

Complement Raising, Extraction and Adposition Stranding in Dutch

Frank Van Eynde and Liesbeth Augustinus
Centre for Computational Linguistics, University of Leuven, Belgium

Let us take a look at some typical examples of adposition stranding in Dutch.

- (1) *Waar* denk je dat ze [... op] wachten?
where think you that they [... up] wait
'What do you think they are waiting for?'
- (2) Ze zegt dat ze *daar* soms nog [... aan] denkt.
she says that she there sometimes still [... on] thinks
'She says that she still thinks about it from time to time.'
- (3) We hebben *er* haar een liedje [... over] horen zingen.
we have there her a song [... about] hear sing
'We heard her sing a song about it.'

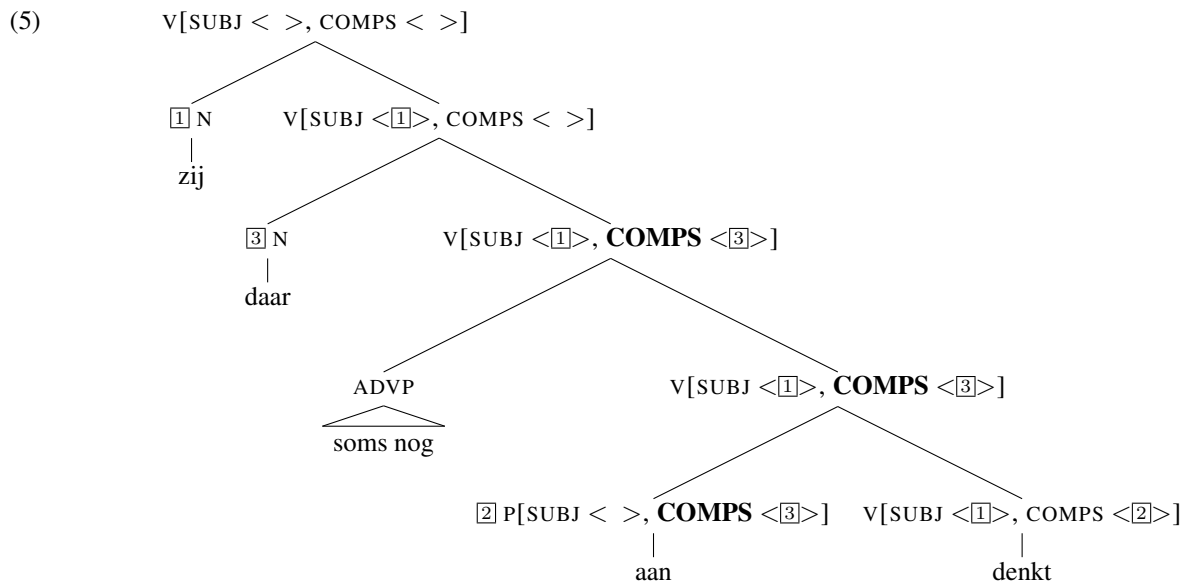
The stranded adposition is in the right part of the Mittelfeld, close to the verb cluster, while its complement is realized in the Vorfeld, as in (1), or in the left part of the Mittelfeld, as in (2–3). It is usually (not always) one of the R-pronouns, so-called because they contain the R-sound (*er, d'r, daar, hier, waar, ergens, nergens, overal*). The phenomenon has been studied extensively, see Haeseryn et al. (1997) and Broekhuis (2013) for a descriptive overview and Van Riemsdijk (1978) and Bennis (1986) for a transformational treatment. Rentier (1993) and Bouma (2000) provide an HPSG analysis; both treat the phenomenon in terms of extraction, employing nonlocal devices such as SLASH and BIND. This paper endorses the extraction treatment for (1), but for (2–3) it proposes an alternative which is based on argument inheritance. Section 1 presents the analysis, section 2 discusses Bouma's reasons for not adopting the argument inheritance approach, and section 3 draws some conclusions.

1 The analysis

Comparing (1) with (2–3) we observe some obvious differences: the stranding in (1) shows the typical characteristics of a long-distance dependency, crossing clause boundaries, but the stranding in (2–3) concerns a more bounded type of dependency, in the sense that the adposition and its complement are in the same Mittelfeld. In HPSG long-distance dependencies are standardly treated in terms of nonlocal devices, but the more bounded types of dependency are commonly dealt with in terms of other devices, such as argument inheritance (aka generalized raising). It was first proposed in Hinrichs & Nakazawa (1994) for a treatment of the German verb clusters, and was adopted and adapted by various authors to deal with similar phenomena in other languages, such as the Dutch verb clusters in Bouma & van Noord (1998) and clitic climbing in French and Italian, see Abeillé et al. (1998) and Monachesi (1998). We will adopt the same device to deal with the adposition stranding in (2–3), albeit with a twist, in the sense that we adopt the treatment of argument inheritance that is proposed in Van Eynde & Augustinus (2013). A distinctive property of that treatment is that it differentiates complement raising from subject raising: While subject raising is treated in terms of lexical constraints on ARG-ST lists (one for subject-to-subject raising lexemes and one for subject-to-object raising lexemes, see Ginzburg & Sag (2000, p.22)), complement raising is treated in terms of a phrasal constraint on COMPS lists:

- (4)
$$\left[\begin{array}{l} hd-ph \\ SS | LOC | CAT | COMPS \ list \oplus \ \boxed{Z} \\ NONHEAD-DTR | SS | LOC | CAT | COMPS \ \boxed{Z} \end{array} \right]$$

(4) allows a headed phrase to inherit the unsaturated COMPS requirements of its non-head daughter (\boxed{Z}). Application to (2) yields the structure in (5).



The verb's requirement for an adpositional complement ([2]) is immediately saturated, but the adposition's requirement for a nominal complement ([3]) is not. It is appended to the one of the mother and discharged after the addition of *daar* 'there'.¹

The major difference between this version of argument inheritance and the original one is that non-saturated COMPS requirements are not inherited by the selecting verb, but propagated directly from the nonhead daughter to the mother. This implies that they are not included in the ARG-ST list of the selecting verb, which in turn has consequences for the treatment of phenomena which are canonically dealt with in terms of ARG-ST lists, such as binding and passive, see Van Eynde & Augustinus (2013) for exemplification and argumentation.

A general constraint on stranded adpositions is that their complement must precede them, also when it is realized within the PP, as in (6–7).

- (6) *Waar op denk je dat ze wachten?*
where up think you that they wait
 'What do you think they are waiting for?'
- (7) *Ze zegt dat ze soms nog daar aan denkt.*
she says that she sometimes still there on thinks
 'She says that she still thinks about it from time to time.'

Complements which follow the adposition cannot be raised nor extracted. The locational adverb in (8), for instance, follows the adposition (in contrast to the homophonous R-pronoun) and cannot be realized out of the PP.

- (8) a. *Hij beweert dat ze volgens hem niet [van hier] zijn.*
 he claims that they according to him not from here are
 'He claims that they are not from here according to him.'
- b. * *Hij beweert dat ze hier volgens hem niet [van ..] zijn.*
 he claims that they here according to him not from are

To model this we use POSITION, a HEAD feature that is assigned to verbs and adpositions. Its function is to spell out whether the head precedes or follows its complement. Its value is of type *position* and is partitioned into *initial* and *final*. In terms of this dichotomy, the Dutch adpositions come in three types. Some are inherently P-final, such as *mee*, *toe*, *af* and *heen*, some are inherently P-initial, such as *met*, *tot*, *te*, *als*, *sinds*, *sedert*, *wegens*, *tijdens*, *volgens*, ..., and some are used either way, such as *in*, *op*, *aan*, *van*, *uit*, *voor*, The latter's POSITION

¹The SUBJ value of the adposition is the empty list, in accordance with the canonical HPSG treatment of argument marking adpositions.

value is underspecified and resolved contextually.² Employing this feature, the constraint which blocks raising and extraction out of P-initial PPs can be formulated as follows:

$$(9) \left[\begin{array}{l} hd-ph \\ SS | LOC | CAT | HEAD \left[\begin{array}{l} adposition \\ POSITION \quad initial \end{array} \right] \end{array} \right] \rightarrow \left[\begin{array}{l} SS \left[\begin{array}{l} LOC | CAT | COMPS \langle \rangle \\ NONLOC | SLASH \{ \} \end{array} \right] \end{array} \right]$$

In other words, P-initial PPs are islands for complement raising and extraction.

The difference between complement raising and extraction is captured by a similar constraint on V-initial VPs (V-first or V-second).

$$(10) \left[\begin{array}{l} hd-ph \\ SS | LOC | CAT | HEAD \left[\begin{array}{l} verb \\ POSITION \quad initial \end{array} \right] \end{array} \right] \rightarrow \left[\begin{array}{l} SS | LOC | CAT | COMPS \langle \rangle \end{array} \right]$$

(10) blocks complement raising, but not extraction. In v-final clauses, it is the complementizer that acts as the barrier: It selects a clause that is v-final and that has an empty COMPS list.³

In sum, we endorse the analysis of (1) as an instance of complement extraction, but treat (2–3) as instances of complement raising. This choice is further motivated in section 2.

2 Bouma’s arguments revisited

Bouma (2000) considers argument inheritance as an alternative for the uniform extraction analysis, but rejects it for no less than four reasons. We will take them one by one.

1. “Prepositions which do not allow extraction (such as *met*) cannot be associated with an R-pronoun in the Mittelfeld either. If two different mechanisms are used to account for these two phenomena, such generalizations are easily lost.” (Bouma, 2000, p.69) Our answer is threefold. First, the constraint which captures the relevant generalization is simple and straightforward, see (9). Second, the two phenomena must be differentiated anyway, since *er* and *d’r* can be raised but not extracted.

- (11) *Daar/*er/*d’r* hadden we een liedje [... over] willen zingen.
 there/*there had we a song [... about] want sing
 ‘That we had wanted to sing a song about.’

The non-extractability is due to the fact that non-subject constituents in the Vorfeld must be able to bear stress. Pronouns with a clear vowel, such as *daar*, can, hence, occur in the Vorfeld, but pronouns with a mute vowel or without vowel cannot.⁴ Third, there are languages, such as English, which allow complement extraction, but not complement raising. This is due to the fact that English abides by the Empty COMPS Constraint, which requires all phrases to have an empty COMPS list, see Ginzburg & Sag (2000, p.33).

2. “As argument inheritance normally involves the composition of two COMPS lists, R-pronouns would have to be allowed on COMPS, even though they can, apart from a few exceptional cases, never appear in a position following the preposition.” (ibid.) This objection is based on the assumption that Dutch adpositions must precede their complement, but this is not the case. It is not only contradicted by (6–7) but also by adpositions with a full NP complement, as in (12).⁵

²The verbs come in three types as well: The nonfinite forms are inherently v-final, the imperatives are inherently v-initial, and the other finite forms are used either way.

³The selection is modeled in terms of the complementizer’s COMPS list if one treats it as the head of a CP, as in Ginzburg & Sag (2000, p.46). If one treats it as a marker, it is modeled in terms of its HEAD|SPEC value, as in Pollard & Sag (1994, p.44).

⁴The other pronouns which cannot undergo complement extraction include the non-nominative personal pronouns *me*, *je*, *ze*, *het*, *d’r*, *m*, *r*, *t* and the reflexive *zich*, see (Van Eynde, 1999).

⁵Rentier (1993, 116) mentions (12) as a possible counterexample for his claim that Dutch has no postpositions, but then casts doubt on the adpositional status of *in*, claiming that it might be a particle. We do not share this doubt, since the adposition in (12) is clearly distinct from the separable verb particle in *inrijden*, a transitive verb denoting the activity of preparing a vehicle (car, bike, bus, ...) for use on the road. For detailed argumentation that postpositions like the one in (12) are distinct from particles, see Van Riemsdijk (1978, 90–108).

- (12) Zij rijdt de auto [*de garage* in].
 she drives the car [*the garage* in]
 ‘She drives the car into the garage.’

Besides, given that Dutch is predominantly head-final, the existence of P-final PPs is just what one expects, along with the ubiquitous V-final VPs and the A-final predicative APs, as in *hij is [haar fratsen beu]* ‘he is fed up with her antics’.

3. “The set of argument inheritance verbs must now not only contain auxiliaries and modals, but all verbs which select a (prepositional) complement.” (ibid.) This holds for the original argument inheritance approach, but not for our version, since the unsaturated COMPS requirements are not added to those of the verb, but propagated directly from the nonhead-daughter to the mother, see (5).

4. “In an argument inheritance approach, the relationship between valence and syntactically realized arguments has to be one-on-one, and thus there is no room for amalgamation of syntactic functions.” (ibid.) This point concerns the analysis of sentences which contain both a raised R-pronoun and an instance of the nominative (existential) *er*, as in (13–14).

- (13) Er heeft *daarl(er)* een artikel [... over] in de krant gestaan.
 there has *there* an article [... about] in the newspaper stood
 ‘There was an article about that/it in the newspaper.’

- (14) ... dat er *daarl(*er)* een artikel [... over] in de krant heeft gestaan.
 ... that there *there* an article [... about] in the newspaper has stood
 ‘... that there was an article about that/it in the newspaper.’

The raised pronoun is realized in the expected position, if it is a full form, such as *daar*, but if it is the phonologically reduced *er*, it may be omitted in (13) and it must be omitted in (14). Bouma (2000, 73) treats the clauses with a single instance of *er* in terms of function amalgamation, claiming that *er* simultaneously fulfills two functions, i.e. subject of the verb and (extracted) complement of the adposition.⁶ This amalgamation he claims is impossible to model in terms of argument inheritance, since that device does not allow for discrepancies between valence and syntactically realized arguments. We see this differently. In our analysis, there is no function amalgamation. Instead, the first *er* tokens in (13) and (14) have only one function, i.e. subject of the verb. The raised *er* is not identified with that subject, but simply omitted, in the same way as it can be omitted in (15).

- (15) Wie is (*er*) voor? En wie is (*er*) tegen?
 who is (*there*) for? And who is (*there*) against?
 ‘Who is in favor? And who is against?’

The optional omission in (13) is just another instance of a very common kind of mismatch, exemplified also by the intransitive use of verbs like *eat* and *read* and by post-auxiliary ellipsis, as in *yes we can*. The obligatory omission in (14), for its part, is due to a constraint on the PHON(OLOGY) value, blocking adjacent instances of *er*.

3 Conclusions

In contrast to English, which allows adposition stranding only as a result of extraction, Dutch allows it both as a result of extraction, as in (1), and as a result of complement raising, as in (2–3). To model the former one can employ the usual nonlocal devices, as in Rentier (1993) and Bouma (2000), but for the latter we propose an alternative, based on argument inheritance. More specifically, we employ the version of argument inheritance in Van Eynde & Augustinus (2013), which differentiates subject raising from complement raising, and show how the phrasal constraint on complement raising in (4), originally motivated to model Dutch verb clusters, can be used to model the adposition stranding in (2–3) as well. This is not only economical, it also accounts for the fact that languages which do not allow the counterparts of (2–3), such as English, have no verb clusters either. The common cause, it appears, is that such languages abide by the Empty COMPS Constraint.

⁶Technically, this is captured in terms of structure sharing: The LOCAL value of the subject is identified with the SLASH value of the adposition as well as with the BIND value of the verb.

References

- Abeillé, A., D. Godard, P. Miller & I. Sag. 1998. French bounded dependencies. In S. Balari & L. Dini (eds.), *Romance in HPSG*, 1–54. Stanford: CSLI Publications.
- Bennis, H. 1986. *Gaps and Dummies*: Katholieke Hogeschool Tilburg dissertation.
- Bouma, G. 2000. Argument realization and Dutch R-pronouns. In R. Cann, C. Grover & P. Miller (eds.), *Grammatical Interfaces in HPSG*, 51–76. CSLI Publications.
- Bouma, G. & G. 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.
- Broekhuis, H. 2013. *Syntax of Dutch. Adpositions and Adpositional Phrases*. Amsterdam: Amsterdam University Press.
- Ginzburg, J. & I. Sag. 2000. *Interrogative Investigations*. Stanford: CSLI Publications.
- Haeseryn, W., K. Romijn, G. Geerts, J. de Rooij & M. van den Toorn. 1997. *Algemene Nederlandse spraakkunst*. Nijhoff and Wolters Plantyn.
- Hinrichs, E. & T. 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.
- Monachesi, P. 1998. Decomposing Italian clitics. In S. Balari & L. Dini (eds.), *Romance in HPSG*, 305–357. Stanford: CSLI Publications.
- Pollard, C. & I. Sag. 1994. *Head-driven Phrase Structure Grammar*. Stanford/Chicago: CSLI Publications and University of Chicago Press.
- Rentier, G. 1993. Dutch Clitic Pronouns and Preposition Stranding. In *Computational Linguistics in the Netherlands. Papers from the Third CLIN Meeting*, 109–120. Tilburg.
- Van Eynde, F. 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, F. & L. Augustinus. 2013. Why and How to Differentiate Complement Raising from Subject Raising in Dutch. In S. Müller (ed.), *Proceedings of HPSG 2013*, 222–242. Berlin: CSLI Publications.
- Van Riemsdijk, H. 1978. *A Case Study in Syntactic Markedness*. Lisse: The Peter de Ridder Press.