

Subject co-reference in antecedent-contained deletion*

Luise Schwarzer

Leipzig University

1. Introduction

English and Dutch both allow the ellipsis of a verbal projection. In contrast to English VP-ellipsis (1), Dutch shows ellipsis of a larger verbal projection, and only allows it in the context of a modal (Aelbrecht 2010). This construction has been called modal complement ellipsis (MCE) and is illustrated in (2).

- (1) *English VP-ellipsis*
Kirsten ate a whole bag of chips but Marina didn't [_{VP} eat a whole bag of chips]
- (2) *Dutch modal complement ellipsis*
Erik is al langsgesproken, maar Jenneke moet nog [_{VOICEP} langsgesproken].
Erik is already by.passed but Jenneke must still
'Erik has already passed by, but Jenneke still has to.' (Aelbrecht 2010:61)

Aelbrecht (2010:139) observes a curious restriction of modal complement ellipsis in antecedent-contained deletion contexts: Dutch only allows antecedent-contained deletion (ACD) with co-referent subjects, (3).¹ Such a restriction is surprising, since it is found neither in English antecedent-contained deletion, (4), nor in modal complement ellipsis in general, (2).

*Thanks to audiences at NELS53, CGSW36, GGS47, and the syntax colloquium at Leipzig University for discussing various versions of this paper with me and giving me valuable feedback, especially Jason Overfelt, Andy Murphy, Martin Salzmann, Gary Thoms. I am grateful to Matthew Cummins, Paula Fenger, Thom van Hugte, Ben L. Sluckin, and Gary Thoms for providing judgments. All errors are entirely my own.

¹The restriction also seems to hold in comparative deletion, (i).

- (i) Maja_i zou meer sap moeten drinken dan {ze_i/ ?*Hans} wilde ~~saap drinken~~
Maja would more juice must drink than she Hans wants
'Maja has to drink more juice than ✓she/ *Hans wants to drink.' (P.Fenger, p.c.)

- (3) *Subject-co-reference in antecedent contained deletion*
 Olaf_i heeft elk boek gelezen dat {hij_i/ *David} moest.
 Olaf has every book read that he David must.PST
 ‘Olaf read every book that he had to.’ (Aelbrecht 2010:139)
- (4) a. Sue likes every boy that Mary does [_{VP} like *t*].
 b. Mary has read every book that June couldn’t [_{VP} read *t*].

This paper aims to explain the subject co-reference constraint in Dutch ACD as a *bound pronoun effect* (Grano and Lasnik 2018, see also Barros and Frank 2022). I propose that generally, a mismatch between the size of the ellipsis site (high verbal projection) and the canonical landing site for quantifier raising (QR; low verbal projection) makes ACD impossible. ACD can only be licensed if QR can exceptionally reach a high verbal projection. Following Grano & Lasnik (2018), a bound pronoun subject can extend the locality domain, thereby allowing such exceptional movement. Thus, I correlate the presence of obligatory subject co-reference with the interaction of (i) the size of the ellipsis site, and (ii) the length of QR. If these independent properties are mismatched in a certain way, ACD can only be licensed in the presence of a bound pronoun. I make the descriptive generalization in (5).

- (5) *Subject co-reference generalization*
 If an elided phrase contains the canonical landing site for QR in a given language L, L can only allow ACD with a bound pronoun subject.

The paper is structured as follows: section 2 provides a background on modal complement ellipsis, antecedent-contained deletion, and quantifier raising. Section 3 discusses the bound pronoun effect. The analysis is developed in section 4. Section 5 concludes.

2. Background

2.1 Modal complement ellipsis

While English shows the well-known VP-ellipsis, in languages like Dutch a verbal projection can only be elided if it is the complement of a modal verb (but not as the complement of a temporal or voice auxiliary). This construction is called modal complement ellipsis in Aelbrecht (2010). Aelbrecht shows that modals are verbs that embed a TP out of which the subject obligatorily raises. The modal verb can trigger deletion of a voice or aspect projection inside the embedded TP, as in (6) (Aelbrecht 2010:54f.). In general, as illustrated in (2), the subject in the antecedent conjunct and the subject in the elliptical clause may be non-identical in modal complement ellipsis (although see Sauerland 2017).

- (6) [_{TP} ... subject ... modal]_[E] [_{TP} *t*_{subj} [_{VoiceP} ... [_{VP} ... [_{VP} ...]]] T]]

2.2 Antecedent-contained deletion

Ellipsis cannot occur spontaneously, but is restricted to contexts in which there is a suitable antecedent for the elided phrase, see (7). The antecedent and the ellipsis site must be sufficiently parallel to each other in order for ellipsis to be grammatical. To which degree this parallelism must be syntactic or semantic is debated (see e.g., Kroll 2019, Rudin 2019).

- (7) Mary was [_{VP} kissing my aunt] and Sue was [_{VP} ~~kissing my aunt~~] too.
antecedent *ellipsis site*

Sometimes the antecedent (= VP₁ in (8)) *contains* the ellipsis site (= VP₂). This is referred to as antecedent-contained deletion (ACD). Such configurations pose challenges for the theory of ellipsis since the otherwise obligatory parallelism requirement cannot be trivially met. Simply copying the antecedent VP into the position of the elided VP would result in an infinite regress problem. This means that ellipsis should not be possible, contrary to fact.

- (8) Sue [_{VP₁} like every boy *Op_k* that Mary does [_{VP₂} ~~like *t_k*~~]]

The standard solution is that parallelism can be created derivationally via Quantifier Raising (QR, e.g., May 1985, Kennedy 1997). The quantified object DP, including the ellipsis site in the relative clause, is moved out of the antecedent VP. It raises and adjoins to the VP, (9), resulting in parallel VPs that each contain the verb and a trace. This analysis is particularly elegant since quantifier raising is independently needed to interpret quantified objects (Heim and Kratzer 1998).²

- (9) [_{VP} Sue [_{VP₁} [_{DP} every boy *Op* that Mary does [_{VP₂} like *t*]]]_j [_{VP₁} like *t_j*]]
-

2.3 Quantifier raising and scope rigidity

English differs from languages like Dutch in that it allows inverse scope readings quite freely. Dutch is scope-rigid (Zwart 1993, De Hoop and Krämer 2006), meaning doubly quantified sentences like (10-b) are not ambiguous. It is traditionally assumed that inverse scope is brought about by covertly raising one quantifier above the other (May 1985).

²As noted in the literature, there are certain issues with the QR account (see e.g., Fox 2000, 2002, Charlow 2014 for discussion and alternatives). As will become clear below, the analysis developed here and its inspiration in Grano & Lasnik (2018) are only compatible with the QR-approach to ACD. In the alternative based on extraposition and late merger, only the head of the relative clause undergoes movement to derive parallelism, the relative clause is adjoined counter-cyclically in the final landing site. In such a configuration, a pronoun inside the relative clause comes too late to have any influence on the locality of movement. An extraposition account cannot easily predict the Dutch pattern: unlike QR, extraposition can in principle reach vP, which would create parallelism and license ellipsis even without bound pronouns.

- (10) a. A shark nibbled on every pirate. ($\exists > \forall, \forall > \exists$)
 b. Een haai heeft elke piraat gebeten.
 a shark has every pirate bitten ($\exists > \forall, *\forall > \exists$, P.Fenger, p.c.)

The conditions that make inverse scope interpretations possible in scope-rigid languages are subject of much empirical and theoretical debate and cannot be covered adequately here. I will only discuss the core assumptions that I make for the analysis of subject co-reference in ACD. I assume with Heim and Kratzer (1998), Abels and Grabska (2022) that QR is in principle obligatory and also happens in scope-rigid languages, but is more restricted in these languages: QR obligatorily lands in a position beneath the subject, i.e., it adheres to the Isomorphic Principle, (11) (Aoun and Li 1993). I assume that in languages like English, (11) can be violated, leading to inverse scope.

- (11) *The Isomorphic Principle* (Aoun and Li 1993:15)
 Suppose A and B are Quantifier Phrases. Then if A c-commands B at S(urface)-Structure, A c-commands B at LF.

3. The bound pronoun effect

Grano and Lasnik (2018) observe that a bound pronoun in the subject of an embedded clause can render the clause boundary transparent. A number of phenomena that are generally clause-bound can exceptionally cross clause boundaries if the embedded subject is bound by the matrix subject (e.g., *tough*-movement, multiple sluicing, ACD). They call this extension of locality the *bound pronoun effect*. Take for example English ACD. In general, QR may not cross a finite clause boundary, (12-b). However, if the embedded subject is a pronoun bound by the matrix subject, the clause boundary is significantly weakened, (12-c). The central claim of this paper is that Dutch ACD belongs to the same class of phenomena as exceptionally cross-clausal dependencies in English. Grano & Lasnik's unified account for such extensions of the locality domain is given in (13).

- (12) a. John reads everything Bill does read.
 b. *John claims that Mark reads everything Bill does elaim [~~that Mark reads~~].
 c. ?John₁ claims that he₁ reads everything Bill₂ does elaim [~~that he₂ reads~~].
 (Grano & Lasnik 2018:466f.)

- (13) *Grano & Lasnik's (2018:482) account of the Bound Pronoun Effect*
 a. Unvalued features on the head of the complement to the phase head keep the phase open.
 b. The locality domain for the phenomena that give rise to the bound pronoun effect is the phase.
 c. Bound pronouns optionally enter the derivation with unvalued ϕ -features.

enough to create parallelism between antecedent and ellipsis site. While the ellipsis site contains the traces of both the object and the subject, the antecedent VP that is created by evacuating QR of the object only contains that trace, (15-b). ACD with a non-pronoun subject in Dutch is (correctly) ungrammatical in this analysis.

(15) *The ungrammatical case: non-bound-pronoun subject*

- a. *Olaf heeft [_{VP} t elk boek gelezen dat David moest [_{VP} t t lezen]]
 Olaf has every book read that David must.PST read
- b. [_{VP} [_{DP} elk boek dat David moest [_{VP} ⟨subj⟩ ⟨obj⟩ lezen]] [_{VP} ⟨obj⟩ lezen]]

Parallelism would require adjunction to the phrase that is elided, i.e., VoiceP or vP, which is usually ruled out. I claim that a bound pronoun as the relative clause subject can extend the locality domain in the main clause such that QR to VoiceP/vP becomes possible. The result is that ACD can only be licensed if there is a co-referent pronoun subject.

In the configuration in which ACD is grammatical, (16), there is a pronoun with unvalued ϕ -features inside the relative clause on the object. An additional assumption I have to make is that the unvalued feature can percolate up the relative clause and be visible on the DP. The quantified object undergoes standard QR and adjoins to matrix VP. At this point, the value provider for the pronoun, the matrix subject, has not been merged yet. This allows QR to go on and adjoin to the next higher projection, vP.³ vP is the projection in which the subject is merged. The subject binds the pronoun, transmitting a feature value to it.⁴ Incidentally, QR has now also progressed high enough to create a parallel vP antecedent for modal complement ellipsis.⁵

³Aelbrecht (2010:54) analyzes modal complement ellipsis as ellipsis of VoiceP. The evidence for this comes from the obligatory deletion of voice auxiliaries. For simplicity, I assume that voice is realized in vP and that vP is elided. If one assumes that the external argument is introduced in VoiceP, as in e.g., Legate 2014 Aelbrecht's analysis can be adopted as is (with the provision in fn. 4). There is only a problem for my proposal if the subject is merged in a lower projection than is targeted for ellipsis, as in (i). Then, QR could proceed to adjoin to vP where the pronoun could receive a value from the subject. This position is not high enough to create parallelism for VoiceP-ellipsis and ACD would be incorrectly ruled out.

(i) [_{VP} modal_[E] [_{TP} T [_{VoiceP} voice auxiliary [_{VP} subject [_{VP} ...]]]]]

⁴In Grano & Lasnik's approach, feature transmission would also re-instate the phase boundary. I assume that a phase boundary can never be reactivated. Instead, movement processes can go on until the next higher phase. This is necessary to allow additional QR to cross the subject. Without this assumption, the subject agrees with the pronoun in its c-command domain, i.e., after the first step of QR. That means the phase would be dissolved *and* re-instated at VP, prohibiting further QR. Another technical way of ensuring additional QR is to make use of workspaces: after adjunction to VP, the unvalued feature could initiate further movement by moving to a workspace. This means that the feature is not in the c-command domain when the subject is merged in Spec,vP. After re-merge of the object at vP, the subject move to Spec,TP. From there, it can agree with the pronoun, now again in its c-command domain and close off the phase in Grano & Lasnik's sense.

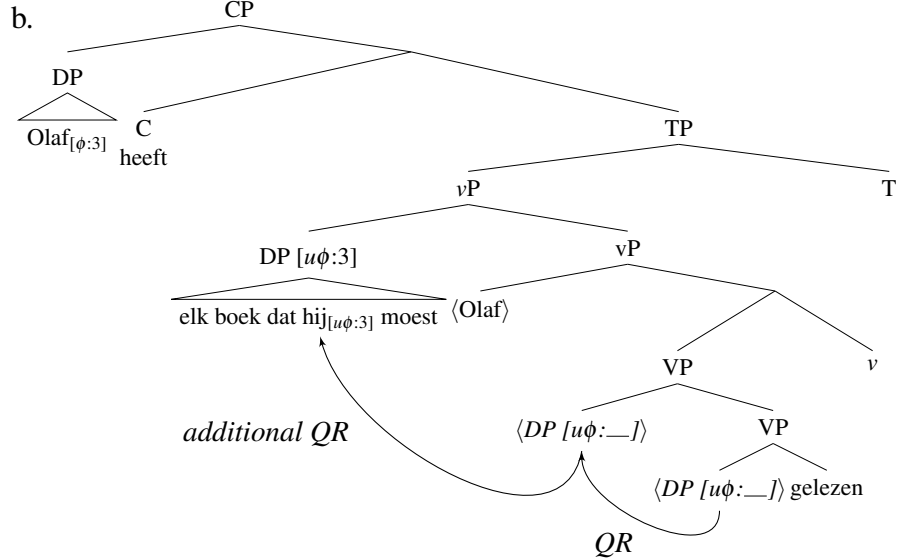
⁵Even though the relative clause is extraposed in (15a), this is not obligatory in ACD in Dutch, see (i).

(i) Olaf heeft elk boek dat hij moest lezen, gelezen.
 Olaf has every book that he must read read
 'Olaf has read every book that he had to.' (Th. van Hugte, p.c.)

Subject co-reference in ACD

(16) *The grammatical case: bound pronoun*

- a. Olaf_i heeft elk boek gelezen dat hij_i moest.
 Olaf has every book read that he must.PST
 ‘Olaf read every book that he had to.’



- c. [_{vP} *t_{subj}* *t_{obj}* lezen]
 d. [_{VP} *t_{obj}* lezen]
 e. [_{vP} *t_{subj}* *t_{obj}* lezen]

elided phrase
non-parallel antecedent after first QR
parallel antecedent after additional QR

Grano & Lasnik’s (2018) analysis of cross-clausal ACD in English is not completely parallel to what I propose here. In my proposal, the unvalued feature stems from a subject, but percolates up the relative clause and has its ultimate effect as the *object* of the main clause. Such a bound-object-pronoun effect is explicitly ruled out in Grano & Lasnik (2018): an unvalued feature on the object should not be close enough to the only phase head that they assume, i.e., C, to extend the phase. So at this point, the question of the right locality domain for QR returns. If we follow Grano & Lasnik’s mechanism closely, we would need to assume that additionally to CP, VP is a phase in Dutch, which is an issue I leave for future research for now. If we assume VP to be a phase, their account can be directly adopted: an unvalued feature on the head (D) of the complement to the phase head (V) can extend the phase. However, as discussed in section 3, it is not clear that QR is restricted by phases at all. Nonetheless, there must be some mechanism that restricts QR in some languages such that a quantified object cannot raise above the subject, thereby leading to surface scope readings in neutral settings. Grano & Lasnik’s phase-based analysis may not be the correct implementation. However, I want to argue that obligatory subject co-reference in Dutch ACD belongs to the same family of phenomena discussed in Grano & Lasnik (2018) that show a bound pronoun effect. Whatever derives cross-clausal ACD in English should also derive subject co-reference in languages like Dutch.

As for English ACD, there is no bound pronoun requirement since the ellipsis site and the landing site for QR coincide (VP). With VP-ellipsis, the minimal QR to VP already creates a parallel antecedent, compare (9) above.

4.2 Implications

A different approach to bound pronoun phenomena has been put forward by Barros and Frank (2022). They propose a discourse-based account in which clause-boundedness only holds if the subjects refer to *different* salient referents, i.e., if the subjects cause an *Attention Shift*. If the embedded subject does not require a shift of attention, i.e., when it is non-referential or anaphoric, processes should be able to cross the clause boundary in their account. The speakers I have consulted report that ACD is only possible if a co-referent pronoun is involved. If the subjects are co-referent but non-pronominal, the structure is not as acceptable, (17-a). Non-referential subjects do not allow ACD either, (17-b)

- (17) a. *?Olaf_i heeft elk boek gelezen dat {Olaf_i/ die idioot_j} wilde.
 Olaf has every book read that Olaf the idiot wanted
- b. ??Olaf heeft de boeken gelezen die geen student/ niemand mocht.
 Olaf has the books read that no student no.one was.allowed.to
 (Th. v. Hughte, P. Fenger, p.c.)

The view of subject co-reference argued for here receives additional support from British *do*-ellipsis. British *do*-ellipsis, illustrated in (18), has been argued to involve a bigger ellipsis site than VP-ellipsis, namely *vP* (e.g., Lewis To appear). It also contrasts with VP-ellipsis in that it disallows inverse scope (Thoms and Sailor 2018), (19).

- (18) Tom should write a paper and Emma should do too. (Lewis To appear)
- (19) a. Some man will read every book, and some woman will too. $\forall > \exists$
 b. Some man will read every book, and some woman will do too. $*\forall > \exists$
 (Thoms and Sailor 2018)

So far, British *do*-ellipsis looks a lot like Dutch MCE: the ellipsis site is *big*, whereas the landing site for QR is *small* (leading to surface scope readings). Under this analysis, we would expect British *do*-ellipsis to behave like Dutch MCE in antecedent-contained deletion in that it is only possible with a bound subject, which seems to be the case, (20).

- (20) a. Jane_i will read every book that {she_i/ *Sue} should do.
 (G. Thoms, B. L Sluckin, p.c.)

The analysis also predicts that inverse scope readings should become available in ACD, since the bound pronoun allows for an additional step of QR, crossing the subject. This, however, does not seem to be borne out, (21).

Subject co-reference in ACD

- (21) a. At least one student_i will read every book that they_i should do. * $[\forall > \exists]$
b. At least one student_i will read every book that they_i should. * $[\forall > \exists]$
(G. Thoms, p.c.)

There could be many reasons for this. First, after exceptionally high QR of the object, the subject undergoes additional A-movement to Spec,TP, across the quantified object, destroying the inverse configuration. Second, even if EPP-movement did not count for scope or the subject had to reconstruct, scope parallelism (Hirschbühler 1982, Fox 2000) would force parallel scope interpretations in antecedent and ellipsis site. Since there is no additional QR that would lead to an inverse reading in the ellipsis site, the only possible interpretation in ACD is surface scope.

Lastly, the analysis makes cross-linguistic predictions. I argue that the presence of obligatory subject co-reference can be reduced to two independent properties: (i) the size of the elided phrase, and (ii) the canonical length of QR. In languages where these parameters mismatch such that the elided phrase is larger than the adjunction site of QR, ACD should only be possible if there is a bound pronoun subject in the relative clause. Candidates for this type of language are French, Italian, and Spanish, which all pattern like Dutch and British English (with *do*-ellipsis; see Dagnac 2010). Languages where the two relevant phrases are identical are predicted to show no co-reference restriction. This is instantiated in English with VP-ellipsis, where both phrases are VPs, and potentially Czech, where both phrases are *v*P/VoiceP (Gruet-Skrabalova 2020).

5. Conclusion

Antecedent-contained deletion has a curious restriction in languages like Dutch: it is only possible if the subject of the embedded clause both clauses is a pronoun that is co-referent with the matrix subject. I propose that this phenomenon can be viewed as a bound pronoun effect: a bound pronoun subject in an embedded clause can extend the locality domain for QR that is necessary to license ellipsis. The analysis makes clear cross-linguistic predictions which seem to be borne out in British *do*-ellipsis, Romance, and Czech. This proposal supports the view that QR cannot apply as high as necessary to license ACD (Overfelt 2020, contra Cecchetto 2004), but instead has to be triggered by something else, like a bound pronoun, and can license ACD as a by-product.

References

- Abels, Klaus, and Dagmara Grabska. 2022. On the distribution of scope ambiguities in Polish. *Glossa* 7:1–44.
- Aelbrecht, Lobke. 2010. *The Syntactic Licensing of Ellipsis*. Amsterdam: John Benjamins.
- Aoun, Joseph, and Yen-hui Audrey Li. 1993. *Syntax of Scope*. Cambridge: MIT Press.
- Barros, Matthew, and Robert Frank. 2022. Attention and locality: On clause-boundedness and its exceptions in multiple sluicing. *Linguistic Inquiry* 1–69.

- Cecchetto, Carlo. 2004. Reconstruction in relative clauses and the copy theory of traces. In *Linguistic Variation Yearbook*, ed. by Pierre Pica, Johan Rooryck, and Jeroen van Craenenbroeck, 3–35.
- Charlow, Simon. 2014. On the semantics of exceptional scope. Doctoral dissertation, New York University.
- Dagnac, Anne. 2010. Modal ellipsis in French, Spanish and Italian. In *Romance linguistics 2008. interactions in romance*, ed. by Karlos Arregi et al., 157–170.
- De Hoop, Helen, and Irene Krämer. 2006. Children's optimal interpretations of indefinite subjects and objects. *Language Acquisition* 13:103–123.
- Fox, Danny. 2000. *Economy and Semantic Interpretation*. Cambridge: MIT Press.
- Fox, Danny. 2002. Antecedent-contained deletion and the copy theory of movement. *Linguistic Inquiry* 33:63–96.
- Grano, Thomas, and Howard Lasnik. 2018. How to neutralize a finite clause boundary: Phase theory and the grammar of bound pronouns. *Linguistic Inquiry* 49:465–499.
- Gruet-Skrabalova, Hannah. 2020. Czech modal complement ellipsis from a comparative perspective. In *Advances in Formal Slavic Linguistics 2017*, ed. by Franc Marušič, Petra Mišmaš, and Rok Žaucer, 97–122.
- Heim, Irene, and Angelika Kratzer. 1998. *Semantics in Generative Grammar*. Oxford: Blackwell.
- Hirschbühler, Paul. 1982. VP-deletion and across-the board quantifier scope. In *Proceedings of NELS 12*, 132–139.
- Huang, C-T. James. 1995. Logical form. In *Government and Binding Theory and the Minimalist Program*, ed. by Gert Webelhuth, 125–176. Hoboken: Wiley-Blackwell.
- Kennedy, Christopher. 1997. Antecedent-contained deletion and the syntax of quantification. *Linguistic Inquiry* 28:662–688.
- Kroll, Margaret. 2019. Polarity reversals under sluicing. *Semantics and Pragmatics* 12:1–49.
- Legate, Julie Anne. 2014. *Voice and v: Lessons from Acehnese*. Cambridge: MIT Press.
- Lewis, Beccy. To appear. British English *do*-ellipsis: An argument for phasal ellipsis. In *Proceedings of WCCFL 40*.
- May, Robert. 1985. *Logical form*. Cambridge: MIT Press.
- Overfelt, Jason. 2020. A remark on the economics of Quantifier Raising. *Linguistic Inquiry* 51:366–394.
- Rudin, Deniz. 2019. Head-based syntactic identity in sluicing. *Linguistic Inquiry* 50:253–283.
- Sauerland, Uli. 2017. On antecedent contained ellipsis in Continental West Germanic. *Wiener Linguistische Gazette* 82:263–270.
- Thoms, Gary, and Craig Sailor. 2018. When silence gets in the way: Extraction from *do*-ellipsis in British dialects. In *Proceedings of NELS 48*, 145–155.
- Zwart, Jan-Wouter. 1993. Dutch syntax: A minimalist approach. Doctoral dissertation, Rijksuniversiteit Groningen.

Luise Schwarzer

marie-luise.schwarzer@uni-leipzig.de