, just as in (35). However, unlike what happens in the latter, it is a lower copy of movement that is pronounced, a possibility that is independently attested (cf. (36b)).

This leaves us with property (30e). In its original formulation, the MTC predicts that there is a full representation of the controller in the embedded clause — this is the position it occupies before moving

into an additional θ-position in the matrix clause. There are, then, two possible realizations: either PRO is a full copy of the controller or it is a deleted chain of movement. As mentioned, these options readily account for backwards and copy control, as well as a phonologically null PRO. A third option is provided by Sulemana (2018) and Fong (To appear), who propose that a pronominal overt PRO is the result of partially exponing the residue of DP movement (Van Urk, 2018). We can see, then, that the MTC, when coupled with additional assumptions can also be made to be compatible with pronominal overt instances of PRO. In section 4.2, I will propose a fourth avenue, under the version of the MTC that results from combining it with Wholesale Late Merge.

3.3 Interim summary

In this section, we examined two competing theories of control, Chomsky's (1981) PRO-based theory and Hornstein's (1999) Movement Theory of Control. We focused on the account that these proposals provide for the obligatory control properties listed in (30). The table in (38) summarizes and compares these results. For convenience, (38) also previews the accomplishments of the control theory to be advanced in the next section.

(38)

		PRO-based theory	Movement-based theory	WLMTC
a.	PRO is a bound variable	✓	✓	✓
b.	PRO is usually phonologically null	✓	✓	✓
c.	PRO is pronominal	✓	*	✓
d.	PRO can be controlled backwards	*	✓	✓
e.	PRO can be an overt pronoun	*	✓	✓

We can see that neither of these theories is successful in accounting for the totality of desiderata in (30). The quality of the differences (and not just their amount) should be taken into consideration as well. While Chomsky (1981) offers an account of PRO's phonological nullness, this property is in fact a stipulation put in place for theory-internal coherence. Furthermore, little wiggle room is available in the account of various instances of pronounced PRO's. In the MTC, on the other hand, what is stipulated is that θ-role assignment can be a trigger of syntactic operations like movement. PRO's phonological nullness follows independently from principles that regulate movement chains.

Because of its advantages, I will then keep the core of proposal put forth by the MTC, but combine it with independently supported mechanisms in order to account for the control properties synthesized in (30). The result is a control theory that inherits the positive results afforded by the MTC, while also accounting for why PRO is pronominal, a result that, thus far, only Chomsky's (1981) can boast.

As will be elaborated on below, while the theory to be proposed in section 4.2 inherits from the MTC its ability to account for backwwards control, an advantage that this proposal has over the MTC is its ability to predict when the residue of movement into an additional θ-position is pronominal and when it is a full copy. The crucial ingredient, as I will show below, is independently detectable phasehood properties. Phases will also be instrumental in accounting for the correlation between the size of a control clause and the phonological properties of PRO in languages that allow for an alternation between a null and an overt PRO (29a).

4 A Wholesale Late Merge Theory of Control

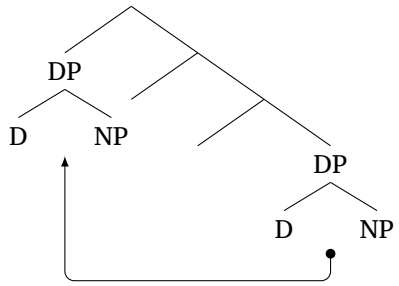
In this section, I will lay out the main proposal to be advocated for in this paper, namely, the idea that, in the derivation of a control sentence, a lone determiner can be merged in the embedded PRO position before it moves into the matrix clause, upon which point the NP portion of the controller is Wholesale Late Merged. Before that, section 4.1 offers an overview of WLM. In 4.2 is the main proposal. In the same section, I spell out how this proposal can account for the control properties listed in (30). In sections 4.4 and 4.5, we discuss the consequences of a theory that couples WLM with the MTC. Specifically, we will examine how it teases apart backwards control from cases where PRO is realized as an overt pronoun. This

is an important result, since, as mentioned above, the original version of the MTC considers that only a full representation of the controller in the embedded clause is possible.

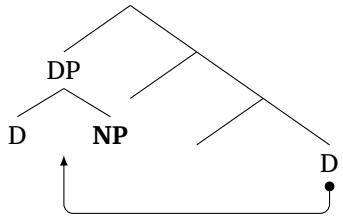
4.1 Wholesale Late Merge

Wholesale Late Merge (WLM, Takahashi & Hulsey 2009; Stanton 2016; Gong 2022) consists in the countercyclic merge of the NP complement of a DP, with D merging first in a lower position. This is diagrammed in (40), which can be contrasted with the standard, non-countercyclic derivation in (39).

(39) *Early merge of* [*determiner* + *NP*]



(40) *Wholesale Late Merge of NP*



The main empirical motivation behind WLM is an asymmetry in Condition C reconstruction. It is well-known that A-movement and Ā-movement are distinct in the possibility of reconstruction (see e.g. Safir 2019 and references therein). More precisely, A-movement does not have to reconstruct for Condition C (41a), while Ā-movement must do so (41b).

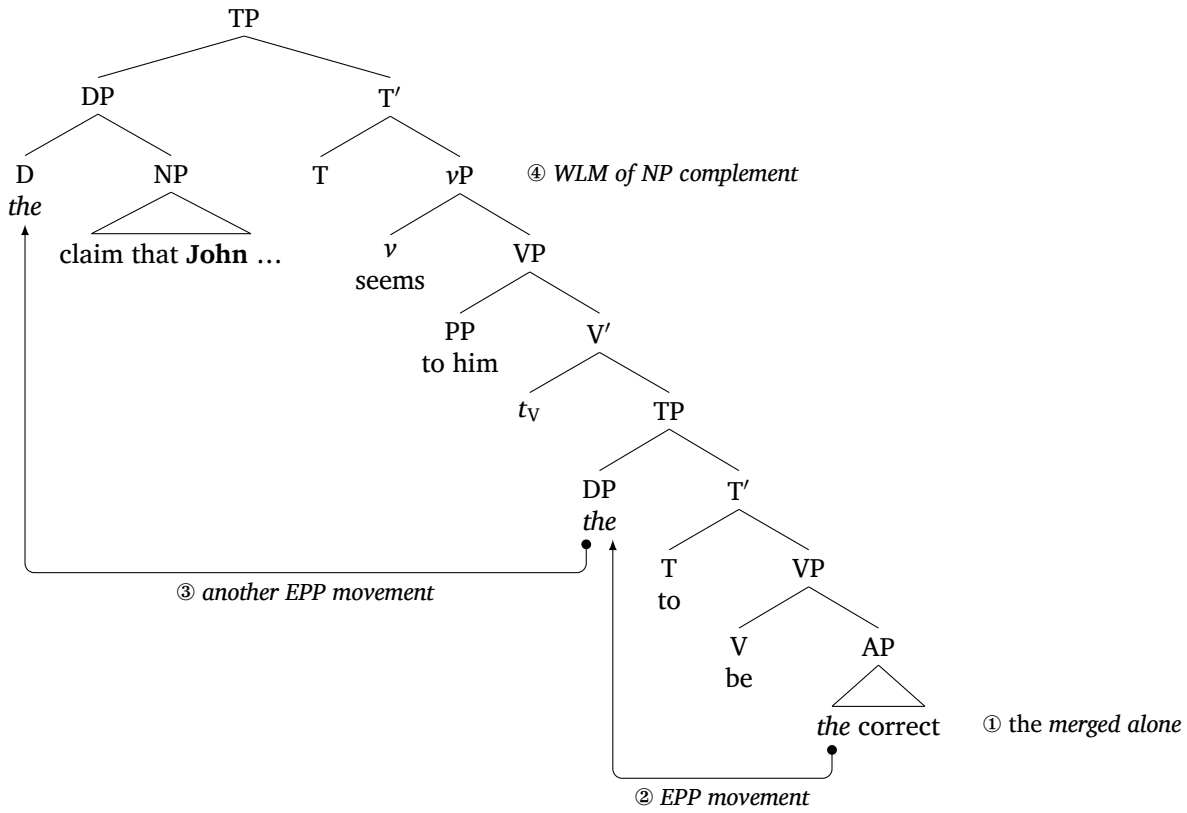
(41) a. The claim that John$_k$ was asleep seems to him$_k$ *t* to be correct.
b. * Which argument that John$_k$ is a genius did he$_k$ believe *t*?

This asymmetry poses a challenge to the Copy Theory of Movement (Chomsky, 1995). The lack of reconstruction for Condition C in A-movement suggests that it does not leave behind a contentful trace. Conversely, the obligatory reconstruction for Condition C in Ā-movement suggests that movement must leave behind a contentful copy. It is, of course, undesirable for two instances of movement to differ regarding whether they leave behind a contentful copy or not.

In view of this problem, Takahashi & Hulsey (2009) propose that a determiner can be merged alone in the base position. After it moves to a higher position, its NP complement can then merge with it. This is known as *Wholesale Late Merge*.

With WLM available in the grammar, the derivation of a sentence like (41a) would be as follows:

(42)



In this derivation, *the* is merged alone in the base-position. It then raises to the matrix Spec-TP. Subsequenctly, the NP complement *claim that John was asleep* is WLM-ed in this higher position. No Condition C violation is induced because *John* (and the rest of the NP complement) does not occupy any position that is c-commanded by *he*.

Takahashi & Hulsey conclude that movement always leaves a contentful copy behind, but WLM allows for reconstruction effects not to obtain by virtue of the NP complement that contains the would-be offending R-expression is countercyclically merged into the structure, so that it never occupies a position where Condition C is violated; that lower position is occupied by the determiner that was merged alone.

However, the obligatory reconstruction for Condition C in Ā-movement (41b) must still be accounted for. Takahashi & Hulsey propose that WLM is governed by independent restrictions. First, any derivation has to satisfy the Case Filter, a restriction that is also supported by Gong (2022). Second, the result of a derivation involving WLM (and, indeed, of any derivation) must be semantically interpretable. Takahashi & Hulsey follow Fox (2002) in assuming that movement chains created by Copy are interpreted via the LF procedure Trace Conversion.

(43) *Trace Conversion* (Fox, 2002)

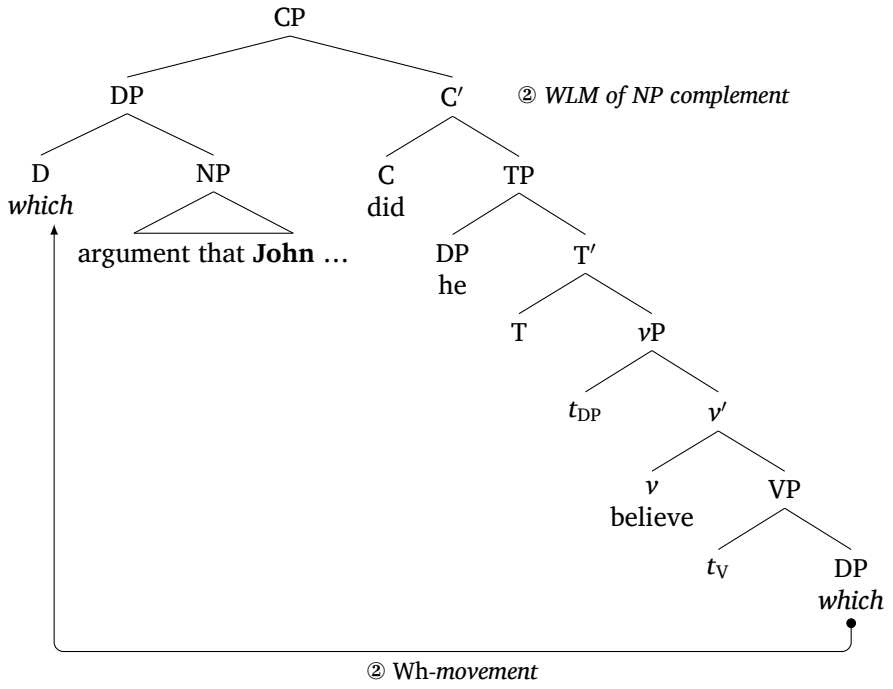
a. Variable Insertion: (Det) Pred → (Det) [Pred λy(y=x)]

b. Determiner Replacement: (Det) [Pred λy(y=x)] → the [Pred λy(y=x)]

Variable Insertion (43a) introduces an *et* predicate in the lower position. A binding dependency is established between the lower copy of the lone determiner and the higher position it moves into. Determiner Replacement (43b) converts that lower copy into a definite description.

While WLM creates interpretable representations in both A- and Ā-movement (see details in Takahashi & Hulsey 2009), Ā-movement creates a representation where case cannot be assigned. Ā-movement usually targets Spec-CP, a position that is outside the c-command domain of any case assigner. The derivation of a sentence like (41b), employing WLM, would go as follows:

(44)



The derivation in (44) is identical to that in (42), except for the position where the full DP sits in (44), which is not a position that can be reached by any case assigner. This causes the derivation to crash due to a violation of the Case Filter. The instance of A-movement depicted in (42) targets a case position. Indeed, the matrix T is able to assign nominative case to the full DP.

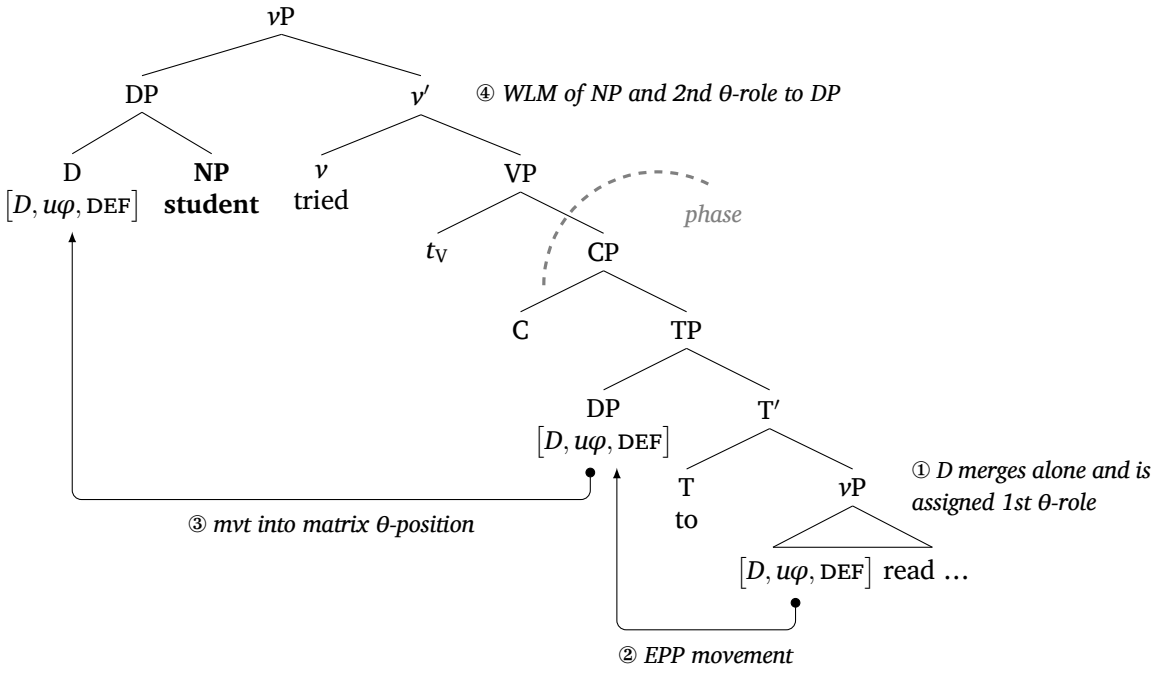
Takahashi & Hulsey emphasize that WLM is a free option in the grammar. Whether or not it applies, the result of the derivation is governed by independent principles of the grammar, such as the LF-interpretability and Case Filter requirements mentioned above. The application of WLM in the derivation of the A-movement sentence in (42) results in a representation that meets both of these requirements. The result is a sentence where Condition C can be avoided. If WLM had not been applied, the derivation would have been semantically interpretable and it would have complied with the Case Filter. However, a Condition C violation would have been incurred, since, in this case, a full representation of the DP containing the R-expression *John* would be c-commanded by the coindexed pronoun *him*. (44) represents a derivation of (41b) where WLM applies and the result crashes due to a Case Filter violation. In the alternative derivation where WLM does not apply, Case is correctly assigned to the early merged full DP, but the derivation crashes due to a Condition C violation. Hence, while there is one convergent derivation for (41a) that does not incur in either Case or binding violations, the same cannot be said of (41b).

4.2 Proposal: WLM meets the MTC

With background on WLM in place, we can now turn to the proposal made in this paper. Recall from section 3.2 that the MTC has not only quantitative advantages over a PRO-based theory of control, but also qualitative differences. As such, I will maintain the core of the MTC, but combine it with independently supported mechanisms, in order to expand its empirical reach.

I propose that obligatory control is the result of early merging a determiner in the embedded clause and then moving it into an additional θ-position in the matrix clause, at which point the NP complement is WLM-ed. A step-by-step derivation is represented in (45). It includes movement of a subject for EPP reasons, though nothing hinges on this part of the derivation.

(45) The student tried [PRO to read the book].



Importantly, the movement of the lone determiner in (45) crosses a phase boundary. The role that phasehood plays in the analysis will be spelled-out momentarily. The derivation continues with the building of the matrix TP, case assignment to the subject (the controller), and EPP movement. For concreteness, I assume that case assignment is a reflex of Agree (Chomsky, 2000, 2001).

To recall, the result of any syntactic derivation, including those where WLM has applied, must be interpretable at LF. I follow Takahashi & Hulsey (2009) in assuming Fox's (2002) Trace Conversion, though I propose the following modifications:

(46) *Early Trace Conversion* (modified from Fox 2002)

a. Variable Insertion: (Det) Pred → (Det) [Pred λy(y=x)]

b. Determiner Replacement: (Det) [Pred λy(y=x)] → [*D*, φ] [Pred λy(y=x)]

c. Trace Conversion takes place at the Narrow Syntax, at the phase level, assuming that operations at a given phase are only triggered at the next higher phase.

I assume Distributed Morphology, (Halle & Marantz, 1993, 1994) a realizational framework in which the syntax operates on phonology-free sets of abstract features. Under this view, Determiner Replacement (46b) consists in substituting the original determiner (a definite determiner DEF in (45)) with a set of features [*D*, *uφ*].[6] More precisely, each determiner is defined by the set of abstract features in (47).

(47) *Features of abstract (i.e. phonology-free) determiners*

$[D, u\varphi, \mathscr{F}]$, where $\mathscr{F}$ is the set of features that define a given determiner (e.g. definite, indefinite, universal quantifier, demonstrative etc).

The modified Trace Conversion (46b) then transforms this abstract determiner into a simpler set of features, interpreted at LF as a definite determiner (48).

(48) *Semantic interpretation of replaced determiner* (based on Heim & Kratzer 1998, 75)

$[\![[D, u\varphi]]\!] = \lambda f : f \in D_{et}$. there is exactly one x such that $f(x)=1$. the unique y such that $f(y) = 1$

[6] This holds of all terminal nodes, though, for simplicity, I represent terminal nodes with abstract features only where relevant.

The interpretation in (48) obtains when $[D, u\varphi]$ is subcategorizes for an NP. Othwerwise, if $[D, u\varphi]$ is intransitive and bound by the controller, it is interpreted as a bound variable, as desired for obligatory control PRO.

I propose further that Trace Conversion can apply not LF, as it was originally proposed, but at the Narrow Syntax (46b). It is usually assumed that Trace Conversion is an LF procedure (Fox, 2002) that does not have any effect on the phonology of a sentence. If we allow Trace Conversion to apply at the Narrow Syntax (i.e. before Spell-Out), we expect that, if something prevents the early merged determiner from being deleted via Chain Reduction, it could actually have overt exponence. As pointed by Takahashi (2019), the set of features $[D, u\varphi]$ is also how a pronoun is defined, assuming that pronouns are simply the realization of the head of a DP (Abney, 1987). In section 4.5, this is exactly how I will model the pronounced PRO's examined in §2.2. This is not a prediction that could be made in an framework where Trace Conversion is circumscribed to LF.

Takahashi (2019) has already proposed that the result of Trace Conversion can be a pronoun. The author's empirical motivation is the fact that, while anti-pronominality effects are not enforced in raising (cf. (11b)), they are in remnant-inducing raising:

(49) (Postal 2004, via Takahashi 2019, (7c))

* How likely to be wrong with her liver is [something really grave]?

According to Takahashi, the derivation of a sentence like (49) would proceed as follows:

(50) [DET$_1$ something really grave λx is how likely [$[D, u\varphi]$ x to be wrong with her liver]]

(50) is generated by merging a lone determiner in the subject position of the predicate *be wrong with her liver*. After the lone determiner raises, the NP complement *something really grave* is WLM-ed. The lone determiner in the base-position is converted into a set of features that is indistinguishable from a pronoun (represented in (50) as '$[D, u\varphi]$' for uniformity), inducing an anti-pronominality effect.

Erlewine & Gould (2016) is an important precedent of Early Trace Conversion. Erlewine & Gould propose that Trace Conversion or, more precisely, Inverse Trace Conversion, can apply at the Narrow Syntax, before Spell-Out. Their empirical motivation is provided by doubly headed relative clauses in Japanese. A relative clause like the one in (51) is called 'doubly headed' because it contains both a head that is external to the relative CP (viz. *sono-ringo-o* 'that-apple-ACC') as well as one that is internal to it (viz. *ringo-o hanbun* 'apple-ACC half').

(51) *Japanese* (Erlewine & Gould, 2016, (11), adapted)

Junya-wa [$_{RC}$ Ayaka-ga **ringo-o** hanbun mui-ta] **sono-ringo-o** zenbu tabe-ta.
Junya-TOP [Ayaka-NOM apple-ACC half peel-PAST] that-apple-ACC all eat-PAST
Literally: 'Junya ate all of those apples [that Ayaka peeled half of the apples].'

The demonstrative *sono* in (51) is obligatory in doubly headed relative clauses. Erlewine & Gould model it as the exponence of the determiner that results from Inverse Trace Conversion, whereby it is the highest copy of movement that undergoes Determiner Replacement. The determiner is available for realization at PF because Inverse Trace Conversion applies at the Narrow Syntax, prior to LF. No linearization issue (e.g. a contradictory linearization statement or a violation of the ban on anti-reflexivity of copy pronunciation, cf. Nunes 2004) because, as a consequence of Inverse Trace Conversion applying at the Narrow Syntax, the copies of movement are no longer identical.

With the bulk of the WLM take on the Movement Theory of Control made explicit, we can turn to how the theory models the properties of obligatory control PRO in (30), repeated below.

(52) a. PRO is a bound variable.
b. PRO is usually phonologically null.
c. PRO is pronominal.
d. PRO can be controlled backwards.
e. PRO can be pronounced as an overt pronoun.

Since the control theory put forward in this paper builds on the Movement Theory of Control, it inherits its account of the main semantic and phonological properties of PRO. In the derivation of an obligatory control sentence (45), the lone determiner merged in the embedded clause undergoes movement into a θ-position in the matrix clause. The higher DP, which contains also the WLM-ed NP complement, binds its lower copy of movement, capturing the fact that PRO is a bound variable (52a). Likewise, the reason PRO is phonologically null (52b), as in the Movement Theory of Control, is that it is a lower copy of movement. The theories differ in the account of the pronominal nature of PRO (52c), the place of backwards (52d) in the theory, and the range of possibilities as to how PRO can be pronounced (52e). We discuss each of these differences in turn.

Recall that the pronominal nature of obligatory control PRO can be demonstrated by anti-pronominality effects and that, crucially, control (53a) and raising (53b) differ in that regard. Some of the relevant data is repeated below:

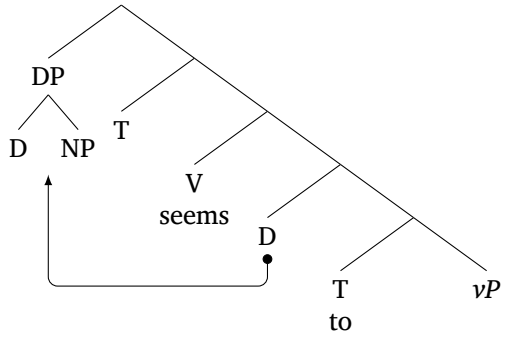
(53) (Landau, 2013, 16f)

a. * Lots of things can be the matter with your transmission [without **PRO** being the matter with mine].

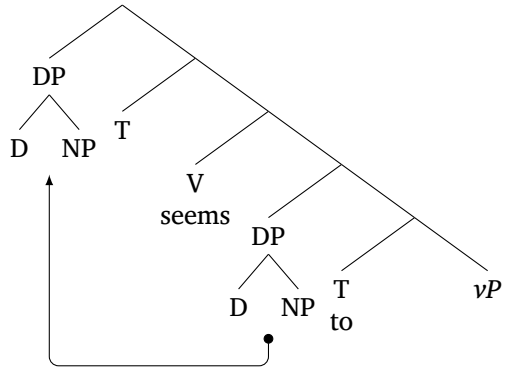
b. Lots of things seem [*t* to be the matter with your transmission].

I assume that, in English, raising complements are not phases, while control complements are phases. Because raising complements are not phases, movement of either a non-countercyclically built DP or a determiner that merges with its NP complement early can move out of them without violating the Phase Impenetrability Condition Chomsky (2001). The two options are depicted in (54) and (55), respectively.

(54) *Raising with WLM*

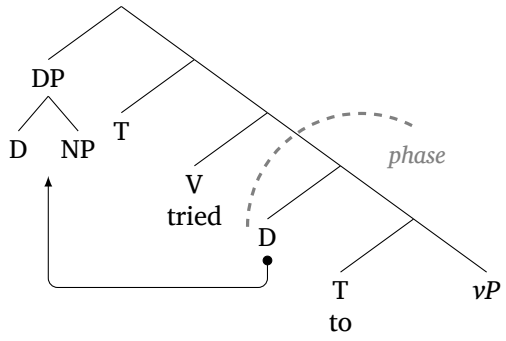


(55) *Raising with early merge*

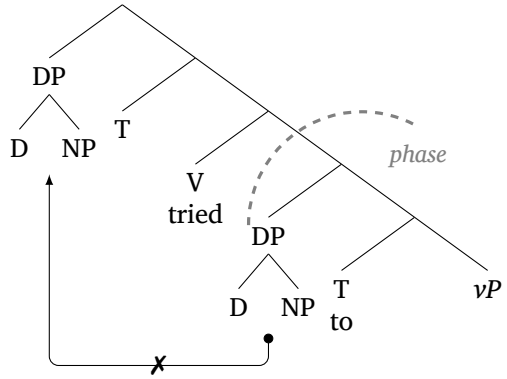


The same range of possibilities is, in principle, available for obligatory control:

(56) *Control with WLM*



(57) *Control with early merge*



Because control clauses are phasal, a DP cannot move out of it (57), due to a violation of the Phase Impenetrability Condition. However, the derivation of WLM sketched in (45) above crucially relies on the

early merged determiner being able to move out of the embedded clause (and into a matrix θ-position). The grammar must then exclude (57), while allowing (56), at least in a language like English.

I had proposed above in (47) that the determiner that is early merged is a set of abstract features, which is then replaced with the features $[D, u\varphi]$, upon the completion of a phase like the embedded CP in the obligatory control sentences. Importantly, $[D, u\varphi]$ are features that can be used to define a pronoun, which places a variable at the edge of a phase in the derivation of an obligatory control sentence proposed above. According to the Grano & Lasnik (2018), this is precisely the configuration when a phase can be neutralized:

(58) *Phase Neutralization* (Grano & Lasnik, 2018, (36a))

Unvalued features on the head of the complement to the phase head keep the phase open.

The empirical motivation for this proposal are banned dependencies across a phase (e.g. tough-movement) that become at least ameliorated if there is a variable near the edge of a phase. (59a) contrasts with (59b), indicating that a dependency like tough-movement can be established across a nonfinite clause, not across a phasal CP. (59c) indicates that the presence of a variable in the Spec position of the phase complement ameliorates tough movement across that phase. This is the phase neutralization effect. Finally, (59d) shows that the presence of a variable within the phase is necessary, though not sufficient to neutralize the phase edge.

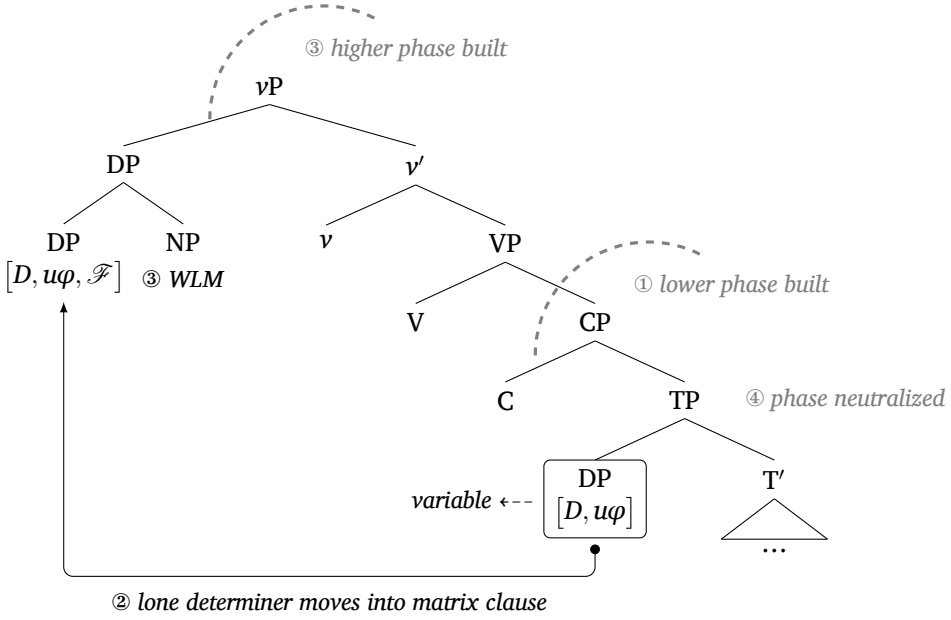
(59) *Tough-movement* (Grano & Lasnik, 2018, (34), with (a) added)

a. The book is too dear for [$_{TP}$ Mary to lend __ to Ann].
b. * The book is too dear for Jim to claim [$_{CP}$ that Mary lent __ to Ann].
c. ? The book is too dear for Jim$_1$ to claim [$_{CP}$ that **he**$_1$ lent __ to Ann].
d. * The book is too dear for Jim$_1$ to claim [$_{CP}$ that Ann lent __ to **him**$_1$].

The intuition that the features of the subject in Spec-TP can influence the phasehood properties of the CP that dominates it is already present in Livitz (2014, §3).

In the proposed obligatory control sentences, this is the point of the derivation when there is an unvalued set of φ-features near the edge of a phase:

(60)



In an extension of the so-called Weak Phase Impenetrability Condition (Chomsky, 2001), I assume that whether or not the phase edge is neutralized (58) is evaluated only at the next higher phase. This is the so-called "Weak Phase Impenetrability Condition" (see also Sheehan & Cyrino 2023):

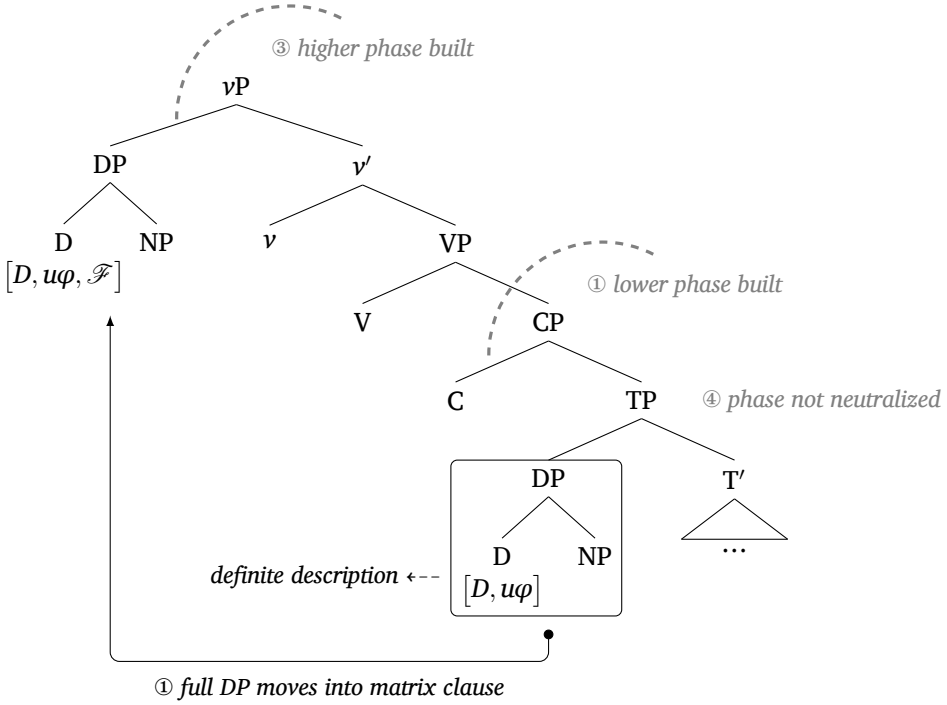
(61) *"Weak" Phase Impenetrability Condition* (Chomsky, 2001, 13)

> The domain of H is not accessible to operations outside HP [a phase]; only H and its edge [i.e. Spec, HP and elements adjoined to HP] are accessible to such operations. […] [A phase] Ph_1 is interpreted/evaluated at the next relevant phase Ph_2.

In order to build the matrix phase *v*P in (60), the lone determiner of the embedded clause moves into its subject position. The early merged determiner then undergoes Early Trace Conversion (46). Under the proposal made in this paper, the lone determiner is converted into a set of features that render it formally indistinguishable from a pronoun. In the configuration of (60), the phase edge is then neutralized, so that the derivation of an obligatory control sentence involving movement across a finite CP does not violate the Phase Impenetrability Condition.

Conversely, if the NP complement had been early merged in the embedded clause, a whole [$_{DP}$ D NP] would be placed near the phase edge, so the result of Early Trace Conversion would be a definite description (Fox, 2002), rather than a variable, under the same assumptions above about the timing of syntactic operations at a given phase:

(62)



In this case, the embedded phase CP would not be neutralized and the derivation in (62) would crash due to a violation of the Phase Impenetrability Condition.

The net result is that, while a derivation with WLM and a derivation without it both converge in the case of raising (cf. (54) and (55)), only a derivation with WLM converge in the case of obligatory control clauses that are phasal (cf. (56) vs. (57)). Additionally, in the convergent derivation (56) (also depicted in more detail in (60)), the residue of raising of the lone determiner is a pronoun. This captures the pronominal character of obligatory control PRO (52c), as well as the fact that raising and control differ in the enforcement of anti-pronominality effects. Importantly, because WLM does not have to apply in the derivation of a raising sentence (55), the residue of standard raising does not have to pronominal (as it would be in (54)).

The occurrence of a pronoun in the subject position of obligatory control, interpreted as a set of features $[D, u\varphi]$ that results from Early Trace Conversion is indisguishable from minimal pronouns (Kratzer, 2009; Safir, 2014). According to these approaches, the grammar has at its disposal an abstract pronoun whose behavior is contextually determined. Minimal pronouns have been defined in various ways, with the different definitions converging in the assumption that minimal pronouns are a set of unvalued φ-features. For concreteness, I assume the following definition:

(63) (Landau, 2015, (28))

X is a minimal pronoun if and only if X = $[D, u\varphi]$.

The minimal pronoun thus defined can then have different realizations depending on the environment where it occurs, including anaphors, referential and bound pronouns, fake indexicals, as well as obligatory control PRO. A theory of control that is based on minimal pronouns has been advanced by Livitz (2014); Landau (2015); McFadden & Sundaresan (2018); Sulemana (2021); Ostrove (To Appear). This type of theory can also be classified as a reductionist approach (along with the MTC discussed in section 3.2), since it does not posit that control PRO is a primitive of the grammar.

The proposal put forward in this paper aligned with such minimal pronoun-based reductionist approaches. However, while minimal pronouns are taken to be base-generated terminal nodes, presumably taken from the Lexicon, I propose a novel way for minimal pronouns to be available at the Narrow Syntax, i.e. the a minimal pronoun $[D, u\varphi]$ is *created* during the course of the syntactic derivation: it is the result of the application of procedures that the grammar employs to yield an interpretable movement chain.

Theories that propose a minimal pronoun usually also assume a separation between form and meaning (cf. a realizational framework such as Distributed Morphology). The result of such assumption is that, as mentioned, the minimal pronoun can then be realized as an anaphor, fake indexical, deictic or bound pronoun, obligatory control PRO, etc. Such a separation is also assumed in this paper. The semantics of obligatory control $[D, u\varphi]$ are a consequence of the fact that it is created by movement: it is locally bound by the constituent that has movement, under the assumption that movement creates a λ-abstraction configuration (Heim & Kratzer, 1998). On the PF side, the residue of movement is usually phonologically (e.g. Nunes 2004), though it can be overtly realized. This is how pronounced PRO's will be modeled in 4.5. Before that, in section 4.4, we turn to the status of backwards control in the current theory.

4.3 Ruling out hyperraising

Before we proceed, a comment is in order regarding a prediction that the current proposal seemingly makes. In the derivation of raising and control sentences, the analysis proposed envisions the following logical possibilities:

(64) Raising

- a. Complement is not a phase (e.g. TP)
  - i. WLM applies.
  - ii. WLM does not apply.
- b. Complement is a phase (e.g. CP)
  - i. WLM applies.
  - ii. WLM does not apply.

(65) Control

- a. Complement is not a phase (e.g. TP)
  - i. WLM applies.
  - ii. WLM does not apply.
- b. Complement is a phase (e.g. CP)
  - i. WLM applies.
  - ii. WLM does not apply.

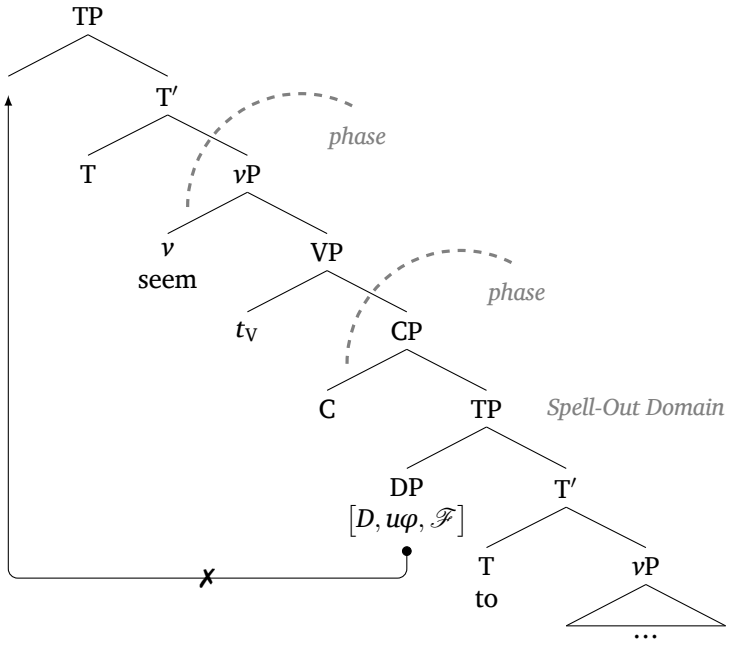
We considered in the previous section the derivation of control clauses with a phasal CP complement. If WLM does not apply, the derivation crashes due to a violation of the Phase Impenetrability Condition. If WLM does apply, the derivation converges, since the Early Trace Conversion transforms the residue of WLM into a pronoun that neutralizes the phase. We will consider the consequences of the clause subcategorized for by a control predicate not being a phase in sections 4.4 and 4.5.

In a raising sentence, if the complement is an infinitival TP, the derivation always converges, whether or not WLM has applied. Independent principles, such as a Condition C or the rules that govern anti-pronominality effects impose additional restrictions. Raising predicates can also subcategorize for CP complements. If WLM does not apply, a Phase Impenetrability Condition violation is induced, similarly to the analogous control derivation mentioned above. An apparent issue arises in the derivation does involve WLM. As in the control case, the residue of WLM would be converted into a pronoun that is in the right configuration to neutralize a phase. The incorrect prediction, then, is that hyperrasing, i.e. raising out of a finite clause (Zyman, 2023) in a language like English should be possible:

(66) * Sindhu seems/is likely/is expected [$_{CP}$ (that) *t* will watch *Grave of the Fireflies*].

In order to rule out this possibility, I follow Legate (2003) in assuming that unaccusative *v*P's are phasal and extend it to raising predicates. Under this assumption, the derivation of a sentence like (66) would go as follows:

(67)



The matrix *v*P in (67) is an unaccusative phase. Assuming the Weak Phase Impenetrability Condition (61), the embedded subject, a determiner that is merged alone, is contained inside a Spell-Out Domain. As such, it is sent off to the interfaces before it can hyperraise to Spec-TP. The same assumptions underlie the control derivation (45), though the embedded subject can move into the matrix because it targets a θ-position inside the phase *v*P.

While the derivation depicted in (67) crashes, hyperraising can still be ruled in in the languages that allow for this construction. Additional mechanisms such as movement of the embedded subject to the edge of the embedded clause or phase deactivation may be available in such languages. See an overview of these strategies in Zyman (2023).

With this apparent issue solved, we can resume the discussion of the consequences of the analysis proposed.

4.4 Phasehood and backwards control in Modern Greek

As mentioned before, Takahashi & Hulsey (2009) remark that WLM is a possibility that is available in the grammar, though it is not obligatory. Independent principles of the grammar rule in or out the representations created by WLM or early merge (e.g. LF-interpretability and the Case Filter). In the control derivation proposed above, an independent restriction is recruited, namely, the Phase Impenetrability Condition. What was crucial was that the embedded clause in the control was a phase, so movement of a full DP was blocked (while it was allowed with early merge of a determiner alone, yielding a Phase Neutralization effect).

A prediction that arises from this phase-based proposal is that, if the embedded control clause is not a phase, movement of an early merged [$_{DP}$ D NP] should be possible without a Phase Impenetrability Condition violation, so that a complete representation of [$_{DP}$ D NP] is present in the embedded clause. Languages with backwards control may instantiate exactly this pattern, with the moved [$_{DP}$ D NP] conspicuously pronounced in the embedded clause.

Recall from section 2.1 that Modern Greek displays backwards control. That the embedded clause where the controller is realized is not a phase, as predicted by the present proposal, is demonstrated by the licensing of the emphatic negative indefinite KANENAS 'any'. According to Grano (2012), KANENAS must

be licensed by clause-mate negation. This negative indefinite cannot be licensed by negation across an *oti* clause:

(68) *Modern Greek* (Grano, 2012, 190, adapted)

* O Yanis **den** ipe [oti elise KANENA provlima].
the Yanis not say.PP.3SG that solve.PP.3SG any problem
Intended: 'Yanis didn't say that he solved any problem.'

But licensing is possible across a backwards control *na* clause:

(69) *Modern Greek* (Christos Christopoulos, p.c.)

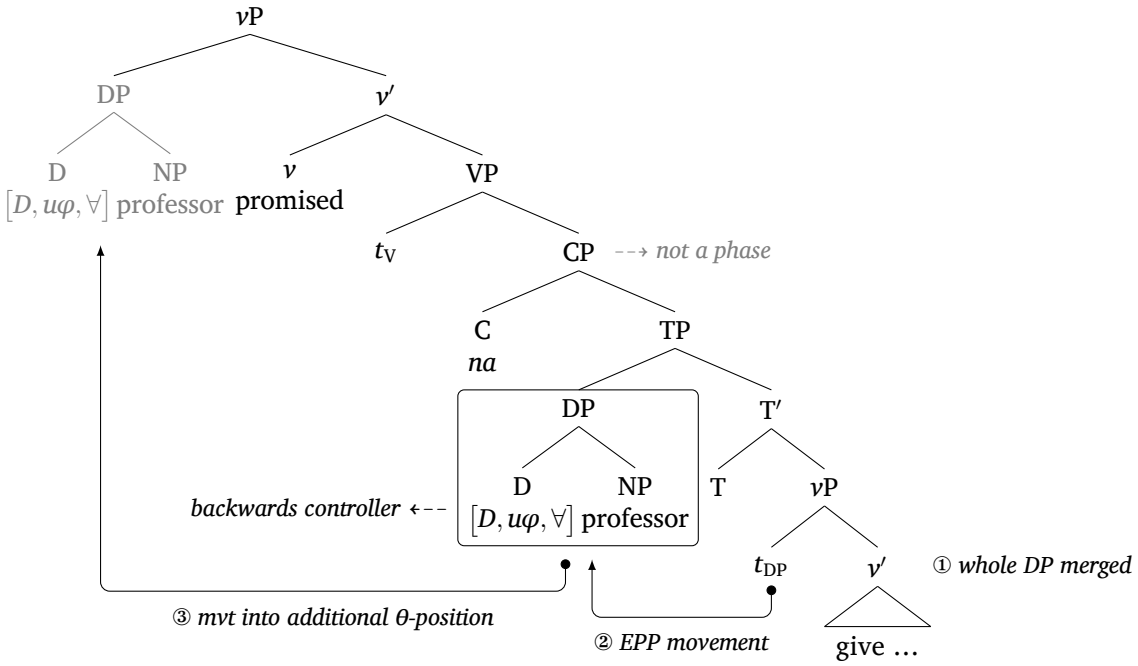
I kathigités **dhen** iposkhethikame tis Marias [na lisoume **i** **kathighites** KANENA provlima].
not promised.1PL the.GEN Maria.GEN SBJV solve.1PL the teachers any problem
'We teachers did not promise Maria to solve any problem.'

In order for the contrast between (68) and (69) to be accounted for, a simplex clause-mate condition does not suffice. A contrast between the *oti* clause in (68) and the *na* clause in (69) can, however, be stated in terms of phases: *oti* clauses are phases, so that they prevent the licensing of KANENAS, while the *na* clause in obligatory control sentences is not a phase, so that they do not prevent the licensing of the negative indefinite.

Because the *na* clause in a sentence like (69) is not a phase, A-movement can cross it without incurring a violation of the Phase Impenetrability Condition. In other words, the early merge of a lone determiner resulting in the occurrence of a pronoun at the edge of the embedded clause (60) is not a necessary condition for the derivation of a sentence like (69) to converge. The controller realized in the embedded clause is the full DP that A-moves across the *na* clause.

A snapshot of the derivation of a sentence like (17a), repeated below, is given in (70) (with irrelevant details omitted for simplicity). For concreteness, I assume that *na* is the head of a non-phasal CP.

(70) káthe kathigitís iposkhethike stus mathites tu [na tus dhosi **kathe** **kathighitis** kalus vathmus]. 'Every professor$_1$ promised their$_{1/2}$ students to give them good grades.'
promised.3SG to-the.PL.ACC student.PL.ACC 3SG.GEN [SBJV 3PL.GEN give.3SG every professor.NOM good.ACC grades.ACC]



In sum, in backwards control in Modern Greek, there is a correlation between where the controller is pronounced (inside the embedded clause) and the non-phasal status of the complement clause. This exactly what is expected under the theory proposed here, with the additional assumption, independently needed, that syntactic operations are restricted by phasehood.

As alluded to above, Polinsky & Potsdam (2002) and Alexiadou *et al.* (2010) remark that backwards control poses a challenge to theories of obligatory control that are based on PRO as a sui generis category. The reason is that the controller is pronounced as a full DP in the position where the phonologically null PRO occurs. Polinsky & Potsdam and Alexiadou *et al.* also observe that the this phenomenon constitutes strong evidence in favor of the Movement Theory of Control. The ability to account for backwards control (52d) is another property that the current theory inherits from the Movement Theory.

However, according to the original version of the MTC, all instances of obligatory control resemble backwards control, in that there is a full representation of the controller in all such sentences. As mentioned above, this undergenerates the pronominal nature of PRO, as well as the contrast between control and raising regarding anti-pronominality effects in a language like English. The proposed version of the MTC, in contrast, does not incur into this issue, since not all instances of obligatory control involve the full representation of the controller in the embedded clause. Critically, whether or not a full representation of the controller or just a pronoun occurs in the embedded clause is predictable from an independent property, namely, the phasehood properties of the control clause. In other words, the current theory differs from the original MTC not only in its ability to account for the pronominal nature of PRO in English, but also in its ability to predict whether or not the residue of the controller's raising is *not* pronominal.

4.5 Modelling pronominal overt PRO's

The Copy Theory of Movement (Chomsky, 1995) provides a typology of movement realization, with languages varying on which copies are pronounced. Assuming the simplest non-trivial chain, some of the logical possibilities are listed in (71), with examples from Ā-movement phenomena:

(71) a. Highest copy pronounced (e.g. overt *Wh*-movement in e.g. English)
b. Highest and lowest copy pronounced (e.g. multiple *Wh*-copy pronunciation in e.g. German)
c. Lowest copy pronounced (e.g. *Wh*-in situ in e.g. Mongolian)

Interestingly, obligatory control displays the same range of possibilities. If obligatory control is derived by movement, then we should also expect variability in the realization of the raised controller. In canonical

instances of control, the higher copy of movement is pronounced, while the tail of the chain is deleted (72a). If both copies are pronounced, the result is copy control (72b) (Lee, 2003). Finally, if the movement into an additional θ-position is "covert", the result is backwards control (72c) (Polinsky & Potsdam, 2002).

(72) a. *Highest copy pronounced: "canonical" control*
Sindhu tried [Sindhu to eat natto].

b. *Highest and lowest copy pronounced: copy control in San Lucas Quiaviní Zapotec* (Lee, 2003, (62), adapted))
R-cààa'z **Lia** **Paamm** [g-ahcnèe **Lia** **Paamm** Gye'eihlly].
HAB-want FEM Pam IRR-help FEM Pam Mike
'Pam wants to help Mike.'

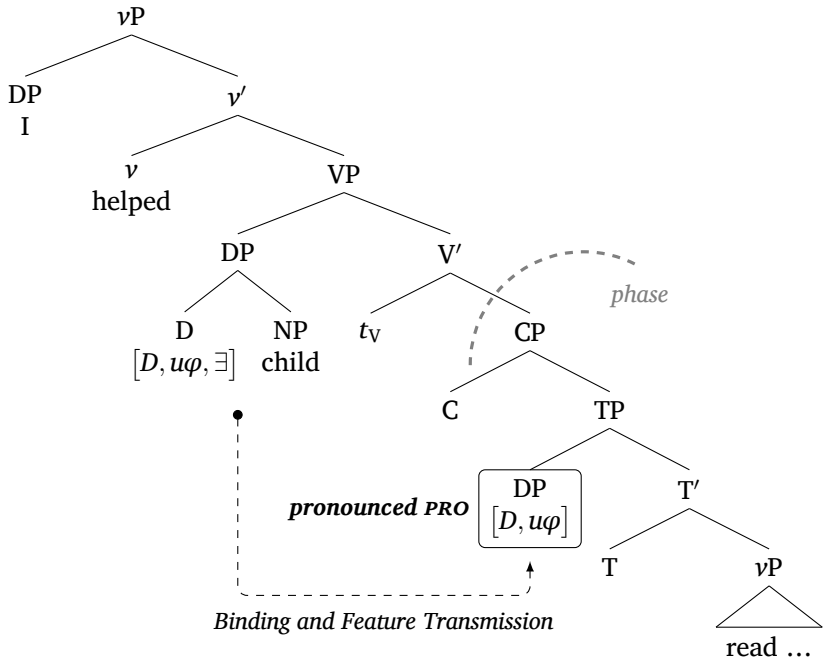
c. *Lowest copy pronounced: backwards control in Tsez* (Polinsky & Potsdam, 2002, (2))
kidbā [**kidbā** ziya bišra] yoqsi.
girl.ERG cow.ABS feed.INF began
'The girl began to feed the cow.'

In the context of the theory put forth in this paper, a forth possibility of chain resolution is made available. Given the possibility of pronouncing the tail of a movement chain, if obligatory control can be derived by merging a lone determiner before the rest of the controller is WLM-ed, then we expect the lone determiner, the tail of movement in this case, to be amenable to overt exponence as well.

(73) *Highest copy pronounced (revised)*
a. **Sindhu** tried [Sindhu to eat natto].
b. **Sindhu** tried [$[D,\ u\varphi]$ to eat natto].

(73b) is how I propose the pronounced PRO's examined in section 2.2 to be derived, since the lower copy, after Early Trace Conversion (46), is indistinguishable from a pronoun. The analysis proposed is in (74). For concreteness, I use Wolof as an exemplar.

(74) Dimbali-na-a a-b xale [*PRO / **mu** jàng téere b-i].
help-NA-1SG INDEF-CM.SG child 3SG.SUBJ read book CM.SG-DEF
'I helped a child read the book.'



Binding gives rise to Feature Transmission, defined in (75). I assume Heim's (2008) definition (though also see Kratzer 2009).

(75) *Feature Transmission Under Variable Binding* (Heim, 2008, (40))

In the derivation of PF, features of a DP may be copied onto variables that it binds.

The pronoun ends up with φ-features that match that of the WLM-ed controller as a consequence of Feature Transmission. I assume the application of Feature Transmission is uniform, so it also takes place in cases where PRO is phonological null, as in the sample derivation in section 4.2. In a language like English, PRO's φ-features can be seen in an anaphor bound by it:

(76) a. I tried [PRO not to criticize myself].

b. Jaimie tried [PRO not to criticize themselves].

The table below summarizes the syntactic, phonological, and semantic properties of the lone determiner whose movement into an additional θ-position yields obligatory control in a language like Wolof. In the derivation depicted in (74), the lone determiner is an indefinite. After Early Trace Conversion (46), it becomes $[D, u\varphi]$, with the φ-features then valued as 3SG as a consequence of Feature Transmission (75). On the PF branch, the original indefinite determiner, which was moved to a higher position and to which the NP *xale* 'child' is countercyclically merged, is exponed as the indefinite determiner *a*. $[D, u\varphi : 3\text{SG}]$, in turn, is exponed as the pronoun *mu*. At LF, it is interpreted as a bound variable, under the assumption that movement creates a λ-abstraction configuration (Heim & Kratzer, 1998).

(77)

	Narrow Syntax	Exponent	Interpretation
Matrix controller	$[D, u\varphi, \exists]$	*/a/*	indefinite determiner
Embedded PRO	$[D, u\varphi : 3\text{SG}]$	*/mu/*	bound variable

We are now in the position to fulfill the desideratum (52e), i.e. the need to account for the fact that obligatory control PRO may be exponed as an overt pronoun. This instance of pronounced PRO is analyzed as the realization of the tail of the movement chain created by movement, where the element that moves is a lone determiner, which is then converted into a set of features that renders it identical to a pronoun. It is crucial that the procedure responsible for such conversion apply at the Narrow Syntax, so that it can feed a PF operation like Vocabulary Insertion. Under this view, pronounced PRO's in e.g. Wolof, Bùlì, and Mandarin can be taken to be on a part with Japanese doubly headed relative clauses (51), in that their analysis crucially involves the result of Trace Conversion being visible to PF, instead of being restricted to LF.

The analysis proposed also offers an account of the generalizations in (29), repeated below.

(78) a. When an alternation is available between a phonologically null PRO and a pronominal overt PRO, the former tends to occurs in a clause that is structurally less complex than the clause where the overt PRO tends to occur. This correlation can be profitably restated in terms of phasehood.

b. A pronominal overt PRO is never a separate category in the languages that allow for them. Rather, it is identical to pronouns that occur independently in these languages, albeit one that is necessarily locally bound.

Starting with (78a), the correlation between clause size and the phonological realization of PRO is captured in the present proposal by the role that phases play in the analysis. Recall from (60) that a pronoun is the result of Early Trace Conversion applying to the lone determiner that is merged in the embedded control clause. Critically, the latter is phase that can only be neutralized if its subject position is filled with a lone determiner — is a full DP is base-generated at this position, the embedded phase cannot be neutralized because, after Early Trace Conversion, causing the derivation to crash due to a violation of the Phase Impenetrability Condition. In languages like Wolof and Bùlì, the null variant occurs in clauses that are reduced or restructured, so the absence of a pronounced follows trivially from the assumption that these clauses lack a subject position altogether (Wurmbrand, 1998). In languages like Mandarin, where the null PRO occurs in CP clause (Li, 2021), this category can be taken to be non-phasal, so that movement out of it

(and into a matrix θ-position) does not have to leave behind a pronominal residue. In sum, an account of the tendency expressed in (78a) is afforded in the theory proposed in this paper by phasehood. Importantly, whether or not a control clause is a phase can be demonstrated by independent properties (see e.g. clitic climbing and other language-specific diagnostics in section 2.2).

However, other logical possibilities have to be taken into account. Within the theory advanced here, it is also possible that the derivation of a control sentence involves a non-phasal complement. This is precisely how backwards control was analyzed in section 4.4. There are two other convergent possibilities to be considered. First, it is possible that the derivation proceeds just as it does in the backwards control case, but the fully assembled DP in the embedded clause is deleted (i.e. the control movement is overt and not convert, as in backwards control). In this case, we expect a control clause whose PRO subject does not have pronominal properties. Finally, the last logical possibility to be considered is one that involves a control clause that is not phasal and where WLM does apply, but the lone determiner is pronounced. While this option is not available for restructuring cases like Wolof and Bùlì, this is perhaps the case for languages with an obligatorily overt PRO such as San Martín Peras Mixtec. Ostrove (To Appear) shows that the obligatorily overt PRO occurs in untensed subjunctive clauses that are structurally reduced. This can be demonstrated by the fact that these clauses cannot the landing site of quantifier raising. What is required instead is that quantifier raising targets the edge of the matrix clauses.

(79) *San Martín Peras Mixtec* (Ostrove, To Appear, (24/25))

{ Kwa'ă ko'ŏ$_i$ } nàntŏso Mateo [{ *kwa'ă ko'ŏ$_i$ } nakatsya =**rà** __$_i$].
many plate forget:COMP Mateo many plate wash:IRR =he
'Mateo forgot to wash many plates.'

In contrast, in finite and tensed subjunctive embedded clauses, quantifier raising must target their edge (Ostrove, To Appear, (20–23)). In order to account for this contrast, one could appeal once again to a contrast based on phasehood: non-phasal embedded clauses cannot be a target of quantifier raising, while their phasal counterparts can. Furthermore, because the embedded clause in (79) contains an overt subject, it cannot be an instance of restructuring in Wurmbrand's (1998) sense, i.e. a subject-less VP. Taken together, these observations can lead to the conclusion that the embedded clause in (79) is a non-phasal projection such as a TP. While non-phasal (hence the impossibility of intermediate quantifier raising), the TP is complex enough for a subject position (i.e. the overt PRO). The overt PRO in a language like San Martín Peras Mixtec could thus instantiate the movement of a lone determiner out of non-phasal clausal complement and which is pronounced.[7]

As mentioned before, the proposal made here gives rise to a set of logically possible derivations for an obligatory control sentence. Independent constraints such as the Weak Phase Impenetrability Condition (61) determine whether or not such possibilities are in fact convergent. Additionally, I assume that languages vary with respect to the type of complement that their control predicates can subcategorize for, which I take to be a non-controversial assumption. The theory proposed here, then, gives rise to the following typology of obligatory control PRO realization:

(80) *PRO realization typology according to the Wholesale Late Merge Theory of Control*

a. Complement is a phase
   i. WLM applies
      - Lower copy deleted: null PRO with pronominal properties, e.g. English
      - Lower copy realized: overt PRO in alternating languages, e.g. Wolof and Bùlì (null PRO variant: restructuring), as well as Mandarin
   ii. WLM does not apply: derivation crashes due to a Phase Impenetrability Condition violation.
b. Complement is not a phase
   i. WLM applies.
      - Lower copy deleted: null PRO with pronominal properties.

[7] Ostrove (To Appear) argues convincingly that the overt PRO in San Martín Peras Mixtec cannot be derived by movement. I will comment on this issue in section 5. (79) is taken here to exemplify what a language could look like that displays an overt PRO sitting in a clause that is not a phase.

- Lower copy realized: obligatorily overt PRO, e.g. San Martín Peras Mixtec

ii. WLM does not apply.
- Lower copy deleted: null PRO without pronominal properties.
- Lower copy realized: backwards control, e.g. Modern Greek

It remains to be seen whether all the expected variants listed in (80) are in fact attested. Nonetheless, I believe the ability of theory proposed in this paper to yield a typology of possible overt PRO languages based on an independently diagnosable property (viz. phasehood) to represent of a contribution to the advancement of control theory, specially as it regards an underexamined control phenomenon from a crosslinguistic perspective.

We can now turn to generalization (78b). In the context of the Wholesale Late Merge Theory of Control, an account is provided by the realizational framework this theory is couched in. As mentioned above, following the tenets of Distributed Morphology, I assume a dissociation between form and meaning, with the Narrow Syntax operating on abstract sets of features. The $[D, u\varphi]$ that result from Early Trace Conversion can be deleted, as it is usually the case for the residue of movement. However, it is also available for exponence. An auxiliary assumption I make is that what we usually call pronouns consists in an underspecified Vocabulary Item that can expone different, though related terminal nodes, e.g. deictic and bound pronouns. For concreteness, some Vocabulary Items for pronouns in Wolof, listed in the table in (81), are offered in (82). Pronounced PRO's in the language are realized as nominative pronouns (Fong, To appear).

(81) *The pronominal system of Wolof* (based on Zribi-Hertz & Diagne 2002, (29))

	ACC clitics	OBL pronouns	NOM pronouns
1SG	ma	man	ma
2SG	la	yaw	nga
3SG	ko	moom	mu
1PL	ñu	ñoom	ñu
2PL	leen	yeen	ngeen
3PL	leen	ñoom	ñu

(82) *Some pronominal Vocabulary Items in Wolof*

a. $\begin{bmatrix} \text{Cat}: & \text{D} \\ u\varphi: & \text{3SG} \end{bmatrix} \leftrightarrow$ /mu/

b. $\begin{bmatrix} \text{Cat}: & \text{D} \\ u\varphi: & \text{3SG} \\ \text{Case}: & \text{ACC} \end{bmatrix} \leftrightarrow$ /ko/

c. $\begin{bmatrix} \text{Cat}: & \text{D} \\ u\varphi: & \text{3SG} \\ \text{Case}: & \text{OBL} \end{bmatrix} \leftrightarrow$ /moom/

(82a) is chosen to expone the subject position in (74) because it is an Vocabulary Item that matches its features.[8] Because this Vocabulary Item is underspecified, it can also expone other terminal nodes with matching features, such as deictic and other bound pronouns. Even though (74) is a Vocabulary Item that can expone different pronominal terminal nodes, each with its own semantic properties, the fact that the obligatory control PRO that it is exponed with is necessarily a bound variable is enforced by the fact that it is the residue of A-movement, which creates a λ-abstraction configuration. The fact that pronominal PRO's are not a sui generis category, but rather a pronoun that is employed elsewhere in a given language is a consequence of the availability of underspecified Vocabulary Items that can expone different pronominal terminal nodes. Given the dissociation between meaning and form, the syntactic and semantic properties of such pronominal terminal nodes are due to other principles, e.g. how A-movement chains are interpreted.

The account proposed for the generalization (78b), thus, relies on the specifications of particular Vocabulary Items that are available in a given language (coupled with the possibility of realizing the tail of a movement chain, as opposed to its deletion). A prediction that arises from this proposal is that a PRO that is

[8] As mentioned above, in Wolof, overt PRO's are nominative pronouns. Given that they occur in nonfinite clauses, it is possible that they are not assigned any case. Following Preminger (2011), I assume that an unvalued case feature does not cause a derivation to crash. Nominative pronouns can either lack a case specification, as in (82a), or be specified as having an unvalued case feature. Nothing hinges on the choice between these options. Other appropriate Vocabulary Items can be proposed for pronominal PRO's in other languages. In Mandarin, for instance, pronominal Vocabulary Items presumably lack a case specification altogether.

available for exponence can be realized by a Vocabulary Item that is not pronominal. This is perhaps how overt PRO's that are realized as anaphors can be captured. Reflexive overt PRO's have been documented in Tamil, illustrated in (83), Korean (Lee, 2009), and Chirag (Ganenkov, To Appear).

(83) *Tamil* (Sundaresan, 2010, brackets added)

ramani [**taan** saadatt.ai saappi.ɖ.a] paa.tt.aan.
Raman.NOM self.NOM rice.ACC eat.INF try.PST.3M.SG
'Raman tried to eat the rice.'

I leave a WLMTC account for anaphoric overt PRO's for future research. Tentatively, the anaphoric realization of a pronounced PRO could be tied to the mechanics of Feature Transmission (cf. the transmission of a feature [REFL] in Kratzer 2009).

Finally, I would like to wrap up this section with a speculative discussion about the rarity of pronounced PRO's. While pronounced PRO's of different varieties (viz. the pronominal (and anaphoric) PRO discussed in this section, backwards, and copy control) have been documented in a few, mostly unrelated languages, to the best of my knowledge, in the overwhelming majority of languages, obligatory control PRO is phonologically null. A factor that must perhaps be considered is that the fact that, under the Movement Theory of Control, control is an instance of A-movement — to the very least, it can be considered as a type of A-dependency. While covert Ā-movement is widely attested, analyzed as the realization of a lower copy, covert A-movement has been attested less frequently (e.g. backwards raising, Potsdam & Polinsky 2012). The MTC-based analysis proposed implies that these facts should be related.

Another point of consideration are language-specific factors conditioning the exponence of certain terminal nodes. A notable proposal in this regard is Allotey's (2021) analysis of pronounced PRO's in Gã. In this language, an irrealis morpheme is suffixed to the pronounced PRO:

(84) *Gã* (Allotey, 2021, (2a))

Mi$_i$ tao [(ni) **ma**$_i$ na bo].
1SG want C 1SG see.ING 2SG
'I want to see you.'

In (84), the embedded pronominal subject *ma* is to be decomposed into *mi* '1SG' (cf. matrix subject) and *-a* 'IRREALIS'. Allotey (2021, 28) proposes that PRO must be overt in Gã in order for a host to be provided to the irrealis suffix. This is essentially the spirit of Landau's (2006) P(honological)-Recoverability:

(85) *P-Recoverability* (Landau, 2006, (49))

In a chain $< X_1 \ \ldots \ X_i \ \ldots \ X_n >$, where some X_i is associated with phonetic content, X_i must be pronounced.

Building on Allotey's insight, it is possible that the rarity of pronounced PRO's is related to language-specific features and constraints, such as the occurrence of an affix that is attached to subjects and a ban on stranded affixes.

An analysis that takes into consideration phonological constraints may also be applicable to languages where a pronounced PRO must be associated with focus. This group includes European Portuguese, Hungarian, and Italian (Livitz 2014, chapter 4 and references therein), as well as Chirag (Ganenkov, To Appear, (17)), which is illustrated below.

(86) *Chirag* (Ganenkov, To Appear, (18/20a))

a. dami [**du**=**jal** š:a w-ač'-i] b-ik:-an-da.
1SG.DAT 1SG.ABS=only home M.SG-come:PF-INF N.SG-want:IPF-DUR-1
'I want to come back home ALONE.'

b. * dami [**du** š:a w-ač'-i] b-ik:-an-da.
1SG.DAT 1SG.ABS home M.SG-come:PF-INF N.SG-want:IPF-DUR-1
Intended: 'I want to come back home.'

This is in fact the proposal made by Livitz (2014). Livitz proposes that control PRO is a minimal pronoun (cf. (63)) whose φ-features are valued via Agree (Landau, 2001). More precisely, the minimal pronoun is a defective goal, which ends up identical to the probe after Agree. The minimal pronoun is then deleted under identity. While Livitz specifically argues against the MTC, the relevant point of resemblance is that, in their theory, PRO's phonological nullness is epiphenomenal and reduces to independent restrictions on movement chain resolution, just as in the MTC. The two proposals differ, however, in how the chain in question is created. Specifically, according to Livitz, PRO, as a minimal pronoun, participates in a dependency with the controller that is created by Agree (and not movement), but which is formally indistinguishable from a movement chain. A prediction that arises from this proposal is that, if identity cannot be obtained between the minimal pronoun and its antecedent, the former is not deleted under identity and can, thus, be overtly realized. This is how Livitz (2014, chapter 4) models pronounced PRO's which are associated with focus. More precisely, the author proposes that overt PRO's associated with focus merge with a null focus operator that prevents them from being identical to the controller and, therefore, resist deletion under identity.

In the next section, we turn to a comparison between the current analysis and other accounts of pronounced PRO's, with a focus on whether or not their logic is consistent with the generalization that pronounced PRO's are never a separate category in the languages that feature them (78b).

4.6 Previous accounts of pronounced PRO's

Three types of approaches can be distinguished that account for pronounced PRO's. First, there are approaches that propose that the pronounced PRO that a given language exhibits is a lexical item that is similar or identical to PRO, except for its phonological content. Second, another class of analysis models the pronounced PRO as an instance of the realization of more than one link in a movement chain. The theory advanced in this paper belongs to this category. Finally, a set of proposals is based on analyzing PRO as a minimal pronoun, which can undergo exponence. In this section, we will briefly examine these approaches, with a focus on how they account for the generalization that pronominal overt PRO's are not a sui generis category, but rather a pronoun that already exists in a given language (78b).

As mentioned above, Allotey (2021) analyses obligatorily overt pronominal controlled subjects in Gã and proposes that the overt pronoun in this language is a lexical item just like PRO, except that it has overt exponence. The pronoun must match the features of the controller, which is achieved via Upwards Agree. It must be bound because of its lexical specification as a control PRO. In a similar vein, Li (2021) analyzes overt PRO's in Mandarin and argues that, because the overt PRO behaves just like other pronouns and reflexives in the language and unlike the null PRO, then overt PRO must be a category of its own. Such proposals do not account for generalization (78b). Because the overt PRO is a pronoun with its own properties, the fact that it is identical to other pronouns found in a given language is an accident. Alternatively said, the logic of these proposals is consistent with an overt PRO whose form is different from that of other pronouns in a given language. As discussed in section 2.2, the available data contradicts this possibility. Likewise, another question that could be asked regards the semantics of pronominal overt PRO's. The fact that they are obligatorily bound variables must come about as a stipulation. One may wonder why other identical forms in the same languages can occur as either a referential or a bound pronoun.

Another set of proposals is couched in the MTC. Within this class, a division can be drawn between partial deletion approaches and the no-deletion approach advocated for in the current paper. Sulemana (2018) and Fong (To appear) analyze pronominal overt PRO's in Bùlì and Wolof, respectively, under a MTC framework. More precisely, the authors propose that the overt PRO that appears in these languages is the result of partially exponing the copy of the controller that sits in the embedded subject position (i.e. the position it occupies prior to moving into an additional θ-position in the matrix clause). More precisely, a subpart of the DP is deleted (viz. the NP complement), so that the remaining structure is also indistinguishable from a pronoun (cf. Van Urk 2018). This accounts for generalization (78b). While these approaches can account for pronominal overt PRO's, one may wonder about the matching between the pronominal PRO and those of the controller. The evidence available indicates that full matching is necessary. However, partial deletion has been resorted to in order to account for cases of resumptive pronouns that have a proper subset of the features of their antecedent (Georgi & Amaechi, 2022; Yip & Ahenkorah, 2023). The current approach does not predict that a proper subset relation is possible between the features of the pronominal PRO and the controller, under the assumption that Feature Transmission

(75) results in full matching. It remains to be seen whether partially matching pronominal PRO's are in fact attested. In this regard, it is worth pointing out that, in Akan, obligatory control PRO is necessarily an overt pronoun (Allotey 2021, Comfort Ahenkorah, p.c.) that fully matches the features of the controller (Comfort Ahenkorah, p.c.), even though resumptive pronouns in this language only partially match the features of its antecedent (Yip & Ahenkorah, 2023).

Lastly, McFadden & Sundaresan (2018), Sulemana (2021), and Ostrove (To Appear) assume that control is derived based on a minimal pronoun which can be realized via a process such as Vocabulary Insertion. As mentioned above, the current proposal can also be taken as an instance of control theory that is based on a minimal pronoun, though one that is not base-generated (as in the proposals just listed), but rather one that results from a determiner that merges alone in the embedded clause and which undergoes Early Trace Conversion (46). This is a new version of the MTC which results from combining it with WLM. As mentioned in footnote 5, the MTC has faced much criticism, specially as it regards partial control and case connectivity facts. However, phenomena like backwards and copy control also provide strong support in favor of the MTC and challenge theories that do not resort to movement. A solution to this tension has been offered by hybrid theories such as Van Urk (2010) and Grano (2012), which combine some version of PRO, as well as movement in order to account for different facets of control phenomena. Such proposals take for granted that each type of control theory has its virtues and limitations, doing away with the latter by restricting the application of PRO or movement to appropriate phenomena.

I would like to suggest that the proposal made in this paper, when considered side-by-side with the proposal made by McFadden & Sundaresan (2018), Sulemana (2021), and Ostrove (To Appear), could be part of such a hybrid account of control. The distinction between a base-generated minimal pronoun and a minimal pronoun that is built during the course of the derivation is parallel to a contrast between PRO as a primitive of the grammar and a PRO that results from movement in the above-mentioned hybrid theories. Ostrove, for instance, argues convincingly that pronounced the PRO in Martín Peras Mixtec cannot be derived by movement. However, as we have discussed a few times, a theory of control that is completely devoid of movement is hard-pressed to account for phenomena like backwards and copy control, which the MTC-based theory I proposed here can predict. It is, thus, possible that a different incarnation of a hybrid theory of control could be formulated that involved both a base-generated minimal pronoun, as well as one that is created by movement, each applied to an appropriate control phenomenon.

5 Concluding remarks

The theory proposed in this paper attempts to advances control theory by proposing a combination between the Movement Theory of Control with Wholesale Late Merge, which preserves the virtues of the original formulation, but also combines it with the virtue of a PRO-based theory of control like Chomsky's (1981), notably its ability to account for the enforcement of anti-pronominality effects in control. In addition, it expands its empirical domain by accounting for pronounced instances of PRO, but without the stipulation of additional categories. Finally, the analysis continues a discussion about the timing of procedures like Trace Conversion (cf. Erlewine & Gould 2016), which is usually considered part of LF, exclusively. The availability of a pronoun at the Narrow Syntax was instrumental in the modelling of pronominal overt PRO's. If Trace Conversion were restricted to LF, a component of the grammar that does not interface with PF, such modelling would not have been possible.

A Deriving Visser's Generalization without Agree

Visser's Generalization is "the observation that verbs whose complements are predicated of their subjects do not passivize" (Bresnan, 1982, 402). In (87a), the main predicate *regard* takes an adjectival predicate as its complement (viz. *poompous*), which is predicated of the matrix object *him*. In (88a), *strike* subcategorizes for the same adjectival complement, though in this case, it is predicated of the subject. While the former can passivize (87b), the latter cannot (88b).

(87) (Bresnan, 1982, (83a) and (85a))

a. His friends regard him as pompous.
b. He is regarded by his friends as pompous.

(88) (Bresnan, 1982, (84a) and (86a))
a. He strikes his friends as pompous.
b. * His friends are struck (by him) as pompous.

In the domain of control, this generalization translates into the impossibility of an implicit agent to bind PRO. More precisely, subject control predicates cannot passivize (89), but object controls predicates can (90). In (89b), as soon as the subject control predicate is passivized, only an object control reading becomes available. PRO cannot be interpreted as coindexed with the matrix subject that is demoted to the status of an implicit agent as a consequence of passivization.

(89) a. Solfrid$_1$ promised Faatu$_2$ [PRO$_{1/*2}$ to watch *Grave of the Fireflies*].
b. Faatu$_2$ was promised [PRO$_{*1/2}$ to watch *Grave of the Fireflies*].

If the matrix predicate is an object control verb, the object control construal is trivially preserved, since there is no binding relationship between the external argument of such verbs and PRO to begin with.

(90) a. Solfrid$_1$ persuaded Faatu$_2$ [PRO$_{*1/2}$ to watch *Grave of the Fireflies*].
b. Faatu$_2$ was persuaded [PRO$_2$ to watch *Grave of the Fireflies*].

Visser's Generalization is made particularly clear by control shift predicates like *ask*. In the active, (91a) is ambiguous between a reading where PRO is bound by the matrix subject or the matrix object. However, if *ask* is passivized (91b), only the matrix indirect object is a viable controller.

(91) a. Solfrid$_1$ asked Faatu$_2$ [PRO$_{1/2}$ to leave the room].
b. Faatu$_2$ was asked [PRO$_{*1/2}$ to leave the room].

Van Urk (2013) observes that, in a language like Dutch, an implicit agent can control PRO, as long as the ditransitive control predicate is able to assign inherent case to its indirect object. This holds of verbs like *beloven* 'promise' and *aanbieden* 'offer' (92a), but not of a verb like *overtuigen* 'convince' (92b).

(92) *Dutch* (Van Urk, 2013, (8) and (10a), adapted)
a. Er werd mij$_2$ beloofd / aangeboden [PRO$_1$ om me$_2$ op de hoogte te houden
there was me.DAT promised / offered INF.COMP me on the height to keep.INF
].

'It was promised /offered to me to keep me informed.'
b. * De leraren$_2$ werden overtuigd [PRO$_1$ om ze$_2$ te mogen kietelen].
the teachers were convinced INF.COMP them to may.INF tickle.INF
Literally: 'The teachers were convinced to be allowed to tickle them.'

Based on data like this, Van Urk argues that Visser's Generalization must be revised, so that it regards only cases of passivization that involve Agree between T and an overt argument of a control predicate:

(93) *Revised Visser's Generalization* (Van Urk, 2013, (12))
Obligatory control by an implicit subject is impossible if an overt DP agrees with T.

Wurmbrand (2021) adds to puzzle that, by observing that, even though an implicit argument cannot control PRO, a *by*-phrase can:

(94) *Dutch* (Wurmbrand, 2021, (5a), adapted)

De leraren$_2$ werden door de kinderen$_1$ overtuigd [PRO$_1$ ze$_2$ te mogen kietelen].
the teachers were by the children convinced [them to may tickle]
'The teachers were convinced by the children that they (the children) would be allowed to tickle them (the teachers).

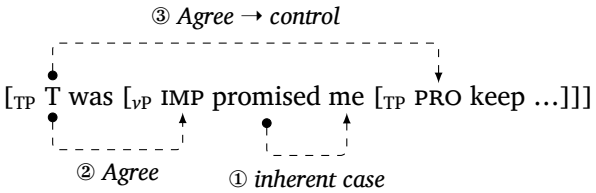
An explanation was to why Visser's Generalization, now revised in (93), should hold must also account for data like (94).

A.1 An Agree analysis of Visser's Generalization

In order to provide an explanation for (93), Van Urk proposes that Visser's Generalization is a syntactic effect that results from the failure of the operation Agree to be established between the implicit agent of the passive and PRO. Van Urk follows Landau (2001) in assuming that obligatory control is the result of Agree between the controller, a functional head *F* in the matrix clause, and PRO. Van Urk also assumes that, syntactically, implicit arguments are intransitive Ds, and that, semantically, they are bare existential quantifiers that existentially closes the argument position they are merged in. In the diagrams in (95) and (96), the implicit argument is represented as IMP.

Under these assumptions, a sentence like (92a), repeated below, is derived as follows:

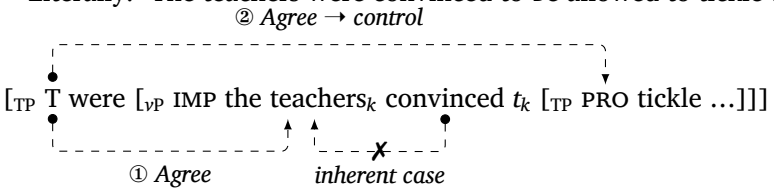
(95) Er werd mij$_2$ beloofd / aangeboden [PRO$_1$ om me$_2$ op de hoogte te houden].
there was me.DAT promised / offered INF.COMP me on the height to keep.INF
'It was promised /offered to me to keep me informed.'



In (95), the matrix functional head *F* responsible for mediating Agree between the controller IMP and PRO is T. It Agrees first with the former and then with the latter, which results in the control relationship, according to Landau's (2001) framework. Importantly, the matrix indirect object *mij* 'me' is assigned inherent case by the matrix control predicate.

A prediction that arises from this proposal is that, if the functional head *F* must Agree with a DP other than the implicit argument IMP, an Agree relationship between IMP and PRO cannot be mediated by *F* and, hence, the implicit argument cannot control PRO. Indeed this is what happens in the derivation of a sentence like (92b), where the matrix predicate cannot assign inherent to its indirect object:

(96) * De leraren$_2$ werden overtuigd [PRO$_1$ om ze$_2$ te mogen kietelen].
the teachers were convinced INF.COMP them to may.INF tickle.INF
Literally: 'The teachers were convinced to be allowed to tickle them.'



This is how Van Urk derives Visser's Generalization: the implicit agent in (96) cannot control PRO because it is not Agreed with the matrix T. The matrix verb cannot assign inherent case to its indirect object (*de leraren* 'the teachers'). As such, this DP has to rely on Agree with T in order to be licensed, so that the derivation satisfies the Case Filter.[9]

[9]For concreteness, Van Urk assumes that the indirect object shifts to the edge of the *v*P, where it is as distant as IMP is from T. Because of this configuration, Agree between T and the indirect object across IMP does not violate Minimality.

The net result is that control by the implicit agent is possible in (95) because the indirect object is assigned inherent case, freeing T to Agree with IMP, which can then control PRO (after T Agrees with PRO). In contrast, in (96), the derivation only converges if T Agrees with the indirect object, since it is not assigned inherent case. This prevents IMP from controlling PRO, since T cannot Agree with IMP.

While successful in accounting for the revised version of Visser's Generalization (93), this Agree-based analysis undergenerates the effect of an overt *by*-phrase. In (94), repeated below, the matrix verb is not able to assign inherent case to its indirect object — the control verb in (97) is the same as that in (92b), where exemplifies Visser's Generalization in its original version.

(97) *Dutch* (Wurmbrand, 2021, (5a), adapted)

De leraren$_1$ werden door de kinderen overtuigd [PRO ze$_1$ te mogen kietelen].
the teachers were by the children convinced [them to may tickle]
'The teachers were convinced by the children that they (the children) would be allowed to tickle them (the teachers).

In (97), the indirect object is assigned nominative case, just as in (92b). Yet, (97) is grammatical. Bu the derivation of both sentences would be the same in Van Urk's analysis. More precisely, T Agrees with the indirect object in order to assign case to it. T then Agrees with PRO. The result should be a sentence where the indirect object controls PRO, since T mediates Agree between both. The prediction, then, is that a sentence like (97) should be also fall under Visser's Generalization and, thus, be ungrammatical. The prediction is not borbe out by facts, as there is a grammaticality contrast between (97) and (92b).

A.2 Visser's Generalization without Agree: a WLMTC Analysis

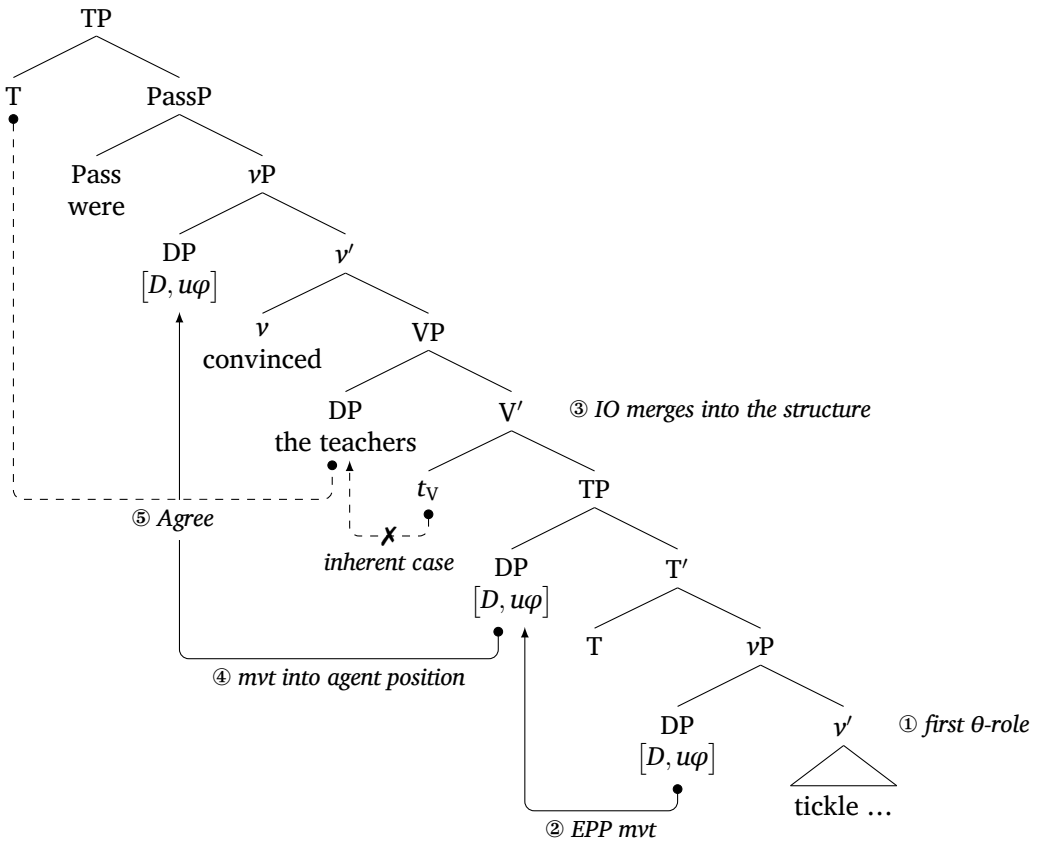
I show in this section that the theory proposed in the main portion of this paper can account for the Revised Visser's Generalization (93) and, in addition, for the *by*-phrase effect in (97).

In this paper, I proposed that obligatory control can be derived by merging a determiner alone in the embedded clause, with its NP complement being WLM-ed in the matrix controller position. There is, however, no requirement that the determiner have an NP complement. However, whether or not there is nominal material to construct an NP complement to the lone determiner, the resulting representation must comply with independent principles of the grammar. Here, the Case Filter will be relevant. I propose further that this is exactly what happens to the implicit agent of passives: $\left[D, u\varphi\right]$ is merged as the agent of the verb to be passivized, à la Legate (2014). The lone determiner is existentially bound in the absence of a *by*-phrase, also à la Legate 2014. Additionally, $\left[D, u\varphi\right]$ must be assigned case, since it is a DP. This is not an additional stipulation — rather, it is commonly assumed that DPs must be assigned case.[10]

With these assumptions in place, we can turn to the derivation of a sentence where the original Visser's Generalization is complied with, i.e. (92b).

(98) * De leraren$_2$ werden overtuigd [PRO$_1$ om ze$_2$ te mogen kietelen].
the teachers were convinced INF.COMP them to may.INF tickle.INF
Literally: 'The teachers were convinced to be allowed to tickle them.'

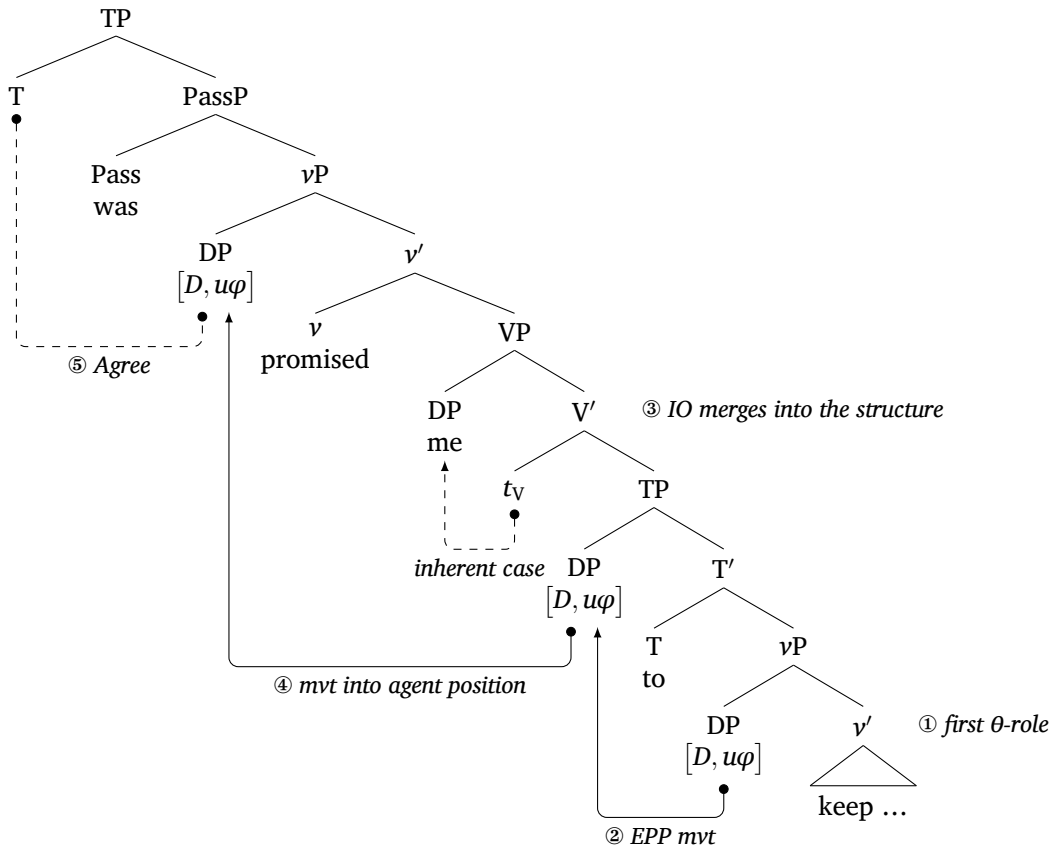
[10] In Legate (2014), the passivized predicate combines with a set of ϕ-features, which Restricts it (Chung & Ladusaw, 2003). Because the implicit agent in the sentences under investigation is a controller, I assume that it is a fully fledged DP, albeit one that is merged as a lone determiner. Being a full DP, the implicit agent must be assigned case.



Lower copies of $[D, u\varphi]$ undergo Early Trace Conversion, as proposed in this paper. The higher copy of $[D, u\varphi]$ binds the embedded subject position, resulting in obligatory control. This higher copy of $[D, u\varphi]$ is, in turn, existentially bound at *v*P level, assuming the passivization theory in Legate (2003). The derivation diagrammed in (98) results in an ungrammatical sentence due to a violation of the Case Filter. In the matrix clause, there is just one case assigner (the finite T), but two DPs that must be assigned case, viz. the implicit agent $[D, u\varphi]$ and the indirect object (*de leraren* 'the teachers'). The data that motivate Visser's Generalization, thus, reduce, to a Case Filter violation.

This analysis is also able to account for the effect that the availability of inherent case in the matrix clause has to Visser's Generalization. This is demonstrated in (99), a representation of (92a).

(99) Er werd mij$_2$ beloofd / aangeboden [PRO$_1$ om me$_2$ op de hoogte te houden].
there was me.DAT promised / offered INF.COMP me on the height to keep.INF
'It was promised /offered to me to keep me informed.'



This derivation converges because there is just one DP to be assigned case in the matrix clause, namely, the implicit agent. The indirect object does not compete to Agree with the matrix case assigner T because it is assigned inherent case. An implicit agent control construal results because the implicit agent is base-generated inside the embedded clause, where is it assigned a θ-role, prior to becoming the agent in the matrix clause. Control by the

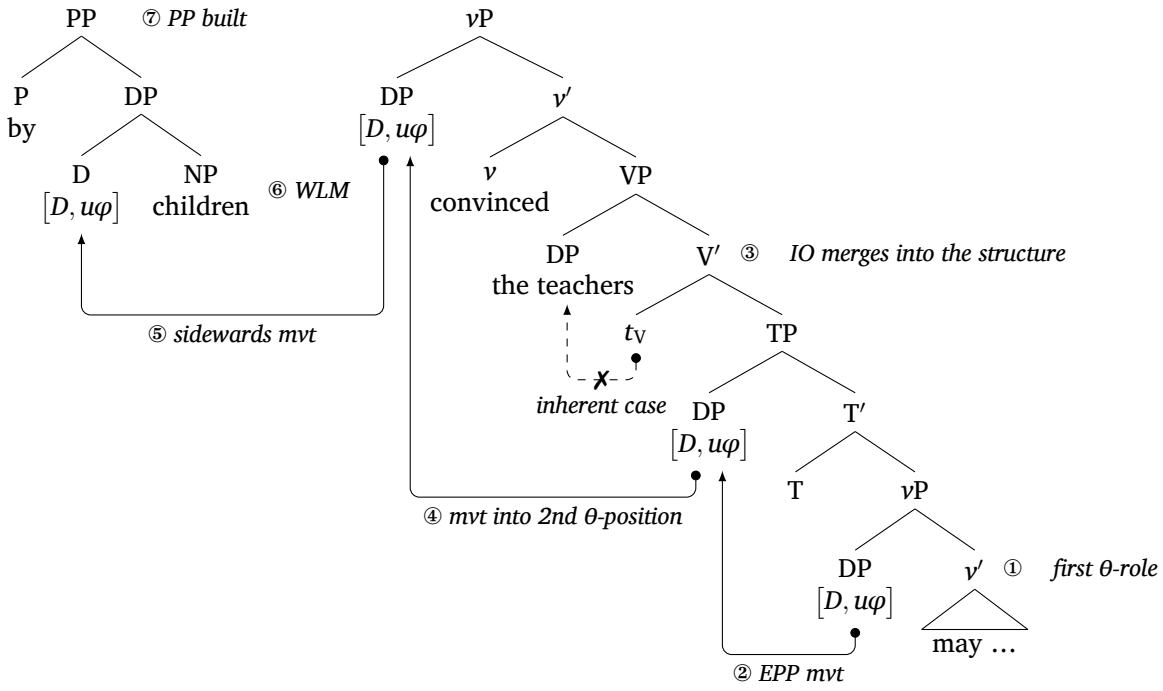
So far, the analysis, which was depicted in (98) and (99), is quite similar to that proposed in Van Urk (2013), since it is provides a syntactic explanation to Visser's Generalization that is tied to the case assignment possibilities of a given sentence. Empirically, both analyses are hitherto equivalent. They differ, however, in the ability to account for the *by*-phrase effect, noted by Wurmbrand (2021) and illustrated in (97), repeated below.

(100) *Dutch* (Wurmbrand, 2021, (5a), adapted)

De leraren$_1$ werden door de kinderen overtuigd [PRO ze$_1$ te mogen kietelen].
the teachers were by the children convinced [them to may tickle]
'The teachers were convinced by the children that they (the children) would be allowed to tickle them (the teachers).

To recall, the analysis proposed in Van Urk (2013), anchored on Landau's (2001) Agree framework, predicted that a sentence like (100) is as ungrammatical as (92b), since, in both, T Agrees with the indirect object DP. Under the WLMTC proposed here, the derivation of (100) goes as in (101). I assume that movement can be sidewards (Nunes, 2004), i.e. movement can be a dependency between structures that are not dominated by a single node.

(101)



The derivation in (101) continues by merging the *by*-phrase as an adjunct to the *v*P. Instead of being existentially bound, the *v*P combines with the *by*-phrase PP, assuming the semantics proposed by Legate (2014). The indirect object *de leraren* 'the teachers' is assigned nominative by T. $[D, u\varphi]$ must also be assigned case (as before), but it is assigned case by the preposition *by*. Control by the agent in the *by*-phrase results from a determiner merging alone in the embedded clause and then moving sidewards into the **by**-phrase in the matrix clause.

We are in the position to propose a reformulation of Visser's Generalization. I take for granted the corrected version of Van Urk's (2013) Visser's Generalization (93), though I restate it in terms of a correlation between case the case of the indirect object and the availability of subject control. Agree between a DP and T is a *derivative* property: it happens as a consequence of whether or not that DP has been assigned case, which allows it to dispense with assignment of nominative case by T.

(102) *Revised Revised Visser's Generalization*

Obligatory control by an implicit subject is impossible if an overt DP is not assigned case.

In the derivation depicted in (101), T does not Agree with T and control is still possible (unlike what happens in Van Urk's analysis). In the present analysis, Visser's Generalization is a consequence of a violation of the Case Filter: if an overt DP cannot be assigned case (lexical or structural), it ends up competing with $[D, u\varphi]$ for case. However, if that DP is assigned lexical case, T is freed up to assign nominative case to $[D, u\varphi]$, analogously to Van Urk's analysis. However, unlike what happens in the latter, in the present analysis, control by an implicit argument is not dependent on Agree with T, allowing the *by*-phrase effect to be captured.

A.3 Concluding remarks

In this brief appendix, I showed that the Wholesale Late Merge Theory of control proposed in this paper can successfully account for the data that originally motivates Visser's Generalization, the apparent counterexamples pointed out by Van Urk (2013), as well as for the *by*-phrase effect pointed out by Wurmbrand (2021). The gist of the proposal was to model the implicit argument as the lone determiner merged in the control clause and allow it to remain intransitive throughout the derivation. The existential reading usually attributed to implicit agents is afforded by existential closure (Legate, 2014). Otherwise, if a *by*-phrase is

merged into the passive structure, the lone determiner can merge with an NP complement, which is introduced via WLM, following the logic of the theory proposed here. Furthermore, I proposed a revision of Van Urk's revision of Visser's Generalization. While maintaining the core of the author's insight, I argued that the original revision in fact captured a secondary property of the purported exceptions.

References

Abney, Steven P. 1987. *The English noun phrase in its sentential aspect*. Ph.D. thesis, Massachusetts Institute of Technology. Available at: https://dspace.mit.edu/handle/1721.1/14638.

Alexiadou, Artemis, & Anagnostopoulou, Elena. 2019. An asymmetry in backward control: Subject vs. object control. *Agreement, case and locality in the nominal and verbal domains*, **5**, 55.

Alexiadou, Artemis, Anagnostopoulou, Elena, Iordachioaia, Gianina, & Marchis, Mihaela. 2010. No objections to backward control. *Movement theory of control*, 89–118.

Allotey, Deborah. 2021. Overt pronouns of infinitival predicates of Gã. *Western Papers in Linguistics*, **4**(1), 1–47.

Bobaljik, Jonathan David, & Landau, Idan. 2009. Icelandic control is not A-movement: The case from case. *Linguistic Inquiry*, **40**(1), 113–132. DOI: https://doi.org/10.1162/ling.2009.40.1.113.

Bresnan, Joan. 1982. Control and complementation. *Linguistic Inquiry*, **13**(3), 343–434.

Chomsky, Noam. 1981. *Lectures on government and binding*. Mouton de Gruyter.

Chomsky, Noam. 1995. *The minimalist program*. MIT press.

Chomsky, Noam. 2000. Minimalist inquiries: The framework. *In:* Martin, Roger, Michaels, David, & Uriagereka, Juan (eds), *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*. Cambridge, MA: MIT Press.

Chomsky, Noam. 2001. Derivation by phase. *Pages 1–52 of:* Kenstowicz, Michael (ed), *Ken Hale: A life in language*. Cambridge, MA: MIT Press.

Chung, Sandra, & Ladusaw, William. 2003. *Restriction and saturation*. Linguistic Inquiry Monographs, vol. 42. MIT Press.

Culicover, Peter W, & Jackendoff, Ray. 2001. Control is not movement. *Linguistic Inquiry*, **32**(3), 493–512. DOI: https://doi.org/10.1162/002438901750372531.

Erlewine, Michael Yoshitaka, & Gould, Isaac. 2016. Unifying Japanese relative clauses: copy-chains and context-sensitivity. *Glossa: a journal of general linguistics*, **1**(1). DOI: https://doi.org/10.5334/gjgl.174.

Ferreira, Marcelo. 2000. *Argumentos nulos em português brasileiro*. Ph.D. thesis, Universidade de Campinas.

Ferreira, Marcelo. 2009. Null subjects and finite control in Brazilian Portuguese. *In:* Nunes, Jairo (ed), *Minimalist essays on Brazilian Portuguese syntax*. John Benjamins Amsterdam.

Fong, Suzana. To appear. Pronouncing PRO in Wolof. *In: Proceedings of ACAL 53*. Language Science Press.

Fox, Danny. 2002. Antecedent-contained deletion and the copy theory of movement. *Linguistic Inquiry*, **33**(1), 63–96. DOI: https://doi.org/10.1162/002438902317382189.

Ganenkov, Dmitry. To Appear. Partial control with overt embedded subjects in Chirag. *Language*.

Georgi, Doreen, & Amaechi, Mary. 2022. Resumption in Igbo: Two types of resumptives, complex phi-mismatches, and dynamic deletion domains. *Natural Language & Linguistic Theory*, 1–68. DOI: https://doi.org/10.1007/s11049-022-09558-x.

Gong, Zhiyu Mia. 2022. Case in Wholesale Late Merger: Evidence from Mongolian Scrambling. *Linguistic Inquiry*, 1–66. DOI: https://doi.org/10.1162/ling_a_00494.

Grano, Thomas, & Lasnik, Howard. 2018. How to neutralize a finite clause boundary: Phase theory and the grammar of bound pronouns. *Linguistic Inquiry*, **49**(3), 465–499. DOI: https://doi.org/10.1162/ling_a_00279.

Grano, Thomas Angelo. 2012. *Control and restructuring at the syntax-semantics interface*. The University of Chicago.

Haddad, Youssef A. 2009. Copy Control in Telugu1. *Journal of linguistics*, **45**(1), 69–109. DOI: https://doi.org/10.1017/S0022226708005525.

Halle, Morris, & Marantz, Alec. 1993. Distributed morphology and the pieces of inflection. *Pages 111–176 of:* Hale, Kenneth, & Keyser, Samuel Jay (eds), *The view from Building 20: Essays in linguistics in honour of Sylvain Bromberger*. MIT press.

Halle, Morris, & Marantz, Alec. 1994. Some key features of Distributed Morphology. *Page 88 of:* Carnie, Andrew, Harley, Heidi, & Tony, Bures (eds), *Papers on phonology and morphology*, vol. 21. MIT press.

Heim, Irene. 2008. Features on bound pronouns. *In:* Harbour, Daniel, Adger, David, & Béjar, Susana (eds), *Phi theory: Phi-features across modules and interfaces*. Oxford University Press.

Heim, Irene, & Kratzer, Angelika. 1998. *Semantics in generative grammar*. Blackwell Oxford.

Hornstein, Norbert. 1999. Movement and control. *Linguistic Inquiry*, **30**(1), 69–96. DOI: https://doi.org/10.1162/002438999553968.

Huang, C-T James. 2022. Finiteness, opacity, and Chinese. *Page 17 of:* Simpson, Andrew (ed), *New Explorations in Chinese Theoretical Syntax: Studies in honor of Yen-Hui Audrey Li*, vol. 272. John Benjamins Publishing Company.

Kratzer, Angelika. 2009. Making a pronoun: Fake indexicals as windows into the properties of pronouns. *Linguistic Inquiry*, **40**(2), 187–237. DOI: https://doi.org/10.1162/ling.2009.40.2.187.

Kratzer, Angelika, & Heim, Irene. 1998. *Semantics in generative grammar*. Vol. 1185. Blackwell Oxford.

Landau, Idan. 2001. *Elements of control: Structure and meaning in infinitival constructions*. Vol. 51. Springer Science & Business Media.

Landau, Idan. 2003. Movement out of control. *Linguistic Inquiry*, **34**(3), 471–498. DOI: https://doi.org/10.1162/002438903322247560.

Landau, Idan. 2006. Chain resolution in hebrew V (P)-fronting. *Syntax*, **9**(1), 32–66. DOI: https://doi.org/10.1111/j.1467-9612.2006.00084.x.

Landau, Idan. 2007. Movement-resistant aspects of control. *Pages 293–325 of:* Davies, William D., & Dubinsky, Stanley (eds), *New horizons in the analysis of control and raising*. Springer.

Landau, Idan. 2013. *Control in generative grammar: A research companion*. Cambridge University Press.

Landau, Idan. 2015. *A two-tiered theory of control*. Vol. 71. MIT Press.

Lee, Felicia. 2003. Anaphoric R–expressions as bound variables. *Syntax*, **6**(1), 84–114. https://doi.org/10.1111/1467-9612.00057.

Lee, Kum Young. 2009. *Finite control in Korean*. Ph.D. thesis, The University of Iowa. DOI: https://doi.org/10.17077/etd.qgulnymt.

Legate, Julie Anne. 2003. Some interface properties of the phase. *Linguistic inquiry*, **34**(3), 506–515. DOI: https://doi.org/10.1162/ling.2003.34.3.506.

Legate, Julie Anne. 2014. *Voice and v: Lessons from Acehnese*. Vol. 69. MIT Press.

Li, Daoxin. 2021. Controlling overt subjects in Mandarin. *Proceedings of the Linguistic Society of America*, **6**(1), 303–316. DOI: https://doi.org/10.3765/plsa.v6i1.4973.

Livitz, Inna. 2014. *Deriving silence through dependent reference: Focus on pronouns*. Ph.D. thesis, New York University.

McCloskey, James. 2017. Resumption. *The Wiley Blackwell Companion to Syntax, Second Edition*, 1–30.

McFadden, Thomas, & Sundaresan, Sandhya. 2018. Reducing pro and PRO to a single source. *The Linguistic Review*, **35**(3), 463–518. DOI: https://doi.org/10.1515/tlr-2018-0003.

Modesto, Marcello. 2007. Null subjects in Brazilian Portuguese and Finnish: They are not derived by movement. *Pages 231–248 of: New horizons in the analysis of control and raising*. Springer.

Modesto, Marcello. 2011. Finite control: Where movement goes wrong in Brazilian Portuguese. *Journal of Portuguese Linguistics*, **10**(2). DOI: https://doi.org/10.5334/jpl.95.

Ndayiragije, Juvénal. 2012. On raising out of control. *Linguistic Inquiry*, **43**(1), 275–299. DOI: https://doi.org/10.1162/LING_a_00086.

Nunes, Jairo. 2004. *Linearization of chains and sideward movement*. Vol. 43. MIT Press.

Ostrove, Jason. To Appear. Obligatorily overt PRO in San Martın Peras Mixtec. *Linguistic Inquiry*.

Pietraszko, Asia. 2021. Backward Control without A-movement or φ-agreement. *In: Proceedings of NELS 51*.

Polinsky, Maria, & Potsdam, Eric. 2002. Backward control. *Linguistic Inquiry*, **33**(2), 245–282. DOI: https://doi.org/10.1162/002438902317406713.

Potsdam, Eric, & Polinsky, Maria. 2012. Backward raising. *Syntax*, **15**(1), 75–108. DOI: https://doi.org/10.1111/j.1467-9612.2011.00158.x.

Preminger, Omer. 2011. *Agreement as a fallible operation*. Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, MA. Available at: http://dspace.mit.edu/handle/1721.1/7582.
Safir, Ken. 2014. One True Anaphor. *Linguistic Inquiry*, **45**(1), 91–124. DOI: https://doi.org/10.1162/LING_a_00149.
Safir, Ken. 2019. The A/Ā distinction as an epiphenomenon. *Linguistic Inquiry*, **50**(2), 285–336. DOI: ttps://doi.org/10.1162/ling_a_00305.
Sheehan, Michelle, & Cyrino, Sonia. 2023. Restrictions on long passives in English and Brazilian Portuguese: a phase-based account. *Linguistic Inquiry*, 1–35. DOI: https://doi.org/10.1162/ling_a_00482.
Stanton, Juliet. 2016. Wholesale late merger in Ā-movement: Evidence from preposition stranding. *Linguistic Inquiry*, **47**(1), 89–126. DOI: https://doi.org/10.1162/LING_a_00205.
Sulemana, Abdul-Razak. 2018. PRO as a Partial Copy: a movement-based account of Control. *Pages 493–502 of: Proceedings from the Annual Meeting of the Chicago Linguistic Society*, vol. 54. Chicago Linguistic Society.
Sulemana, Abdul-Razak. 2021. *Non-finite Complementation: A case study of Bùlì*. Ph.D. thesis, Massachusetts Institute of Technology. Available at: https://dspace.mit.edu/handle/1721.1/139142.
Sulemana, Abdul-Razak. 2023. *Reduced complements*. Ms., University of Ghana.
Sundaresan, Sandhya. 2010. A phase-based account of the PRO/anaphor distinction. *Proceedings of ConSOLE XVIII*, **1**, 19.
Takahashi, Shoichi. 2019. Putative null pronominals in English: Causes and consequences. *Studia Linguistica*, **73**(3), 650–682. DOI: https://doi.org/10.1111/stul.12114.
Takahashi, Shoichi, & Hulsey, Sarah. 2009. Wholesale late merger: Beyond the A/Ā distinction. *Linguistic Inquiry*, **40**(3), 387–426. DOI: https://doi.org/10.1162/ling.2009.40.3.387.
Van Urk, Coppe. 2010. *On obligatory control: A movement and PRO approach*. Ms., Utrecht University.
Van Urk, Coppe. 2013. Visser's Generalization: The syntax of control and the passive. *Linguistic Inquiry*, **44**(1), 168–178. DOI: https://doi.org/10.1162/LING_a_00124.
Van Urk, Coppe. 2018. Pronoun copying in Dinka Bor and the copy theory of movement. *Natural Language & Linguistic Theory*, **36**(3), 937–990. DOI: https://doi.org/10.1007/s11049-017-9384-x.
Wood, Jim. 2012. Against the movement theory of control: Another argument from Icelandic. *Linguistic Inquiry*, **43**(2), 322–330. DOI: https://doi.org/10.1162/LING_a_00089.
Wood, Jim. 2017. Icelandic object extraposition is still a problem for the Movement Theory of Control: A reply to Drummond and Hornstein. *Linguistic Inquiry*, **48**(3), 513–527. DOI: https://doi.org/10.1162/ling_a_00252.
Wurmbrand, Susi. 1998. *Infinitives*. Ph.D. thesis, Massachusetts Institute of Technology. Available at: https://dspace.mit.edu/handle/1721.1/9592.
Wurmbrand, Susi. 2021. Rethinking implicit control. *In:* Douglas, Jamie, Biberauer, Theresa, Vikner, Sten, & Bárány, András (eds), *Syntactic architecture and its consequences III: Inside syntax*, vol. 11. Language Science Press.
Yip, Ka-Fai, & Ahenkorah, Comfort. 2023. Non-agreeing Resumptive Pronouns and Partial Copy Deletion. *University of Pennsylvania Working Papers in Linguistics*, **29**(1), 24.
Zribi-Hertz, Anne, & Diagne, Lamine. 2002. Clitic placement after syntax: Evidence from Wolof person and locative markers. *Natural Language & Linguistic Theory*, **20**(4), 823–884. DOI: https://doi.org/10.1023/A:1020494714861.
Zyman, Erik. 2023. Raising out of finite clauses (hyperraising). *Annual Review of Linguistics*, **9**, 29–48. DOI: https://doi.org/10.1146/annurev-linguistics-022421-070658.