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

Liesbeth Augustinus and Frank Van Eynde

Center for Computational Linguistics, KU Leuven, Belgium

Abstract

In V-final languages, such as German and Dutch, the auxiliaries and the main verb tend to form a cluster in which the main verb is separated from its syntactic arguments by one or more auxiliaries. The current treatment to link the main verb to its arguments is known as *argument inheritance* or *generalized raising*. It extends the treatment of subject raising to all arguments of the verb.

This paper argues that subject raising had better be distinguished from complement raising, first because complement raising also occurs with subject control verbs, and second because raised complements behave differently with respect to binding than raised subjects. As an alternative we propose to treat subject raising in terms of the same lexical constraints as for English, and to treat complement raising in terms of a phrasal constraint. The proposal is supported with data from two Dutch treebanks.

1 Introduction

In V-final languages, such as German and Dutch, the auxiliaries and the main verb tend to form a cluster in which the main verb is separated from its syntactic arguments by one or more auxiliaries, as in the Dutch (1), where *vinden* is separated from its arguments by the future auxiliary *zal* and the modal *kunnen*.

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

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

(2)
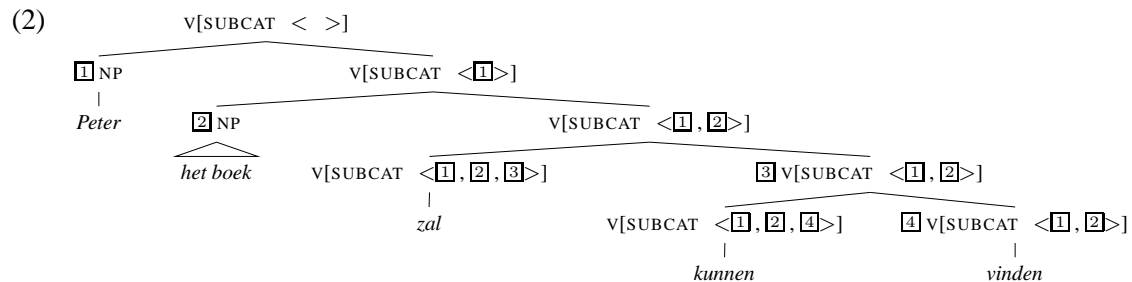
The relation between the main verb and its arguments is modeled in terms of **argument inheritance**: each auxiliary inherits the SUBCAT list of its unsaturated verbal complement, as in (3), after Hinrichs and Nakazawa (1994).

$$(3) \quad \left[\text{SUBCAT} \;\; \boxed{A} \;\; \oplus \;\; \left\langle \left[\text{LOCAL} \,|\, \text{CAT} \begin{bmatrix} \text{HEAD} & verb \\ \text{SUBCAT} & \boxed{A} \end{bmatrix} \right] \right\rangle \right]$$

In more recent versions of HPSG, where SUBCAT is replaced with the valence features SUBJ, COMPS and SPR, and where lexical signs are assigned the feature ARG-ST, (3) is reformulated as in (4).

$$(4) \quad \left[\text{ARG-ST} \;\; \left\langle \boxed{1} \right\rangle \;\; \oplus \;\; \boxed{A} \;\; \oplus \;\; \left\langle \left[\text{LOCAL} \,|\, \text{CAT} \begin{bmatrix} \text{HEAD} & verb \\ \text{SUBJ} & \left\langle \boxed{1} \right\rangle \\ \text{COMPS} & \boxed{A} \end{bmatrix} \right] \right\rangle \right]$$

This inheritance is also known as **generalized raising**, since it extends the treatment of subject raising to all arguments.[1]

The aim of this paper is to demonstrate that subject raising and complement raising are different phenomena and that a treatment which generalizes over both, as in (3) or (4), is less appropriate than one which differentiates them.

2 Why to differentiate complement raising from subject raising

2.1 Complement raising without subject raising

The first argument for differentiating complement raising from subject raising is that complements are not only raised out of the verbal complement of subject raising verbs, as in (1), but also out of the verbal complement of subject control verbs, as in the Dutch (5–7).[2]

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

(6) In 2001 heb ik *saxofoon*   **leren** spelen.
in 2001 have I   saxophone learn play
'In 2001 I learnt to play the saxophone.' [LASSY, dpc-qty-000936-nl-sen.p.36.s.2]

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

That the verbs in bold are indeed control verbs is clear from the fact that they are not compatible with a nonreferential subject. In that respect, they differ from the raising verbs, such as *blijven* 'keep' and *beginnen* 'begin'.[3]
This is confirmed by the passive test: the active sentence in (10) is not semantically equivalent to its passive counterpart in (11).

---

[1]If one requires $\boxed{A}$ to be the empty list, one gets the constraint which is characteristic of subject raising.

[2]The examples are taken from the LASSY treebank, see Van Noord, Bouma, Van Eynde, de Kok, van der Linde, Schuurman, Tjong Kim Sang and Vandeghinste (2013). The control verbs are in bold and the raised complements of the main verb are in

(8) Het blijft/*wil/*leert   regenen.
     it   keeps/*wants/*learns rain
     'It keeps/*wants/*learns raining.'

(9) Het begint/*probeert te regenen.
     it   begins/*tries   to rain
     'It begins/*tries to rain.'

(10) Hij wilde   haar verrassen.
     he wanted her surprise
     'He wanted to surprise her.'

(11) Zij wilde   door hem verrast   worden.
     she wanted by   him surprised be
     'She wanted to be surprised by him.'

In (5–7) both the control verb and its main verb complement are part of the verb cluster, but this is not a necessary condition for complement raising. In (12), for instance, the main verb complement of the control verb is in the Nachfeld, but also here its (italicized) complement is raised out of the infinitival complement.[4]

(12) ... ik heb *'r* **geprobeerd** te bellen maar d'r   werd niet opgenomen ...
     ... I   have her tried   to call   but there was not up-taken   ...
     '... I've tried to call her but there was no reply ...' [CGN, fna000583_351]

This is an instance of the so-called third construction. A conspicuous difference between the third construction and the clustering constructions in (5–7) concerns the form of the control verb: In (12) it is a past participle, as required by the perfect auxiliary, but in (5–7) it is a bare infinitive, in spite of the presence of the perfect auxiliary. This use of the infinitive where a past participle is expected, is known as the IPP-phenomenon (Infinitivus Pro Participio). It only affects the clustering verbs.

Taking stock, subject control verbs allow complement raising, both when they are used as clustering verbs and when they are used as main verbs. The phenomenon is quite common: A quantitative study of the LASSY treebank, supplemented by Google search, shows that the set of clustering control verbs includes at least 21 verbs, see Augustinus and Van Eynde (2012).

2.2 Binding

The second argument for differentiating complement raising from subject raising is that raised subjects interact in the expected way with the binding principles, whereas raised complements do not. To show this, we make use of subject-to-object raisers, also known as AcI-verbs (Accusativus cum Infinitivo).

The HPSG binding principles are defined in terms of obliqueness relations in the ARG-ST list: While anaphoric pronouns must be coindexed with a less oblique argument on the same ARG-ST list, nonanaphoric NPs may not be coindexed with a less oblique argument on the same ARG-ST list, see Sag, Wasow and Bender (2003, 207). Given that raised subjects are integrated in the ARG-ST list of the selecting verb, this makes the right prediction for both (13a), where the raised reflexive must be coindexed with the subject, and (13b), where the raised personal pronoun cannot be coindexed with the subject.

(13) a. ... dat hij$_i$ zich$_i$ *die wedstrijd* niet meteen   **ziet** winnen.
         ... that he$_i$ himself$_i$ that game   not immediately sees win
         '... that he does not expect himself to win that game rightaway.'

italics.

[3]Some verbs have both control and raising uses. *Willen*, for instance, is a control verb in its volitional use, as in (5), and a raising verb in its less common modal use, exemplified in *het wil hier al eens regenen* 'it will occasionally rain here'.

[4]This example is taken from the Spoken Dutch Treebank (CGN), see Oostdijk, Goedertier, Van Eynde, Boves, Martens, Moortgat and Baayen (2002).

b. ... dat hij$_i$ hem$_{j/*i}$ *die wedstrijd* niet meteen **ziet** winnen.
   ... that he$_i$ him$_{j/*i}$ that game   not immediately see win

   '... that he doesn't expect him to win that game rightaway.'

Raised complements, by contrast, show exactly the opposite behavior.

(14) 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.'

The raised reflexive in (14a) cannot be interpreted as coindexed with the subject, and the raised personal pronoun in (14b) can. This suggests that the raised complements had better not be integrated in the ARG-ST list of the perception verb.

3   How to differentiate complement raising from subject raising

To accommodate the observations in the previous section, we need a treatment (1) which does not automatically link complement raising to subject raising, and (2) which does not integrate the raised complements in the ARG-ST list of the dominating verb. To meet these requirements, we treat subject raising in terms of the same **lexical constraints** as those that are used for English, i.e. one for subject raisers and one for subject-to-object raisers, aka AcI verbs.[5]

(15) a. *s-rsg-lx* $\Rightarrow$ [ARG-ST $\langle$ ①, [SUBJ $\langle$①$\rangle$]$\rangle$]
   b. *orv-lx* $\Rightarrow$ [ARG-ST $\langle$NP, ①, [SUBJ $\langle$①$\rangle$]$\rangle$]

   For complement raising, however, we propose the **phrasal constraint** in (16).

(16) $$\begin{bmatrix} \textit{hd-phr} \\ \text{SYNSEM} \,|\, \text{LOC} \,|\, \text{CAT} \,|\, \text{COMPS} \quad \textit{list} \ \oplus \ \boxed{Z} \\ \text{NONHEAD-DTR} \,|\, \text{SYNSEM} \,|\, \text{LOC} \,|\, \text{CAT} \,|\, \text{COMPS} \quad \boxed{Z} \end{bmatrix}$$

In a headed phrase, the COMPS list of the nonhead daughter is appended to the COMPS list of the mother.[6] Application to (1) yields (17).

(17)

The auxiliaries select a verbal complement and inherit its SUBJ list, but not its COMPS list. The latter is propagated directly from the nonhead daughter to the mother. Small as this difference may seem, it provides exactly what we need to meet the two requirements: 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.

3.1 Complement raising out of nonverbal projections

Since (16) does not contain any restrictions on the syntactic category of the mother or the nonhead daughter, it not only allows complement raising out of verbal projections, but also out of adpositional and adjectival projections, as in (18).

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

      b. Ik begrijp   dat we *die hittegolf* nog steeds niet [_ kwijt] zijn!
          I   understand that we that heat wave still still   not   lost   are
          'I understand that we are not finished with that heat wave yet!' [LASSY, WR-P-P-I-0000000219.p.4.s.4]

The pronominal complement of the adposition in (18a) is not realized within the PP, but raised and realized in the left part of the Mittelfeld, preceding the VP adjuncts.[7] The same holds for the nominal complement of the predicative adjective in (18b).

3.2 Constraints on complement raising

Constraints on complement raising take the form of requirements that the COMPS list of some given phrase must be empty. An extreme case is the Empty COMPS Constraint, which requires all phrases to have an empty COMPS list:

(19) Empty COMPS Constraint (Ginzburg and Sag 2000, 364)
      phrase:
      $\begin{bmatrix} \text{COMPS} & \langle\,\rangle \end{bmatrix}$ $\rightarrow$ ...

Languages which abide by the ECC, such as English, do not allow complement raising at all.[8]

Dutch does not abide by the ECC, but has some more specific constraints: It does not allow complement raising out of VPs with V-initial order, out of PPs with P-initial order, and out of CPs. These constraints will be motivated and spelled out in the full paper.

4 Conclusion

To model the link between a verb and its arguments in Dutch and German verb clusters it has become a matter of common practice to treat subject raising and complement raising in the same way. This paper has provided arguments for differentiating complement raising from subject raising (in Dutch) and has

---

[7]Since prepositions which are externally selected by the verb are canonically assumed to have an empty SUBJ list, this is yet another instance of complement raising without subject raising.

[8]Such languages may allow complement extraction, but that is another phenomenon. For comparison, subject extraction is another phenomenon than subject raising.

proposed a way on how to do this: While subject raising is treated in terms of the same lexical constraints as for English, complement raising is treated in terms of a phrasal constraint.

References

Augustinus, L. and Van Eynde, F.(2012), A treebank-based investigation of IPP-triggering verbs in Dutch, *in* I. Hendrickx, S. Kübler and K. Simov (eds), *Proceedings of TLT 11*, Colibri, Lisbon, pp. 7–12.

Ginzburg, J. and Sag, I.(2000), *Interrogative Investigations*, CSLI Publications, Stanford.

Hinrichs, E. and Nakazawa, T.(1989), Flipped out: AUX in German, *Papers from the 25th Regional Meeting of the Chicago Linguistic Society*, Chicago, pp. 193–202.

Hinrichs, E. and Nakazawa, T.(1994), Linearizing AUXs in German verbal complexes, *in* J. Nerbonne, K. Netter and C. Pollard (eds), *German in HPSG*, CSLI Publications, Stanford, pp. 11–37.

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

Oostdijk, N., Goedertier, W., Van Eynde, F., Boves, L., Martens, J.-P., Moortgat, M. and Baayen, H.(2002), Experiences from the Spoken Dutch Corpus Project, *in* M. Gonzalez Rodriguez and C. Saurez Araujo (eds), *Proceedings of LREC 3*, Paris, pp. 340–347.

Pollard, C. and Sag, I.(1994), *Head-driven Phrase Structure Grammar*, CSLI Publications and University of Chicago Press, Stanford/Chicago.

Sag, I. A., Wasow, T. and Bender, E.(2003), *Syntactic theory. A formal introduction. Second edition*, CSLI Publications, Stanford, California.

Van Noord, G., Bouma, G., Van Eynde, F., de Kok, D., van der Linde, J., Schuurman, I., Tjong Kim Sang, E. and Vandeghinste, V.(2013), Large scale syntactic annotation of written Dutch: Lassy, *in* P. Spyns and J. Odijk (eds), *Essential Speech and Language Technology for Dutch*, Springer, pp. 147–164.