Why and how to differentiate complement raising from subject raising in Dutch

Frank Van Eynde
Centre for Computational Linguistics, University of Leuven

Liesbeth Augustinus
Centre for Computational Linguistics, University of Leuven

Proceedings of the 20th International Conference on
Head-Driven Phrase Structure Grammar

Freie Universität Berlin

Stefan Müller (Editor)

2013

Stanford, CA: CSLI Publications

pages 222–242

Van Eynde, Frank & Liesbeth Augustinus. 2013. Why and how to differentiate complement raising from subject raising in Dutch. In Stefan Müller (ed.), *Proceedings of the 20th International Conference on Head-Driven Phrase Structure Grammar, Freie Universität Berlin*, 222–242. Stanford, CA: CSLI Publications.
DOI: 10.21248/hpsg.2013.12.

Abstract

In Dutch V-final clauses the verbs tend to form a cluster in which the main verb is separated from its syntactic arguments by one or more other verbs. In HPSG the link between the main verb and its arguments is canonically modeled in terms of *argument inheritance*, also known as *argument composition* or *generalized raising*. When applied to Dutch, this treatment yields a number of problems, making incorrect predictions about the interaction with the binding principles and the passive lexical rule. To repair them this paper proposes an alternative, in which subject raising and complement raising are modeled in terms of different devices. More specifically, while subject raising is modeled in terms of lexical constraints, as for English, complement raising is modeled in terms of a more general constraint on headed phrases. This new constraint not only accounts for complement raising out of verbal complements, it also deals with complement raising out of adjectival and adpositional complements, as well as with complement raising out of PP adjuncts and subject NPs. It is, hence, a rather powerful device. To prevent overgeneration we add a number of constraints. For Dutch, the relevant constraints block complement raising out of CPs, V-initial VPs and P-initial PPs. For English, the Empty COMPS Constraint is sufficient to block complement raising entirely.

1 Introduction

In Dutch V-final clauses with more than one verb, the verbs tend to form a cluster, as in (1).

(1) ... of Peter het boek *zal kunnen vinden*.
... if Peter the book will can find
'... if Peter will be able to find the book.'

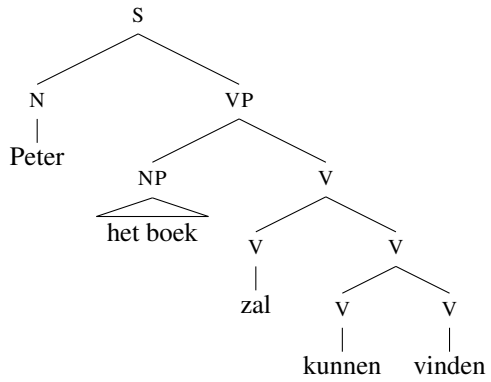
The result of this clustering is that the main verb, i.c. *vinden* 'find', is separated from its syntactic arguments by other verbs, i.c. the future *zal* 'will' and the modal *kunnen* 'can'. To model the relation between the main verb and its arguments, Head-driven Phrase Structure Grammar canonically employs the device of *argument inheritance*, also known as *argument composition* or *generalized raising* (Hinrichs & Nakazawa, 1989, 1994). In this paper we first present the Argument Inheritance analysis and apply it to (1) (section 2). Then, we show that the application to Dutch yields a number of problems (section 3), and we propose an alternative (section 4).

[†]We thank the audiences of the Workshop on Structure and Evidence in Linguistics, better known as the Ivan Sag Fest (Stanford, April 28-30), and of the HPSG 2013 conference (Berlin, August 28-29) for their comments. The research presented in this paper is part of a project on complement raising and cluster formation in Dutch, sponsored by FWO Vlaanderen (2011-2015, G.0.559.11.N.10).

2 Argument inheritance

Building on a GPSG proposal in Johnson (1986), Hinrichs & Nakazawa (1989) argue that the German verb cluster is a constituent with a binary branching structure to which the arguments of the main verb are added one at a time. Applying this to the Dutch construction in (1) yields the phrase structure in (2).

(2)



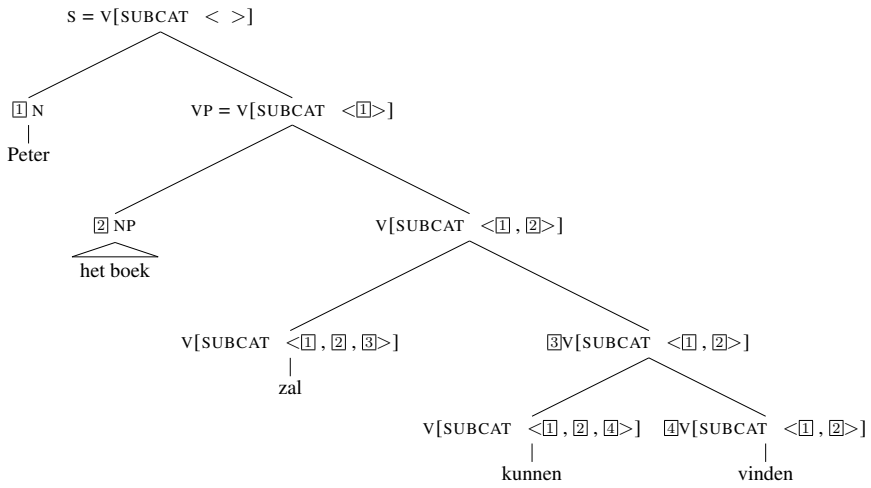
The relation between the main verb and its arguments is modeled in terms of the SUBCAT(EGORIZATION) values of the verbs. The one of the main verb *vinden* 'find' is a list which contains two NPs. The SUBCAT values of the other verbs are more complex: They take a verbal complement as their most oblique argument, and inherit the SUBCAT list of that verbal complement, as in (3), after Hinrichs & Nakazawa (1994).

(3)

$$\left[\text{SUBCAT}\ \boxed{A} \oplus \left\langle \left[\text{LOCAL}\,|\,\text{CAT}\left[\begin{array}{l}\text{HEAD}\ \ \textit{verb}\\ \text{SUBCAT}\ \ \boxed{A}\end{array}\right]\right]\right\rangle\right]$$

Adding this information to (2) yields (4).

(4)



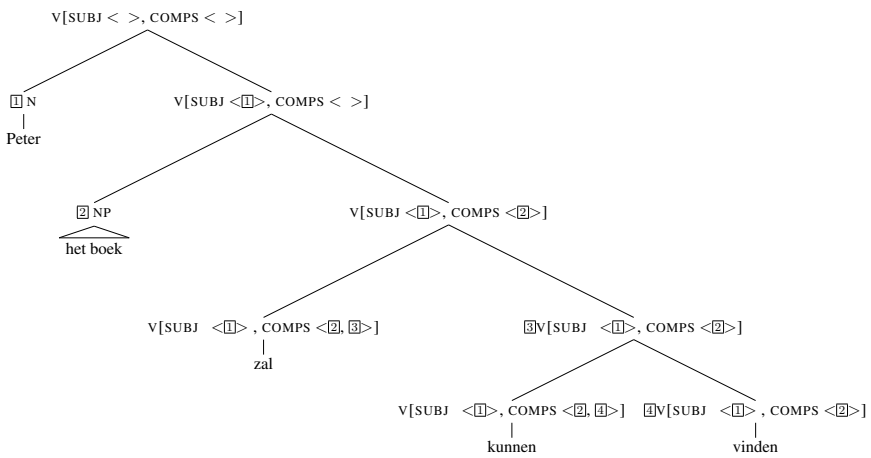
The modal *kunnen* 'can' selects the bare infinitive *vinden* 'find' (4) as well as the arguments of that infinitive (1 and 2), and the resulting cluster (3) is selected by the future *zal* 'will', which also inherits the arguments of the cluster. The net result is that the combination *zal kunnen vinden* 'will be able to find' has the same SUBCAT list as *vinden* 'find'. At that point, the direct object and the subject are added and the corresponding SUBCAT requirements in the verbal projection are discarded in the usual way.

In more recent versions of HPSG, SUBCAT is replaced with the valence features SUBJ(ECT) and COMP(LEMENT)S, on the one hand (Pollard & Sag, 1994, chapter 9),[1] and by the lexical ARG-ST feature, on the other hand (Miller & Sag, 1997). Expressing argument inheritance in this notation yields (5).

(5)
$$\left[\text{ARG-ST}\ \left\langle \boxed{1} \right\rangle \oplus \boxed{A} \oplus \left\langle \left[\text{LOCAL}\,|\,\text{CAT}\ \begin{bmatrix}\text{HEAD} & \textit{verb}\\ \text{SUBJ} & \left\langle \boxed{1} \right\rangle\\ \text{COMPS} & \boxed{A}\end{bmatrix}\right]\right\rangle\right]$$

The arguments are differentiated, depending on whether they are realized as subjects or as complements, and both are added to the ARG-ST list of the selecting verb. Application to (1) yields the phrase structure in (6).[2]

(6)



The argument inheritance treatment has turned out to be very influential in HPSG treatments of German and Dutch: Something along the lines of either (3) or (5) was adopted in amongst others Rentier (1994), Bouma & Van Noord (1998),

[1] There is a third valence feature (SP(ECIFIE)R) that is mainly used to model the selection of a determiner by a nominal. It is omitted here since it does not play any role in the treatment of argument inheritance, see also Van Eynde (2006).

[2] It has been argued that the first argument of a finite verb is a complement in German, see Müller (2002). If that assumption is adopted for Dutch, 1 is on the COMPS list of the future auxiliary. Nothing in this paper hinges on that choice.

Kathol (2000, chapter 8), and Müller (2002, chapter 2). It was also adopted for the treatment of clitic climbing in the Romance languages, as in Abeillé et al. (1998) for French and Monachesi (1998) for Italian.

3 Why to differentiate complement raising from subject raising

A common property of the argument inheritance proposals is that subjects and complements are raised in the same way: They are both integrated in the SUBCAT list, c.q. the ARG-ST list, of the selecting verb. This is in fact the reason why argument inheritance is also known as *generalized raising*. What will be argued now is that complement raising ought to be differentiated from subject raising, at least for Dutch. The evidence comes from three sources. They concern the occurrence of complement raising with subject control verbs, the binding properties of subject-to-object raisers, also known as ACI verbs (*Accusativus cum Infinitivo*), and the interaction of complement raising and the passive lexical rule.

3.1 Subject control verbs and complement raising

The formulation of argument inheritance in (5) allows for the occurrence of subject raising without complement raising, since $\boxed{A}$ may be the empty list.[3] What (5) does not allow, though, is the occurrence of complement raising without subject raising: The SUBJ list of the selected verb is required to contain one *synsem*, and that *synsem* must be identical to the first argument of the selecting verb.

This constraint now is too strict, since complement raising also occurs with subject control verbs, such as *willen* 'want' and *proberen* 'try'. Some instances are provided by the following sentences, taken from LASSY, a treebank for written Dutch (Van Noord et al., 2013). The control verbs are in bold and the raised complements in italics.

(7) Kasparov beschuldigde Gorbatsjov ervan dat hij *het bloedvergieten*
Kasparov accused Gorbatsjov there-of that he the bloodshed
niet had **willen** stoppen.
not had want.IPP stop.INF
'Kasparov accused Gorbatsjov that he had not wanted to stop the bloodshed.' [LASSY, dpc-ind-001648-nl-sen.p.19.s.6]

(8) ... nadat ze *zowel de PS als de PRL te vriend* had **proberen** te
... after she both the PS and the PRL as friend had try.IPP to
houden.
keep.INF

[3] In fact, if $\boxed{A}$ is declared empty, one gets the constraint which is characteristic of the English subject raising verbs (Ginzburg & Sag, 2000, p.22).

'... after she had tried to keep both the PS and the PRL as an ally.' [LASSY, WR-P-P-I-0000000106.p.7.s.6]

Notice that the control verbs in these sentences are affected by the IPP phenomenon (*Infinitivus pro Participio*), i.e. the use of the infinitive instead of the past participle in combination with the perfect auxiliary. They also allow complement raising, though, when they are not affected by IPP, as illustrated by (9), quoted from the CGN treebank, a treebank for spoken Dutch (Oostdijk et al., 2002).[4]

(9) ja en en ik heb *'r* **geprobeerd** te bellen maar d'r werd niet
yes and and I have her try.PSP to call.INF but there was not
opgenomen ...
picked-up ...
'yes and and I've tried to call her but there was no reply' [CGN, fna000583_351]

Summing up, subject control verbs are obviously not subject raisers, but they do allow complement raising, both in clustering constructions and in the third construction. Besides *willen* 'want' and *proberen* 'try', they include the verbs in Table 1 (Augustinus & Van Eynde, 2012).

pogen	'try'	trachten	'try'
dreigen	'threaten'	leren	'learn'
weigeren	'refuse'	menen	'mean, intend'
weten	'manage'	zien	'intend'
zoeken	'intend'	durven	'dare'
komen	'come'	liggen	'lie'
lopen	'walk'	staan	'stand'
zijn	'be in the activity of'	zitten	'sit'

Table 1: Other subject control verbs that allow complement raising

Since the argument inheritance constraint in (5) does not subsume the subject control verbs, we need a separate constraint to model the complement raising in clauses like (7–9):

(10)
$$\left[\text{ARG-ST}\ \left\langle \text{NP}_{\boxed{1}} \right\rangle \oplus \boxed{A} \oplus \left\langle \left[\text{LOCAL}\,|\,\text{CAT}\ \left[\begin{array}{l}\text{HEAD}\ \ \textit{verb}\\ \text{SUBJ}\ \ \left\langle \text{NP}_{\boxed{1}} \right\rangle\\ \text{COMPS}\ \ \boxed{A}\end{array}\right]\right]\right\rangle\right]$$

Complying with the way in which subject control verbs are canonically differentiated from subject raising verbs (Pollard & Sag, 1994; Sag et al., 2003), this

[4]This is an instance of the so-called third construction (den Besten et al., 1988; den Besten & Rutten, 1989).

constraint requires the unexpressed subject of the selected verb to share its index with the first argument of the selecting verb, rather than its entire *synsem* value. This addition of an extra constraint is by itself unobjectionable, but it does raise the suspicion that we are missing a generalization.

3.2 Subject-to-object raisers and the binding principles

In HPSG, the binding principles are canonically defined in terms of obliqueness relations in the ARG-ST list (Pollard & Sag, 1994; Sag et al., 2003).

> Principle A: An anaphoric pronoun must be coindexed with a less oblique argument on the same ARG-ST list.
> Principle B: A nonanaphoric NP may not be coindexed with a less oblique argument on the same ARG-ST list.

Assuming that raised subjects are integrated in the ARG-ST list of the selecting verb, this makes the right prediction for the subject-to-object raiser *ziet* 'sees' in (11).

(11) a. ... dat hij$_i$ zich$_{i/*j}$ die wedstrijd niet meteen **ziet** winnen.
... that he$_i$ himself$_{i/*j}$ that game not immediately sees win
'... that he does not expect himself to win that game rightaway.'

b. ... dat hij$_i$ hem$_{j/*i}$ die wedstrijd niet meteen **ziet** winnen.
... that he$_i$ him$_{j/*i}$ that game not immediately sees win
'... that he doesn't expect him to win that game rightaway.'

The raised reflexive pronoun *zich* 'himself' in (11a) must be coindexed with the subject of *ziet* 'sees', yielding the interpretation that he does not expect himself to win the contest. Similarly, the raised personal pronoun *hem* 'him' in (11b) cannot be coindexed with the subject of *ziet*, yielding the interpretation that he does not expect that person to win the contest. Raised subjects thus behave as *bona fide* arguments of the matrix verb, as illustrated by the ARG-ST list of *ziet* for the sentences in (11).

(12) a. *ziet*: ARG-ST <NP$_i$, 1 NP$_{i/*j}$, (...,) V[SUBJ <1>]>
b. *ziet*: ARG-ST <NP$_i$, 1 NP$_{j/*i}$, (...,) V[SUBJ <1>]>

Raised complements, by contrast, show the opposite behavior.

(13) a. * ... dat hij$_i$ ons *zich*$_i$ niet meteen **ziet** uitschakelen.
... that he$_i$ us himself$_i$ not immediately sees eliminate

b. ... dat hij$_i$ ons *hem*$_{i/j}$ niet meteen **ziet** uitschakelen.
... that he$_i$ us him$_{i/j}$ not immediately sees eliminate
'... that he doesn't expect us to eliminate him rightaway.'

If the raised reflexive pronoun in (13a) is integrated in the ARG-ST list of *ziet* and coindexed with its subject, as in (14a), then it complies with binding principle A, but the sentence is nonetheless illformed. Conversely, if the raised personal pronoun in (13b) is integrated in the ARG-ST list of *ziet* and coindexed with its subject, as in (14b), it violates binding principle B, but this interpretation is impeccable.

(14) a. *ziet*: ARG-ST <NP$_i$, 1 NP , 2 NP$_i$, V[SUBJ <1> , COMPS <2>]>
b. *ziet*: ARG-ST <NP$_i$, 1 NP , 2 NP$_{i/j}$, V[SUBJ <1> , COMPS <2>]>

As a consequence, we either need to tinker with the binding principles, or we have to treat the raised complements in another way than the raised subjects, integrating the latter but not the former in the ARG-ST list of the selecting verb.

3.3 Passive and complement raising

HPSG canonically treats passivization in terms of a lexical rule which reshuffles the order of the arguments on the ARG-ST list, as in (15), after (Sag et al., 2003, p.313).

(15)
$$\begin{bmatrix} \textit{tv-lxm} \\ \text{PHON} \quad \boxed{A} \\ \text{ARG-ST} \quad \left\langle \text{NP}_i \right\rangle \oplus \boxed{B} \end{bmatrix} \Longrightarrow_{LR} \begin{bmatrix} \text{PHON} \quad \text{F}_{psp}\left(\boxed{A}\right) \\ \text{SS} \mid \text{LOC} \mid \text{CAT} \mid \text{HEAD} \mid \text{VFORM} \quad \textit{pas} \\ \text{ARG-ST} \quad \boxed{B} \oplus \left\langle \left(\text{PP}_i \right) \right\rangle \end{bmatrix}$$

This rule relates a transitive verbal lexeme to its participial form, fixing the VFORM value to *passive* and changing the order in the ARG-ST list: The second argument of the verbal lexeme becomes the first argument of its passive counterpart.

Assuming that raised subjects are integrated in the ARG-ST list of the selecting verb, this makes the right prediction for the subject-to-object raiser *expect* in (16).

(16) a. We expect them to leave tomorrow.
b. They are expected to leave tomorrow.

Since the noun phrase which is realized by *them* is the second argument of the lexeme *expect*, it can become the first argument of its passive counterpart *expected*.

Raised complements, by contrast, behave differently, as illustrated in (17).

(17) a. ... dat hij *ons* **probeerde** het huis te verkopen.
... that he us tried the house to sell
'... that he tried to sell us the house.'
b. * ... dat *wij* **werden geprobeerd** het huis te verkopen.
... that we were tried the house to sell

The italicized complement of *verkopen* 'sell' in (17a) is raised and realized as a dependent of the subject control verb *proberen* 'try', but in contrast to the raised subject in (16) it cannot become the first argument of the passive *geprobeerd* 'tried'.

As a consequence, if we want to preserve a treatment of passivization in terms of reshuffling along the lines of (15), then we should not integrate the raised complements in the ARG-ST list of the selecting verb.

4 An alternative treatment of complement raising

The development of the alternative proceeds in four steps. First, we differentiate complement raising from subject raising. Second, we differentiate complement raising from complement extraction. Third, we show how the new treatment of complement raising naturally extends to a number of other phenomena, including adposition stranding. Fourth, we propose some constraints on complement raising.

4.1 Complement raising versus subject raising

In order to avoid the problems in the interaction with the binding principles and the passive lexical rule, we assume that raised subjects are integrated in the ARG-ST list of the selecting verb, while raised complements are not.

For the treatment of subject raising this implies that we can use the same lexical constraints as those that are used for English, i.e. one for subject-to-subject raisers and one for subject-to-object raisers, as in (18) (Ginzburg & Sag, 2000, 22).[5]

(18) a. *s-rsg-lx* ⇒ [ARG-ST ⟨[LOC 1] , [SUBJ ⟨LOC 1⟩]⟩]
b. *orv-lx* ⇒ [ARG-ST ⟨NP , [LOC 1] , [SUBJ ⟨LOC 1⟩]⟩]

A treatment in terms of **lexical** constraints is appropriate since the two types subsume a limited number of verbs. The subject raising lexemes (*s-rsg-lx*), for instance, include the modal, temporal and passive auxiliaries, while the object raising lexemes (*orv-lx*) include a number of perception verbs and causative verbs.

For the treatment of complement raising, by contrast, we adopt the following **phrasal** constraint.

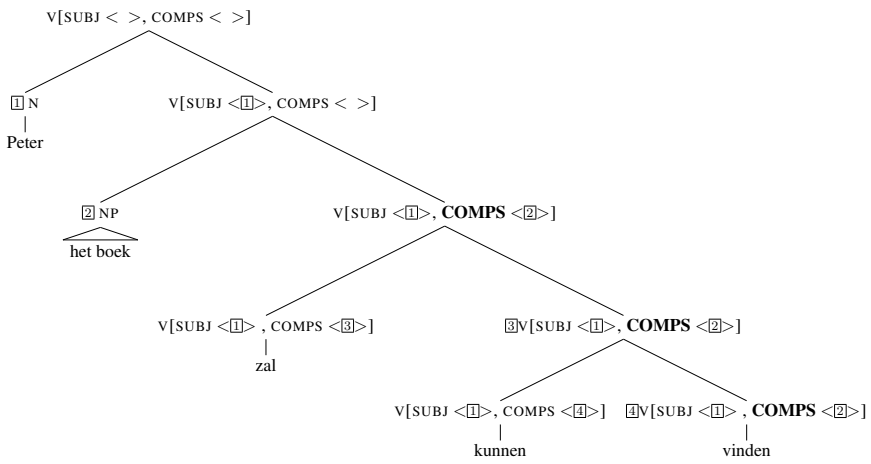
(19)
$$\begin{bmatrix} \textit{hd-ph} \\ \text{SS} \mid \text{LOC} \mid \text{CAT} \mid \text{COMPS} \quad \textit{list} \oplus \boxed{Z} \\ \text{NONHEAD-DTR} \mid \text{SS} \mid \text{LOC} \mid \text{CAT} \mid \text{COMPS} \quad \boxed{Z} \end{bmatrix}$$

In a headed phrase, the COMPS list of the nonhead daughter is appended to the COMPS list of the mother.[6] The $\boxed{Z}$ list may be empty, but it may also contain one or more members. In (20), for instance, which is our representation of (1), $\boxed{Z}$ corresponds to $<\boxed{2}>$.

[5] In this version, the sharing is limited to the objects of type *local*. In other versions, including that of Pollard & Sag (1994), the sharing concerns objects of type *synsem*. Nothing in this paper hinges on that distinction.

[6] In a non-headed phrase, such as a coordinate phrase, the COMPS list of the mother is identical to the COMPS list of each of the conjunct daughters.

(20)



Notice that the modal *kunnen* 'can' inherits the subject requirement of its infinitival complement, but not its COMPS list. The latter is propagated directly from the nonhead daughter to the mother. The same holds for the future *zal* 'will': It inherits the SUBJ list of its infinitival complement, but not its COMPS list.

Small as it is, this difference provides exactly what we need to avoid the problems with the argument inheritance treatment: It allows for complement raising in cases where there is no subject raising, and it does not integrate the raised complements in the ARG-ST list of the selecting verb.

4.2 Complement raising versus complement extraction

Complement raising need not only be differentiated from subject raising, but also from complement extraction. The latter concerns a long distance dependency that may cross clause boundaries, as in (21–22).

(21) Who do you think she said she would date?

(22) Wie beweert ze dat ze in Parijs ontmoet hebben?
who claims she that they in Paris met have
'Who does she claim they met in Paris?'

The complements of *date* and *ontmoet* 'met' are extracted and realized as a filler of the main clause. In HPSG, this is modeled in terms of a lexical rule which subtracts elements from the COMPS list and adds them to the nonlocal SLASH list, see (Ginzburg & Sag, 2000).

Complement raising, by contrast, is a middle distance dependency, and does not cross clause boundaries. To make this more precise let us adopt some notions of topological field theory, i.e. the analysis of the clause in terms of two poles (*Satzklammer*) and three fields (*Vorfeld, Mittelfeld, Nachfeld*). This style of analysis has been very influential in Dutch and German descriptive syntax (Haeseryn et al., 1997; Dudenredaktion, 2006), and some of its insights and terminology have

been imported in HPSG (Kathol, 2000; Müller, 2002). Pursuing in this direction let us make the assumption that complements cannot be raised beyond the first pole (*linke Satzklammer*). This is the position that is taken by the complementizer in verb-final clauses and by the finite verb in verb-initial clauses, i.e. V_1 and V_2 clauses.

That complementizers are a boundary for complement raising is illustrated in (23).

(23) a. . . . dat ze beweert **dat** ze *hem* in Parijs ontmoet hebben.
. . . that she claims that they him in Paris met have
'. . . that she claims that they met him in Paris.'

b. * . . . dat ze *hem* beweert **dat** ze in Parijs ontmoet hebben.
* . . . that she him claims that they in Paris met have

The italicized complement of *ontmoet* 'met' cannot be raised out of the clause that is introduced by the complementizer *dat* 'that'.

That finite verbs are a boundary for complement raising is less obvious, since it is possible to realize the complement of the main verb in the Vorfeld, as in (24).

(24) Dat boek zal Peter toch niet kunnen vinden.
that book shall Peter anyway not can find
'That book, Peter will not be able to find anyway.'

Notice, though, that this is an instance of topicalization, and that topicalization is canonically treated as a long distance dependency in HPSG, amongst others because it can cross clause boundaries, as in (25).

(25) That man I wish I had never known.

A useful test for differentiating topicalization from complement raising in Dutch is exemplified by the contrast in (26).

(26) a. Peter zal jou/je toch niet kunnen vinden.
Peter shall you/you.RED anyway not can find
'Peter will not be able to find you anyway.'

b. Jou/*je zal Peter toch niet kunnen vinden.
you/*you.RED shall Peter anyway not can find
'Me Peter will not be able to find anyway.'

Pronominal complements can be raised out of a verb cluster, as in (26a), no matter whether they take the full form or a phonologically reduced form, i.e. a form with a mute vowel or without vowel. Extraction, by contrast, as in (26b), is only possible for the full form (Van Eynde, 1999).[7]

[7] This restriction holds for extracted *complements*. Subjects may always occur in the Vorfeld, no matter whether they are full forms or reduced forms.

Taking stock, the middle distance nature of complement raising is clear from the fact that it cannot go beyond the first pole: It is bounded by the complementizer in verb-final clauses and by the finite verb in verb-initial clauses. How these constraints can be spelled out in formal terms is discussed in section 4.4.

4.3 Extensions

So far, we have focussed on complement raising out of nonfinite verbal complements. This, however, is not the only type of raising that the phrasal constraint in (19) allows. It also allows raising out of nonverbal complements, since it does not put any constraints on the syntactic category of the nonhead daughter. Moreover, it also allows raising out of subjects and adjuncts, since (19) applies to all headed phrases.

4.3.1 Complement raising out of nonverbal complements

Some examples of complement raising out of adjectival complements are given in (27–28).

(27) . . . dat we *die hittegolf* nog steeds niet **kwijt** zijn!
. . . that we that heat wave still always not lost are
'. . . that we are not finished with that heat wave yet!' [LASSY, WS-U-E-A-0000000221.p.32.s.2]

(28) . . . dat de bevolking van Zimbabwe *haar huisbakken dictator*
. . . that the people of Zimbabwe her home-grown dictator
onderhand meer dan **beu** is.
by now more than fed-up is
'. . . that the people of Zimabwe are more than fed up with their homegrown dictator by now.' [LASSY, WR-P-P-I-0000000219.p.4.s.4]

The italicized nominals are complements of the predicative adjectives in bold, but they are not realized within the AP. Instead, they are raised and realized in the left part of the Mittelfeld, preceding the VP adjuncts.

Complement raising also subsumes the instances of adposition stranding in (29–30).[8]

(29) . . . dat zij *daar* nog wel **van** hield.
. . . that she there still rather of liked
'. . . that she rather liked it.' [CGN, fna000741_12]

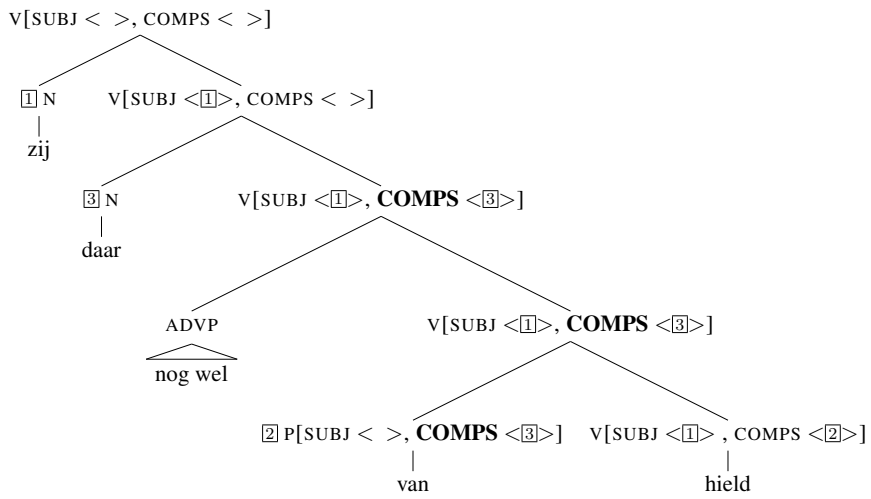
(30) . . . als je *er* pas achteraf **over** nadenkt, is het misschien te laat.
. . . if you there only later about think-of, is it maybe too late

[8]For a treatment of adposition stranding in Dutch, see a.o. Van Riemsdijk (1978) and Beeken (1991).

'... if you only think about it afterwards, it may be too late.' [LASSY, WR-P-P-C-0000000047.txt-10]

Also here, the italicized pronouns are complements of the adpositions in bold, but they are not realized within the PP. Instead, they are raised and realized in the left part of the Mittelfeld, preceding the VP adjuncts, as illustrated by the representation of (29) in (31).

(31)



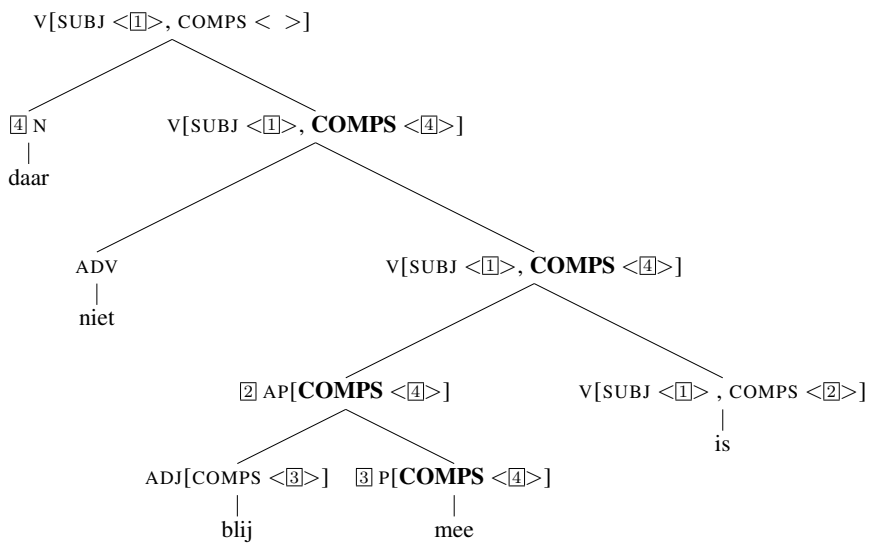
The unsaturated COMPS requirement of the adposition (3) is inherited by the verbal projection and discharged after the addition of *daar* 'there'. Notice that the adposition has an empty SUBJ list, in accordance with the canonical HPSG assumption that argument marking adpositions do not select a subject. In the context of this paper, it provides further evidence for the claim that complement raising may occur in environments where there is no subject raising.

Since there are adjectives which take adpositional complements, complement raising can be applied iteratively, as in (32).

(32) ... dat hij *daar* niet **blij** mee is.
... that he there not happy with is
'... that he is not happy with that.'

The pronominal complement *daar* 'there' is first raised out of the PP, then out of the predicative AP, and finally out of the V-final VP, as illustrated in (33).

(33)



This is comparable to the iterative subject raising in sequences like (34).

(34) He does not seem to be likely to win this game.

The surface subject of *does* is the understood subject of *win this game*, and the relation is mediated by a sequence of subject raising lexemes, including *to*, *likely*, *be*, *seem* and *does*.

4.3.2 Complement raising out of functors and subjects

The previous examples all concerned raising out of complements, but the constraint on complement raising in (19) does not require this: It also allows the nonhead daughter to be a functor or a subject. (35), for instance, is an example of complement raising out of a PP adjunct.

(35) ... dat hij *er* veel verliezen **door** heeft geleden.
... that he there many losses by has suffered
'... that he suffered many losses because of it.'

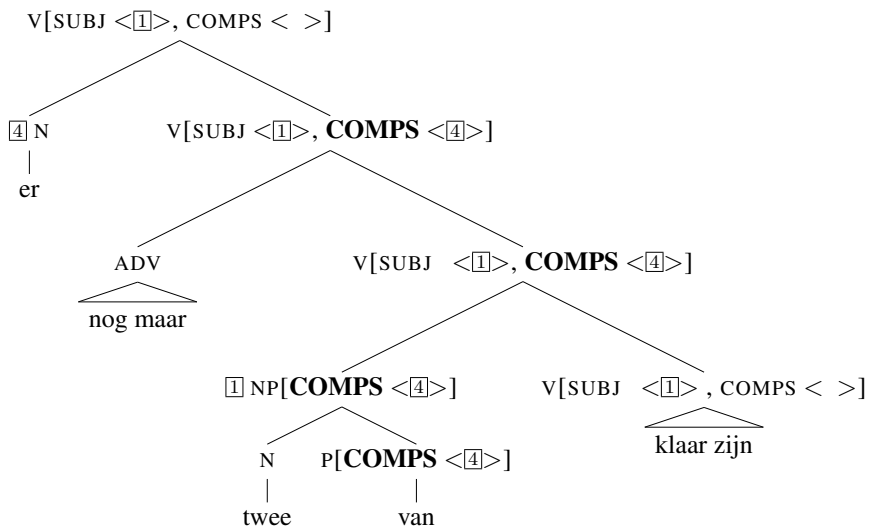
The italicized pronoun is a complement of the adposition *door* 'by' and the latter heads a PP adjunct that specifies the cause of the losses.

Raising out of subjects is exemplified in (36).

(36) ... dat *er* nog maar twee **van** klaar zijn.
... that there still but two of ready are
'... that only two of them are ready.'

The italicized pronoun is a complement of the adposition *van* 'of', which heads the PP adjunct of the cardinal *twee* 'two' which in its turn heads the subject of the clause, as spelled out in (37).

(37)



4.4 Constraints

Given the treatment of complement raising in terms of a phrasal —rather than a lexical— constraint and given the rather permissive nature of its formulation in (19), an obvious question is whether it is not too permissive. The equally obvious answer is that excessive permissivity can be avoided by the addition of extra constraints on (19). To show how this can be done we first discuss English and then return to Dutch.

4.4.1 English

English is a language that allows subject raising and complement extraction, but assuming that it obeys the Empty COMPS Constraint, as defined in Ginzburg & Sag (2000, 33), it does not allow complement raising.

(38) Empty COMPS Constraint (Ginzburg & Sag, 2000, 33)
phrase:
$\left[\text{COMPS} \left\langle \right\rangle\right] \rightarrow \ldots$

Indeed, if phrases are required to have an empty COMPS list, then it follows that the $\boxed{Z}$ list in the phrasal constraint on complement raising in (19) must be empty and, hence, that complement raising is blocked.

The fact that English allows adposition stranding does not provide any evidence against this assumption, since the stranding invariably results from complement extraction, as in (39).

(39) a. *What* do you think they were talking **about**?
b. *This* I would never dare talk **about** in her presence.

Stranding that results from complement raising, as in (40), is not possible.

(40) a. * John heard us *this* talk **about**.
b. * We saw him *that* give a talk **about**.

In that respect, English differs from Dutch, where the equivalents of (40) are well-formed.

(41) a. Jan heeft *er* ons **over** horen praten.
Jan has there us about hear talk
'Jan heard us talk about it.'
b. We hebben hem *daar* een lezing **over** zien geven.
we have him there a talk about see give
'We saw him give a talk about that.'

In sum, the addition of the Empty COMPS Constraint suffices to rule out complement raising from the language.

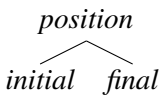
4.4.2 Dutch

Since Dutch does not abide by the Empty COMPS Constraint, it allows complement raising, but this does not mean that its complements can be raised anywhere. For a start, they cannot be raised beyond the first pole, as demonstrated in section 4.2. To model this for the case in which the first pole is a complementizer we add the constraint in (42).

(42)
$$\begin{bmatrix} \textit{hd-ph} & \\ \text{SS} \mid \text{LOC} \mid \text{CAT} & \begin{bmatrix} \text{HEAD} & \textit{complementizer} \\ \text{COMPS} & \langle\,\rangle \end{bmatrix} \end{bmatrix}$$

Phrases which are headed by a complementizer are required to have an empty COMPS list. This suffices to block complement raising out of CPs.[9]

If the first pole is a finite verb, we need an extra feature to model the relevant constraint. We call it POSITION and add it to the HEAD values of verbs. Its possible values are given in (43).

(43)


In terms of this dichotomy, the nonfinite verbs are invariably *final* and the imperative forms *initial*. The other finite forms can occur in either position, and hence receive the underspecified *position* value.[10]

[9] It does not block complement extraction, though, since it does not require the SLASH value of a CP to be empty.

[10] The term *initial* subsumes both V_1 and V_2.

final	nonfinite forms
initial	imperative forms
position	nonimperative finite forms

Table 2: Three types of verb forms

With this addition the constraint which blocks complement raising out of V-initial VPs can now be spelled out as in (44).

(44)
$$\left[\begin{array}{l} \textit{hd-ph} \\ \\ \text{SS} \mid \text{LOC} \mid \text{CAT} \left[\begin{array}{ll} \text{HEAD} & \left[\begin{array}{ll} \textit{verb} & \\ \text{POSITION} & \textit{initial} \end{array}\right] \\ \\ \text{COMPS} & \langle\ \rangle \end{array}\right] \end{array}\right]$$

Phrases which are headed by a verb that is in V-initial position, are required to have an empty COMPS list, just like CPs.

Together, the constraints in (42) and (44) model the fact that complements cannot be raised beyond the first pole. As such, they capture what differentiates complement raising from complement extraction.

A less conspicuous constraint concerns the raising out of PPs. To pave the way for its treatment we start from the observation that Dutch adpositions come in three types: There are those that invariably follow their complement, such as *toe* 'to' and *mee* 'with', there are those that invariably precede their complement, such as *tot* 'to, till' and *met* 'with', and there are those that can precede as well as follow their complement, such as *in* 'in' and *van* 'of'. Table 3 provides a survey.

final	*mee, toe, af, heen*
initial	*met, tot, te, sinds, sedert, als, tijdens, wegens, volgens, ...*
position	*in, op, van, aan, bij, door, ...*

Table 3: Three types of adpositions

The distinction is not only relevant to treat the linear order within the PP, it also correlates with some other facts. Realization in the Nachfeld, for instance, also known as PP-over-V, is possible for P-initial PPs, as shown in (45), but not for P-final PPs, as shown in (46–47).

(45) a. ... dat we nog steeds [**op** een goede afloop] hopen.
... that we still always for a good outcome hope
'... that we are still hoping for a good outcome.'

b. ... dat we nog steeds hopen [**op** een goede afloop].
... that we still always hope for a good outcome
'... that we are still hoping for a good outcome.'

(46) a. ... dat we nog steeds [daar **op**] hopen.
... that we still always there for hope
'... that we are still hoping for that.'

b. * ... dat we nog steeds hopen [daar **op**].
... that we still always hope there for

(47) a. ... dat hij voorzichtig [de garage **in**] reed.
... that he carefully the garage in drove
'... that he drove carefully into the garage.'

b. * ... dat hij voorzichtig reed [de garage **in**].
... that he carefully drove the garage in

Conversely, complement raising is possible out of P-final PPs, as shown in (48–49), but not out of P-initial PPs, as shown in (50).

(48) a. ... dat we nog steeds [daar **op**] hopen.
... that we still always there for hope
'... that we are still hoping for that.'

b. ... dat we *daar* nog steeds **op** hopen.
... that we there still always for hope
'... that we are still hoping for that.'

(49) a. ... dat hij voorzichtig [de garage **in**] reed.
... that he carefully the garage in drove
... that he drove carefully into the garage.'

b. ... dat hij *de garage* voorzichtig **in** reed.
... that he the garage carefully in drove
... that he drove carefully into the garage.'

(50) a. ... dat we nog steeds [**op** een goede afloop] hopen.
... that we still always for a good outcome hope
'... that we are still hoping for a good outcome.'

* ... dat we *een goede afloop* nog steeds **op** hopen.
... that we a good outcome still always for hope

To model the constraint that complements cannot be raised out of P-initial PPs we propose a constraint that resembles the one in (44).

(51)
$$\begin{bmatrix} \textit{hd-ph} & \\ \text{SS} \mid \text{LOC} \mid \text{CAT} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \textit{adposition} \\ \text{POSITION} \quad \textit{initial} \end{bmatrix} \\ \text{COMPS} & \langle\, \rangle \end{bmatrix} \end{bmatrix}$$

(51) requires the P-initial PPs to have an empty COMPS list, just like the V-initial VPs and the CPs.

When this constraint is combined with the observations about PP-over-V, it correctly accounts for the fact that adpositions cannot be stranded in the Nachfeld.

(52) a. * ... dat we *een goede afloop* nog steeds hopen **op**
... that we a good outcome still always hope for

b. * ... dat hij *daar* nog steeds hoopt **op**
... that he there still always hopes for

c. * ... dat hij *de garage* voorzichtig reed **in**
... that he the garage carefully drove in

(52a) is illformed, since (51) does not allow to raise a complement out of a P-initial PP, and (52b–52c) are illformed, since P-final PPs are not allowed in the Nachfeld.

Given that complementizers are invariably CP-initial, at least in the Germanic and the Romance languages, it is tempting to replace the three constraints with one more general constraint, blocking complement raising out of all head-initial phrases. This, however, would be too strict, since it is possible to raise complements out of head-initial APs and NPs, as shown in (33) and (37), respectively.

Further investigation will reveal whether the three constraints suffice to prevent overgeneration and whether it is possible to formulate them in more general terms. What is noteworthy, though, is that they mesh remarkably well with the fact that English does not allow complement raising, since English VPs and PPs are invariably head-initial.

5 Conclusion

To model the raising of complements out of verb clusters HPSG canonically employs the device of *argument inheritance*, also known as *argument composition* or *generalized raising* (section 2). When applied to Dutch, its interaction with the binding principles and the passive lexical rule yields erroneous predictions (section 3). As an alternative, we propose to employ different devices for subject raising and complement raising: While the former is modeled in terms of lexical constraints, as in English, the latter is modeled in terms of a constraint on headed phrases (section 4). This constraint also subsumes other instances of complement raising, such as adposition stranding in Dutch. In order to avoid overgeneration, we added a number of constraints to prevent complement raising out of CPs, V-initial VPs and P-initial PPs.

In future work we will further explore the ramifications of this proposal for Dutch, investigating when complement raising is obligatory and when it is optional. We also intend to explore the potential of this proposal for the treatment of middle distance dependencies in other languages, such as clitic climbing in the Romance languages and clustering in German.

References

Abeillé, Anne, Danièle Godard, Philip Miller & Ivan A. Sag. 1998. French bounded dependencies. In S. Balari & L. Dini (eds.), *Romance in HPSG*, 1–54. Stanford: CSLI Publications.

Augustinus, Liesbeth & Frank Van Eynde. 2012. A Treebank-based Investigation of IPP-triggering Verbs in Dutch. In I. Hendrickx, S. Kübler & K. Simov (eds.), *Proceedings of TLT 11*, 7–12. Lisbon: Colibri.

Beeken, Jeannine. 1991. *Spiegelstructuur en variabiliteit. pre- en postposities in het Nederlands*: University of Leuven dissertation.

den Besten, Hans & Jean Rutten. 1989. On Verb Raising, Extraposition and Free Word Order in Dutch. In D. Jaspers, W. Klooster, Y. Putseys & P. Seuren (eds.), *Sentential Complementation and the Lexicon: Studies in Honour of Wim de Geest*, 41–56. Dordrecht: Foris.

den Besten, Hans, Jean Rutten, Tonjes Veenstra & Joop Veld. 1988. Verb Raising, Extrapositie en de Derde Constructie. Unpublished manuscript.

Bouma, Gosse & Gertjan Van Noord. 1998. Word Order Constraints on Verb Clusters in German and Dutch. In E. Hinrichs (ed.), *Complex Predicates in Nonderivational Syntax*, New York: Academic Press.

Dudenredaktion. 2006. *DUDEN. Die Grammatik. Unentbehrlich für richtiges Deutsch.* Mannheim/Leipzig/Vienna/Zürich: Dudenverlag.

Ginzburg, Jonathan & Ivan A. Sag. 2000. *Interrogative Investigations*. Stanford: CSLI Publications.

Haeseryn, Walter, K. Romijn, Guido Geerts, J. de Rooij & Maarten C. van den Toorn. 1997. *Algemene Nederlandse spraakkunst*. Nijhoff and Wolters Plantyn.

Hinrichs, Erhard & Tsuneko Nakazawa. 1989. Flipped out: AUX in German. In *Papers from the 25th Regional Meeting of the Chicago Linguistic Society*, 193–202. Chicago.

Hinrichs, Erhard & Tsuneko Nakazawa. 1994. Linearizing AUXs in German Verbal Complexes. In J. Nerbonne, K. Netter & C. Pollard (eds.), *German in Head-driven Phrase Structure Grammar*, 11–37. Stanford: CSLI Publications.

Johnson, Mark. 1986. A GPSG account of VP structure in German. *Linguistics* 24. 871–882.

Kathol, Andreas. 2000. *Linear Syntax*. Oxford University Press.

Miller, Philip H. & Ivan A. Sag. 1997. French Clitic Movement without Clitics or Movement. *Natural Language and Linguistic Theory* 15. 573–639.

Monachesi, Paola. 1998. Decomposing Italian clitics. In S. Balari & L. Dini (eds.), *Romance in HPSG*, 305–357. Stanford: CSLI Publications.

Müller, Stefan. 2002. *Complex Predicates*. Stanford: CSLI Publications.

Oostdijk, Nelleke, Wim Goedertier, Frank Van Eynde, Louis Boves, Jean-Pierre Martens, Michael Moortgat & Harald Baayen. 2002. Experiences from the Spoken Dutch Corpus Project. In M. Gonzalez Rodriguez & C.P. Saurez Araujo (eds.), *Proceedings of LREC 3*, 340–347. Paris.

Pollard, Carl & Ivan Sag. 1994. *Head-driven Phrase Structure Grammar*. Stanford/Chicago: CSLI Publications and University of Chicago Press.

Rentier, Gerrit. 1994. Dutch Cross Serial Dependencies in HPSG. In *Proceedings of COLING 94*, 818–822. Kyoto, Japan.

Sag, Ivan A., Thomas Wasow & Emily Bender. 2003. *Syntactic Theory. a Formal Introduction. Second Edition*. Stanford, California: CSLI Publications.

Van Eynde, Frank. 1999. Major and minor pronouns in Dutch. In G. Bouma, E. Hinrichs, G.-J. Kruijff & R. Oehrle (eds.), *Constraints and Resources in Natural Language Syntax and Semantics*, 137–151. Stanford: CSLI Publications.

Van Eynde, Frank. 2006. NP-internal agreement and the structure of the noun phrase. *Journal of Linguistics* 42. 139–186.

Van Noord, Gertjan, Gosse Bouma, Frank Van Eynde, Daniël de Kok, Jelmer van der Linde, Ineke Schuurman, Erik Tjong Kim Sang & Vincent Vandeghinste. 2013. Large Scale Syntactic Annotation of Written Dutch: Lassy. In P. Spyns & J. Odijk (eds.), *Essential Speech and Language Technology for Dutch*, 147–164. Springer.

Van Riemsdijk, Henk. 1978. *A Case Study in Syntactic Markedness*. Lisse: The Peter de Ridder Press.