# Bech's Problem, Again: Linearization and Dutch R-Pronouns*

Kathryn Campbell-Kibler

Stanford University

## 1 Introduction

The Dutch R-pronoun system, particularly the weak pronoun *er*, is noted for its complex and bizarre syntactic behavior. R-pronouns as a class participate in an unusual type of non-local dependency. The R-pronoun *er* in particular is unusual. There are three distinct words *er*: the R-pronoun, the quantitative and the expletive. The three of these interact with each other via function overlap, allowing multiple functions to be fulfilled by a single token of the word.

In sections 2 and 3, I present the basic data of the Dutch R-pronouns and go over a previous HPSG-based proposal for treating them using the SLASH feature. I then introduce the concept of linearization theory and explain its advantages over a SLASH-based treatment of R-pronoun dependencies. In section 5.1, I describe the behavior of two other word *er* and their interaction with the R-pronoun *er*. Finally, I present an analysis of the R-pronouns and the other *er*-words based on linearization.

## 2 The data: R-pronouns

R-pronouns make up a paradigm of locative pronouns, consisting of *hier* 'here', *daar* 'there', *er* 'there', *waar* 'where', *overal* 'everywhere'. They are classed as a group on the basis of this semantic similarity, as well as the syntactic behavior described below, which is unique to them. The name "R-pronoun" derives from the "r" that most of them end with.

(1) *Jan woonde dertig jaar in dit huis.*
Jan lived thirty years in this house
'Jan lived in this house for thirty years.'

*I am extremely grateful to Ivan Sag and Arnold Zwicky for their patient help and guidance in the development of this paper. Thanks also to Tom Wasow, Elizabeth Traugott, Frank Richter, Frederik Fouvry and Andreas Kathol for their helpful comments and suggestions, and to Alex Protopesceau and Rob van Glabbeek for their help with sentence judgments.

Proceedings of the 8th International HPSG Conference, Norwegian University of Science and Technology (3-5 August 2001)
Frank van Eynde, Lars Hellan and Dorothee Beermann (editors)
2002.CSLI Publications. http://csli-publications.stanford.edu/

(2) *Jan woonde hier dertig jaar.*
Jan lived here thirty years
'Jan lived here for thirty years.'

(3) *Jan woonde daar dertig jaar.*
Jan lived there thirty years
'Jan lived there for thirty years.'

(4) *Jan woonde er dertig jaar.*
Jan lived there thirty years
'Jan lived there for thirty years.'

*Er* is the light or minor member of the paradigm, having a comparatively bleached semantics and an unstressed phonological representation. For example, it would be infelicitous to utter (4) while pointing at the house in question, and the appearance of *er* in pre-verbal position is restricted.

In addition to carrying locative meaning, R-pronouns serve as the pronoun forms used for non-human objects of prepositions. Dutch prepositions generally take full NPs and the animate pronouns such as *haar* and *mij*, as shown in (5):

(5) *Kim dacht aan Sandy / haar / het probleem / *het.*
Kim thought on Sandy / her / the problem / *it
'Kim thought about Sandy / her / the problem / *it.'

In order to use a pronominal for a non-human referent, a special class of adposition which selects for an R-pronoun must be used. Frequently the R-selector is phonologically identical to the corresponding preposition (as with *aan*), but some are distinguished formally (*e.g. met Kim* vs. *daarmee*). The R-pronoun may appear either directly before the selecting preposition, or directly after the verb:

(6) *Kim dacht gisteren daar aan.*
Kim thought yesterday there on
'Kim thought about it yesterday.'

(7) *Kim dacht daar gisteren aan.*
Kim thought there yesterday on
'Kim thought about it yesterday.'

The R-selector may be a sister to the verb, as above, or it may be embedded within a noun phrase or adjective phrase. In the latter case, the R-pronoun still appears next to the R-selector or to the right of the verb, creating a dependency across the NP boundary:

(8) *Kim is er tevreden mee.*
Kim is there happy with
'Kim is happy with it.'

(9) *Maar ik heb er wat moeite mee.*
But I have there some difficulty with
'But I have some difficulty with it.'

(10) *Hij was er toch wel een beetje trots op.*
He was there still a little bit proud of
'He was still a little bit proud of it.' (Bouma 2000:17)

Like locative prepositional phrases, R-pronouns, whether a locative or a pronominal, participate in locative inversion, and become the subject of a clause, as in (11):

(11) *Daar kocht Jan een huis.*
There bought Jan a house
'Jan bought a house there.'

R-pronouns may also serve as syntactic subjects in so-called "subjectless sentences" like (12):

(12) *Daar wordt gerookt.*
There was smoked
'People were smoking there.'

Although any of the R-pronouns may participate in locative inversion, the unstressed *er* is prevented from occupying the stressed pre-verbal position in a V2 clause on phonological grounds. As a result, it may only be fronted in subordinate clauses. In a V2 clause it may function as a subject when an adverb is fronted to the pre-verbal spot.

The positional flexibility of the R-pronouns presents an analytic challenge in a framework without movement, since they contain dependencies which are non-local but are clause-bounded. In the following section, I describe the HPSG treatment proposed by Bouma (2000), which classes this phenomenon as a long distance dependency and uses the SLASH feature to account for it. I then bring in data relating to the non-R-pronoun uses of *er* and present an alternative analysis which uses linearization domains to provide a more satisfactory account of the R-pronoun *er* and its interactions with the other *er*s.

# 3 A long-distance approach

Bouma (2000) proposes a solution for the prepositional and locative inversion examples which treats all instances of prepositional R-pronouns as long-distance dependencies. R-pronouns are introduced onto the SLASH list and bound off by verbs with R-pronoun adverbials via a 'Bind R-pronoun' principle.

The lexical entry for each preposition specifies whether its ARGUMENT-STRUCTURE (ARG-ST) list contains *synsem*s with CONT values of *prep-obj*, which includes personal pronouns and non-pronouns, as in (13), or selects for items of content *r-pro* and as a result contain a corresponding *synsem* with CONT *r-pro* in SLASH, as in (14):

(13)
$$\begin{bmatrix} \text{PHON} & \textit{met} \\ \text{COMPS} & \langle \boxed{1} \rangle \\ \text{DEPS} & \langle \boxed{1} \rangle \\ \text{ARG-ST} & \left\langle \boxed{1} \begin{bmatrix} \text{HEAD} & \textit{noun}[\textit{acc}] \\ \text{CONT} & \textit{prep-obj} \end{bmatrix} \right\rangle \\ \text{SLASH} & \{\} \end{bmatrix}$$

(14)
$$\begin{bmatrix} \text{PHON} & \textit{mee} \\ \text{COMPS} & \langle\,\rangle \\ \text{DEPS} & \langle \boxed{1} \rangle \\ \text{ARG-ST} & \left\langle \boxed{1} \begin{bmatrix} \textit{gap} & \\ \text{LOC} & \boxed{2} \begin{bmatrix} \text{HEAD} & \textit{noun}[\textit{acc}] \\ \text{CONT} & \textit{r-pro} \end{bmatrix} \\ \text{SLASH} & \{\boxed{2}\} \end{bmatrix} \right\rangle \\ \text{SLASH} & \{\boxed{2}\} \end{bmatrix}$$

This SLASH member may be bound by a filler-gap construction, as in (15), or through an explicit 'Bind R-pronoun' principle, which allows a verb with an R-pronoun somewhere on its DEPS list[1] to bind a corresponding item:

(15) *Waar denkt Kim aan?*
What thinks Kim about
'What does Kim think about?' (Bouma 2000)

(16)
$$\begin{bmatrix} \text{HEAD} & \textit{verb} \end{bmatrix} \Rightarrow \begin{bmatrix} \text{DEPS} & \left\langle \ldots, \begin{bmatrix} \textit{canon} & \\ \text{LOC} & \boxed{1}\begin{bmatrix} \text{CONT} & \textit{r-pro} \end{bmatrix} \end{bmatrix}, \ldots \right\rangle \\ \text{BIND} & \{\boxed{1}\} \end{bmatrix} \vee \begin{bmatrix} \text{BIND} & \{\} \end{bmatrix}$$

[1]Bouma manipulates adverbials through the DEPS list, which forms an intermediate level between the ARG-ST and the valence lists SUBJ and COMPS:

*(i)* Argument Structure Extension

$$\begin{bmatrix} \text{HEAD} & \textit{verb} \end{bmatrix} \Rightarrow \begin{bmatrix} \text{DEPS} & \boxed{1} \oplus \textit{listof('adverbial')} \\ \text{ARG-ST} & \boxed{1} \end{bmatrix}$$

*(ii)* Argument Realization

$$\begin{bmatrix} \text{HEAD} & \textit{verb} \end{bmatrix} \Rightarrow \begin{bmatrix} \text{SUBJ} & \boxed{1} \\ \text{COMPS} & \boxed{2} \ominus \textit{listof(gap-ss)} \\ \text{DEPS} & \boxed{1} \oplus \boxed{2} \end{bmatrix}$$

This provides a single list in which arguments and adverbials are all listed together (cf. Bouma *et al.* (2001)).

The principle in (16) requires that only one such displaced prepositional R-pronoun may be in play in a given sentence. More than one R-pronoun may be present on the DEPS list, but only one may appear in BIND, preventing a verb from binding more than one extracted prepositional R-pronoun.

Bouma's analysis enables R-pronouns to "move" out of the selecting prepositional phrases and noun phrases, in the same way that Wh-words participate in dependencies across constituent boundaries. In some cases this is the correct prediction, for example in (17), in which the R-pronoun is also a Wh-word, and as such may form a dependency which spans a clause boundary:

(17) *Waar denkt Jan dat Marie een standbeeld van gemaakt had?*
What thinks Jan that Marie a statue of made had
'What does Jan think that Marie had made a statue of?'

However, in (18), we see that other R-pronouns may not "leave" the clause. Under a SLASH analysis, a separate mechanism is required to rule out these examples. Thus, collapsing the Wh-word and R-pronoun mechanisms (by using SLASH for both) doesn't explain the distinction between (17) and (18):

(18) **Jan dacht er dat iedereen tevreden mee was.*
Jan thought there that everybody happy with was
'Jan thought that everybody was happy with it.' (Bouma 2000)

*Waar* is both an R-pronoun (and thus selected by a special r-selector) and a Wh-word (and thus travels outside clauses). Bouma treats the R-pronoun phenomenon as a long distance dependency. But it may be more helpful to consider it a medium distance dependency, since it extends beyond a directly local relation, but does not have the range of true long distance dependencies.

In at least some varieties of Dutch, R-pronouns may not participate in dependencies spanning the boundary of a prepositional phrase, as shown in (19):

(19) **Ik had er niet op een argument tegen gerekend.*
I had there not on an argument against counted
'I had not counted on an argument against it.' (van Riemsdijk 1978)

This too requires an independent mechanism under Bouma's analysis.

# 4 Linearization theory

Linearization theory was first introduced by Reape (1993, 1996), who proposed a new feature DOMAIN. DOMAIN values were lists of signs, which had independent ordering constraints, from which the phonology of a sign was derived. I will be drawing on work by Kathol (1995a, 2000), who combines DOMAIN and traditional notions of topological fields to explain a variety of phenomena in German. In his formulation, topological fields are primitives in the ordering constraints of clause-level DOMAIN lists. This approach allows insights into German (and Dutch) syntax which depend on linear order rather than tree structure to be captured in HPSG. I also adopt the division between signs and constructions introduced in Donohue and Sag (1999) and further elaborated in Sag (2001).

(20) gives the Ordering Principle proposed by Reape (1996), which links the phonology of DOMAIN list members with the phonology of the sign itself:

(20) $$\left[\textit{sign}\right] \Rightarrow \begin{bmatrix} \text{PHON} & \phi_1 \oplus, \ldots, \oplus \phi_n \\ \text{DOM} & \left\langle \left[\text{PHON} \quad \phi_1\right], \ldots, \left[\text{PHON} \quad \phi_n\right] \right\rangle \end{bmatrix}$$

(Donohue and Sag 1999)

This is the engine behind the linearization concept, allowing manipulations of the DOMAIN list to be reflected in the ordering of phonological material. In this way the phonological string may be reshuffled without challenging the tree structure relations.

Reape's DOMAIN list is somewhat awkward as the lists are populated by signs, which themselves may be phrases with daughters. Thus each DOMAIN element on a sign's list is in essence an entire syntactic tree. In order to simplify the picture, (21) and (22) show a reworking of sign and construction such that signs are simple triples of phonology, syntax/semantics and a domain list:

(21) $$\begin{bmatrix} \textit{sign} & \\ \text{PHON} & \textit{list(speech-segment)} \\ \text{SYNSEM} & \textit{synsem} \\ \text{DOMAIN} & \textit{list(sign)} \end{bmatrix}$$

(22) $$\begin{bmatrix} \textit{construction} & \\ \text{MOTHER} & \textit{sign} \\ \text{DTRS} & \textit{list(sign)} \end{bmatrix}$$

(Donohue and Sag 1999)

Rather than incorporating daughter relations into signs, this structure is built entirely with constructions (Donohue and Sag 1999).

Reape proposes that in the formation of a new DOMAIN list, the DOMAIN values of the signs may either be appended into a larger list (DOMAIN union) or compacted. Kathol expands on this idea to propose that constructions may be partially compacting, according to the schema given in (23):

(23) $$\textit{p-compaction}(\boxed{1}, \boxed{2}, \boxed{3}) \equiv \boxed{1}\begin{bmatrix} \textit{sign} & \\ \text{SS} & \boxed{4} \\ \text{DOM} & \boxed{6} \end{bmatrix} \wedge \boxed{2}\left\langle \begin{bmatrix} \textit{dom-obj} & \\ \text{SS} & \boxed{4} \\ \text{PHON} & \boxed{7} \end{bmatrix} \right\rangle$$

$$\wedge \textit{shuffle}(\boxed{5}, \boxed{3}, \boxed{6}) \wedge \textit{join}_{phon}(\boxed{5}, \boxed{7})$$

(Kathol 1995:307)

Note that these HPSG-specific ideas draw on the earlier analyses of Pullum (1982) and Zwicky (1986), who introduced the idea of liberating constructions.

Allowing NPs to be generally compacting but liberating with respect to R-pronouns would allow an account of the "medium distance dependency" behavior discussed in the previous section. The clause-bounding of such liberation would be predicted, since it is at the clause level that the linear precedence rules apply which implement topological fields.

The details of this analysis will be presented in section 6. First, however, I will introduce some additional data on the interaction of the R-pronoun *er* and related phenomena, an interaction which will provide additional evidence for the appropriateness of a linearization based treatment.

# 5 More data

## 5.1 Quantitative and expletive *er*

In addition to the R-pronoun *er*, there are two other functions that the word *er* serves in Dutch: the quantitative (or partitive), which licenses a gapped noun or noun phrase elsewhere in the clause, and the expletive, similar to English's existential expletive *there*.

The quantitative *er* licenses a missing head noun from a quantified NP in certain constructions, like (24) and (25):

(24) *Hoeveel boeken heeft u? Ik heb er/*$\phi$ twee.*
How-many books have you I have there two
'How many books do you have?' 'I have two.'

(25) *Ik heb geen boeken over letterkunde, maar ik heb er wel heel wat over taalkunde.*
I have no books about literature but I have there indeed very many about linguistics.
'I don't have any books about literature, but I have many about linguistics.'
(Geerts *et al.* 1984)

Like the R-pronoun *er*, the quantitative *er* is not found in pre-verbal position in V2 clauses. Since its badness there is not influenced by prosodic constraints, this appears to be a syntactic rather than phonological fact.

Unlike the R-pronoun, the quantitative *er* does not participate in a larger paradigm. In the examples where *er* is used, only *er* appears. However, in addition to those examples, there are two constructions in which a count noun is gapped. In these cases not only is *er* not required, but it is impossible. The first is gapping, as in (26); the second involves non-quantity adjectives, as in (27):

(26) *Jan heeft vier boeken en ik drie.*
Jan has four books and I three
'Jan has four books and I have three.'

(27) *Ik heb (*er) nog drie gele.*
I have (there) still three yellow
'I still have three yellow ones.' (van Riemsdijk 1978)

While a treatment of these competing constructions is beyond the scope of this paper, its study may be useful in determining the exact meaning and use of the quantitative *er* construction.

The expletive *er* is similar to the expletive *there* in English, a semantically empty element which allows verbal arguments to be demoted in the argument hierarchy. Like the English version, the remaining arguments are affected by the definiteness constraint, requiring that they be semantically indefinite. Unlike its English counterpart, *er* may appear with a wide variety of verbs, including transitives:

(28) *Er blaft een hond.*
There barks a dog
'There is a dog barking.'

(29) *Ik denk dat er een hond blaft.*
I think that there a dog barks
'I think that there is a dog barking.'

In examples like (29), the expletive may be difficult to distinguish from an R-pronoun that has undergone locative inversion. However, the expletive, unlike the R-pronoun, is perfectly fine in pre-verbal position, as in (28).

## 5.2 Co-occurrence of functions

In the event that two or more constructions involving *er* are combined within a single clause, there are two possible results: doubling, in which both tokens of *er* are pronounced, as in (30), and overlap, where only one token appears, covering multiple functions (31) and (32):

(30) *Het schijnt dat er één er twee gelezen heeft.*
It seems that there one there two read has
'It seems that one of them has read two of those.' (Kerstens 2000b)

(31) *Vandaag staan er weer twee artikels over in.*
Today stand there again two articles about in
'There are two articles about it in there again today.' (Geerts *et al.* 1984)

(32) *Ik heb er twee op gelegd.*
I have there two on placed
'I've put two on there.' (Kerstens 2000b)

The co-occurrence facts may be expressed in three generalizations:

1. A token of the R-pronoun *er* may not appear in the same clause as (double with) another token of *er*.

2. Two tokens of *er* in succession are not permitted.

3. Overlap is permitted if the tokens could be adjacent, were they both present.

# 6 Analysis

## 6.1 R-pronoun "movement"

I follow Bennis (1986) in classifying R-pronouns as PPs, not NPs. They are like full locative prepositional phrases in meaning: they are selected by the same verbs and prepositions and participate in locative inversion. This classification does not bear significantly on the central aspects of the analysis.

In order to distinguish the R-pronouns and the other *er*s from each other and from other words, I introduce two new binary head features for prepositions, R-PRO and R. The R-PRO value is positive for the R-pronouns (*er, hier, daar, overal, waar*) and negative for everything else. The R value is positive in words which participate in the non-local dependencies discussed above (and distinguished from long distance dependencies like Wh-words), and in the function overlap phenomena. These are the R-pronouns except *waar*, and the quantitative and expletive *er*. The table in (33) shows this distribution, and (34) gives the lexical entry for the R-pronoun *er*:

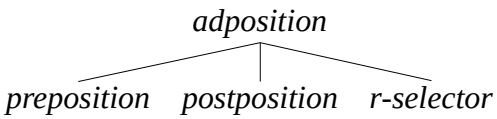
(33)

| | R-PRO + | R-PRO − |
|---|---|---|
| R + | *er (r-pro), daar, hier* | *er (quant.), er (expl.)* |
| R − | *waar* | *aan, met* |

(34)
$$\begin{bmatrix} \text{PHON} & \langle er \rangle \\ \text{HEAD} & \begin{bmatrix} prep \\ \text{R-PRO} & + \\ \text{R} & + \end{bmatrix} \end{bmatrix}$$

The type *adposition* is divided into three subtypes:

(35)




Postpositions are not dealt with here, but they are included on the basis of the discussion regarding them in van Riemsdijk (1978). Example lexical entries for a preposition and an r-selector are given in (36) and (37):

(36)
$$\begin{bmatrix} \text{PHON} & \langle met \rangle \\ \text{HEAD} & \begin{bmatrix} prep \\ \text{R-PRO} & - \\ \text{R} & - \end{bmatrix} \\ \text{COMPS} & \left\langle \begin{bmatrix} \text{HEAD} & noun \end{bmatrix} \right\rangle \end{bmatrix}$$

(37)
$$\begin{bmatrix} \text{PHON} & \langle mee \rangle \\ \text{HEAD} & r\text{-}selector \\ \text{COMPS} & \left\langle \begin{bmatrix} \text{HEAD} & \begin{bmatrix} prep \\ \text{R-PRO} & + \\ \text{R} & + \end{bmatrix} \end{bmatrix} \right\rangle \end{bmatrix}$$

Normal prepositions which are [R-PRO −] and [R −] select for noun phrases which are [R-PRO−]. R-selectors, on the other hand, only select R-pronouns.

These two types are also distinguished by the level of compaction in the formation of phrases. Below are constraints on the the constructions *prepositional-construction* and *r-selector-construction*, detailing the level of compaction appropriate to each.

(38)

$$\textit{prep-cx} \Rightarrow \left[\begin{array}{ll} \text{MOTHER} & \left[\begin{array}{ll} \text{SS} & \boxed{0} \\ \text{DOM} & \left\langle \left[\begin{array}{ll} \text{SS} & \boxed{0} \\ \text{DOM} & \boxed{D}\textit{list}\left(\left[\text{R} \quad -\right]\right) \end{array}\right] \right\rangle \oplus \boxed{C}\ \textit{list}\left(\left[\begin{array}{ll} \textit{rel-cl} \\ \text{TOPO} \quad \textit{nf} \end{array}\right]\right) \end{array}\right] \\ \text{DTRS} & \left\langle \left[\text{DOM} \quad \boxed{A}\right], \left[\text{DOM} \quad \boxed{B}\right] \right\rangle \end{array}\right]$$

where $\boxed{D} \bigcirc \boxed{C} = \boxed{A} \bigcirc \boxed{B}$

(39)

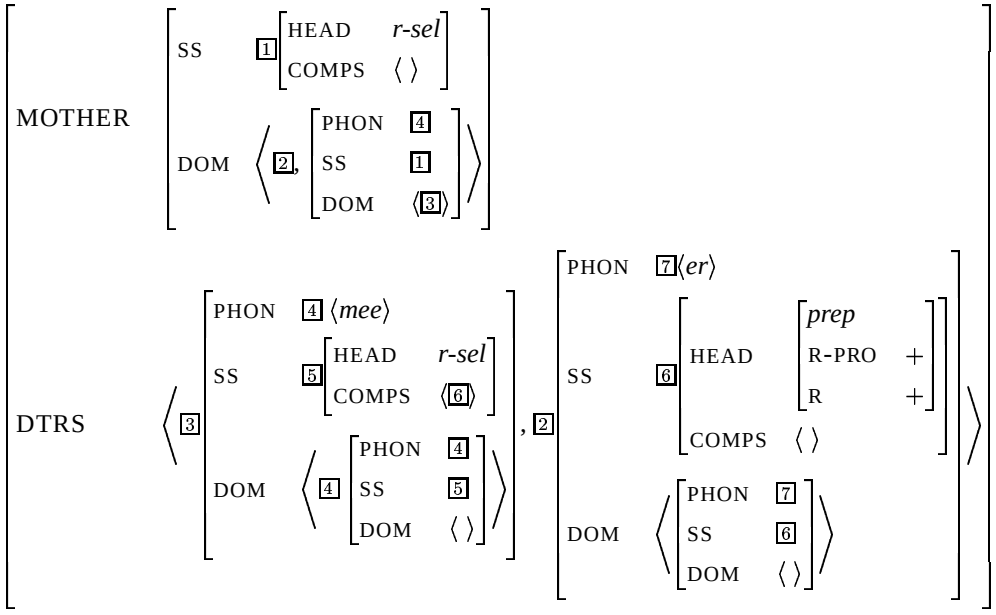
$$\textit{r-selector-cx} \Rightarrow \left[\begin{array}{ll} \text{MOTHER} & \left[\begin{array}{ll} \text{SS} & \boxed{0} \\ \text{DOM} & \left\langle \boxed{2}, \left[\begin{array}{ll} \text{SS} & \boxed{0} \\ \text{DOM} & \langle \boxed{1} \rangle \end{array}\right] \right\rangle \end{array}\right] \\ \text{DTRS} & \left\langle \boxed{1}\left[\text{HEAD} \quad \textit{r-selector}\right], \boxed{2}\left[\text{HEAD} \quad \left[\begin{array}{ll} \textit{prep} \\ \text{R-PRO} \quad + \\ \text{R} \quad + \end{array}\right]\right] \right\rangle \end{array}\right]$$

Prepositional phrases are compacted, except for relative clauses which are liberated and marked as [TOPO *nf*] (since they are mandatorily extraposed (van Riemsdijk 1978)). R-selectors, on the other hand, are completely liberating.

(40) and (41) give example constructions for the phrases *met Sandy* and *ermee*, to show the constraints in practice:

(40)

$$\left[\begin{array}{ll} \text{MOTHER} & \left[\begin{array}{ll} \text{SS} & \boxed{3}\left[\begin{array}{ll} \text{HEAD} \quad \textit{prep} \\ \text{COMPS} \quad \langle\,\rangle \end{array}\right] \\ \text{DOM} & \left\langle \left[\begin{array}{ll} \text{PHON} \quad \boxed{1} \oplus \boxed{2} \\ \text{SS} \quad \boxed{3} \\ \text{DOM} \quad \langle \boxed{4}, \boxed{5} \rangle \end{array}\right] \right\rangle \end{array}\right] \\ \text{DTRS} & \left\langle \left[\begin{array}{ll} \text{PHON} \quad \boxed{1}\langle \textit{met} \rangle \\ \text{SS} \quad \boxed{6}\left[\begin{array}{ll} \text{HEAD} \quad \textit{prep} \\ \text{COMPS} \quad \langle \boxed{7} \rangle \end{array}\right] \\ \text{DOM} \quad \left\langle \boxed{4}\left[\begin{array}{ll} \text{PHON} \quad \boxed{1} \\ \text{SS} \quad \boxed{6} \\ \text{DOM} \quad \langle\,\rangle \end{array}\right] \right\rangle \end{array}\right], \left[\begin{array}{ll} \text{PHON} \quad \boxed{2}\langle \textit{Sandy} \rangle \\ \text{SS} \quad \boxed{7}\left[\begin{array}{ll} \text{HEAD} \quad \textit{noun} \\ \text{COMPS} \quad \langle\,\rangle \end{array}\right] \\ \text{DOM} \quad \left\langle \boxed{5}\left[\begin{array}{ll} \text{PHON} \quad \boxed{2} \\ \text{SS} \quad \boxed{7} \\ \text{DOM} \quad \langle\,\rangle \end{array}\right] \right\rangle \end{array}\right] \right\rangle \end{array}\right]$$

(41)



The difference between prepositional phrases and noun phrases shown in the constraints above predict that R-pronouns will not be able to move outside of a larger prepositional phrase. It also makes a prediction regarding the interaction of PP and NP boundaries and the co-occurrence phenomena introduced in section 5.2. The analysis of these facts will be discussed in depth in section 6.3, but the liberation of the R-pronouns by NPs is necessary for the behavior described. Thus the constraints as written predict divergent behavior for R-pronouns within a PP than those within NPs. For example, an R-pronoun within a PP may appear in the same clause as another token of the same word, as in (42):

(42) *Vandaag staan er weer twee artikels over het probleem ermee in (de krant).*
Today stands there again two articles about the problem there-with in (the newspaper)
'There are two articles about the problem with it in the newspaper again today.'
(Geerts *et al.* 1984)

The permeability of PPs to R-pronouns is a locus of variation which has not been fully explored. Some speakers accept sentences like (19), in which an R-pronoun dependency spans a PP boundary, and others reject it roundly. Variation on this point is easily captured by giving PPs the structure given below for NPs, in which R-pronouns are liberated. This is a minor change which is relatively unsurprising for there to be differences on. My analysis does make the claim that speakers who find (19) to be bad ought to find (42) good, and vice versa. More research is necessary to determine whether this claim is correct.

Noun and adjective phrases are generally compacting, but both of these liberate elements with positive R values, leaving them free to rearrange at the next level:

(43)
$$\textit{noun-cx} \Rightarrow \left[\begin{array}{ll} \text{MOTHER} & \left[\begin{array}{ll} \text{SS} & \boxed{0} \\ \text{DOM} & \left\langle \left[\begin{array}{ll} \text{SS} & \boxed{0} \\ \text{DOM} & \boxed{D} \end{array}\right] \right\rangle \bigcirc \boxed{C}\ \textit{list}\left( \left[\begin{array}{ll} \text{HEAD} & \left[\begin{array}{ll} \textit{prep} & \\ \text{R} & + \end{array}\right] \end{array}\right) \end{array}\right] \\ \text{DTRS} & \left\langle \left[\text{DOM}\ \ \boxed{A}\right], \left[\text{DOM}\ \ \boxed{B}\right] \right\rangle \end{array}\right]$$

where $\boxed{D} \bigcirc \boxed{C} = \boxed{A} \bigcirc \boxed{B}$

This can be seen at work in the construction for the noun phrase *er, twee artikels over*:

(44)
$$\left[\begin{array}{ll} \text{MTHR} & \left[\begin{array}{ll} \text{SS} & \boxed{1}\left[\begin{array}{ll} \text{HEAD} & \textit{noun} \\ \text{COMPS} & \langle\,\rangle \\ \text{SPR} & \langle\,\rangle \end{array}\right] \\ \text{DOM} & \left\langle \boxed{2}, \left[\begin{array}{ll} \text{PHON} & \boxed{3}\langle \textit{twee}\rangle \oplus \boxed{4}\langle \textit{artikels over}\rangle \\ \text{SS} & \boxed{1} \\ \text{DOM} & \langle \boxed{5}, \boxed{6}\rangle \end{array}\right] \right\rangle \end{array}\right] \\ \text{DTRS} & \left\langle \left[\begin{array}{ll} \text{PHON} & \boxed{3} \\ \text{SS} & \boxed{7} \\ \text{DOM} & \left\langle \left[\begin{array}{ll} \text{PHON} & \boxed{3} \\ \text{SS} & \boxed{7} \\ \text{DOM} & \langle\,\rangle \end{array}\right] \right\rangle \end{array}\right], \left[\begin{array}{ll} \text{SS} & \boxed{8}\left[\begin{array}{ll} \text{HEAD} & \textit{noun} \\ \text{COMPS} & \langle\,\rangle \\ \text{SPR} & \langle \boxed{7}\rangle \end{array}\right] \\ \text{DOM} & \left\langle \left[\begin{array}{ll} \text{PHON} & \boxed{4} \\ \text{SS} & \boxed{8} \end{array}\right], \boxed{2}\left[\begin{array}{ll} \text{PHON} & \langle \textit{er}\rangle \\ \text{SS}|\text{HEAD} & \left[\begin{array}{ll} \textit{prep} & \\ \text{R-PRO} & + \\ \text{R} & + \end{array}\right] \end{array}\right] \right\rangle \end{array}\right] \right\rangle \end{array}\right]$$

The constraints given in this section are somewhat complicated, and at present must be stipulated directly. This analysis is meant to fit into a larger understanding of various phenomena in Dutch for which a linearization approach is appropriate, for example extraposition and V2 clause fronting. I am hopeful that future research will allow these constraints to be reduced to a more general formulation.

## 6.2 Accounting for Q-er and E-er

I will not be exploring either of these in depth, and the quantitative *er* in particular has a complex syntax of its own. My treatment is aimed at capturing the basic behavior of the two words, and particularly their interaction with the R-pronoun. Further refinement will undoubtedly be necessary to capture the nuances of their individual syntax.

Like the R-pronouns, the quantitative *er* is selected by a special class of words, in this case nouns produced from quantity adjectives through a lexical rule. Unlike the R-pronoun *er*, the quantitative *er* is syntactically restricted to the middle field of the clause. The quantitative *er* is selected by a noun, usually derived from number words, as the example in (45) shows:

(45)
$$\begin{bmatrix} \text{PHON} & \langle \textit{twee} \rangle \\ \text{HEAD} & \textit{noun} \\ \text{ARG-ST} & \left\langle \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \text{PREP} & \\ \text{R-PRO} & - \\ \text{R} & + \end{bmatrix} \\ \text{TOPO} & \textit{mf} \end{bmatrix} \right\rangle \end{bmatrix}$$

Like the English existential, the expletive is selected for by a verb. (46) shows a lexical rule which takes a regular verb and introduces an expletive *er* as the subject.

(46)
$$\begin{bmatrix} \text{HEAD} & \textit{verb} \\ \text{SUBJ} & \boxed{A} \\ \text{COMPS} & \boxed{B} \end{bmatrix} \Rightarrow \begin{bmatrix} \text{HEAD} & \textit{existential-verb} \\ \text{SUBJ} & \left\langle \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \text{PREP} & \\ \text{R-PRO} & - \\ \text{R} & + \end{bmatrix} \end{bmatrix} \right\rangle \\ \text{COMPS} & \boxed{A} \oplus \boxed{B} \end{bmatrix}$$

## 6.3 Co-occurrence

The co-occurrence facts are formalized using two constraints on the DOMAIN feature of the finite clause, (47) and (48), and a haplology rule as an alternation of the ordering principle, which states that in the case of several consecutive *er*s in the DOMAIN, only one actually appears in the phonology, in (49).[2]

(47) and (48) require that an R-pronoun and an element which is [R +] in the same domain must be located next to each other. The constraint is essentially the proximity portion of immediate precedence as given in Kathol (1995). Rather than distinguishing precedence and immediate precedence as two independent types of constraints, precedence and proximity may be more appropriate as primitives, with immediate precedence as merely the combination of the two.

(47)
$$\neg \left( \begin{bmatrix} \text{PHON} & \alpha \\ \text{R-PRO} & + \end{bmatrix} \prec X \prec \begin{bmatrix} \text{PHON} & \alpha \\ \text{R} & + \end{bmatrix} \right)$$

(48)
$$\neg \left( \begin{bmatrix} \text{PHON} & \alpha \\ \text{R} & + \end{bmatrix} \prec X \prec \begin{bmatrix} \text{PHON} & \alpha \\ \text{R-PRO} & + \end{bmatrix} \right)$$

This proximity constraint applies to the DOMAIN list of a clause, and requires that R-pronouns must appear adjacent to other signs with positive R values and identical phonology values. As a result, all instances of doubling are ruled out unless the tokens are adjacent, or unless none of the tokens concerned are R-pronouns. In the latter case, doubling involving only quantitative and expletive *er* are rightly classed as grammatical. It remains only to rule out doubling of adjacent

[2] *Er* has a phonological alternation between *er* and *d'r*, which I am for the present ignoring. This schema would need to be modified slightly to include that variation.

tokens, and to allow overlap in appropriate cases. Overlap is introduced at the phonological level, by a variation of Reape's Ordering Principle:

(49) $$\begin{bmatrix} \text{PHON} & \phi_1 \oplus \ldots \oplus \textit{er} \oplus \ldots \oplus \phi_n \\ \text{DOM} & \left\langle \begin{bmatrix} \text{PHON} & \phi_1 \end{bmatrix}, \ldots, \oplus \textit{ne-list}(\begin{bmatrix} \text{PHON} & \langle \textit{er} \rangle \\ \text{R} & + \end{bmatrix}) \oplus \ldots, \begin{bmatrix} \text{PHON} & \phi_n \end{bmatrix} \right\rangle \end{bmatrix}$$

To rule out consecutive *er*s we need a further constraint on the phonology of the clause itself, disallowing any examples of */#er##er#/*. Including this as a purely phonological constraint (as opposed to the DOMAIN, for example) reflects the fact that unlike the other aspects of *er* syntax, the constraint on consecutive tokens does not distinguish between the different types. It is also likely that the constraint has a phonological motivation, as there is a strong cross-linguistic pattern for languages to avoid consecutive tokens of an light or unstressed function word (Menn and MacWhinney 1984).

# 7 Conclusion

Using linearization and partial compaction, I provide an account of R-pronoun "movement" which predicts the clause-boundedness of the dependency and allows for a straightforward treatment of the idiosyncrasy of prepositional phrases. As such, this analysis is more comprehensive than the previous SLASH-based HPSG proposal by Bouma (2000) and requires less machinery than earlier transformational accounts.

Manipulating the R-pronoun behavior on the DOMAIN list also allows the analysis to expand to cover the interaction behavior of the R-pronoun *er* with the quantitative and expletive *er*s. The phenomenon of overlap or function amalgamation is explained through linear precedence rules, an alternation to the constraint relating DOMAIN and PHONOLOGY to allow haplology, and a phonological restriction on a sequence of repeated *er*s.

Potential topics for future work include a closer look at the quantitative and expletive *er*s in their own right, and a comprehensive description of the dialectal variation which impacts virtually every aspect of these structures.

# References


BECH, GUNNAR. 1952. Über das niederländische Adverbialpronomen ‘er’. In *Travaux du cercle linguistique de Copenhague*, volume 8, 5–32. Amsterdam: Taalinstituut de natuurmethode.

BENNIS, HANS. 1986. *Gaps and Dummies*. Dordrecht: Foris Publications.

BOUMA, GOSSE. 2000. Argument realization principles and Dutch R-pronouns: Solving Bech’s problem without movement or deletion. In *A collection of papers on Head-driven Phrase Sructure Grammar*, ed. by Claire Grover Ronnie Cann and Philip Miller. Stanford University.

——, ROB MALOUF, and IVAN A. SAG. 2001. Satisfying constraints on extraction and adjunction. *Natural Language and Linguistic Theory* .

DONOHUE, CATHRYN, and IVAN SAG, 1999. Domains in Warlpiri. Paper presented at HPSG-99 – University of Edinburgh.

GEERTS, G., W. HAESERYN, J. DE ROOIJ, and M. C. VAN DEN TOORN. 1984. *Algemene Nederlandse Spraakkunst*. Groningen: Wolters-Noordhoff.

KATHOL, ANDREAS, 1995. *Linearization-Based German Syntax*. Ohio State University dissertation.

——. 2000. *Linear Syntax*. Oxford: Oxford University Press.

KERSTENS, JOHAN G., 2000b. Dat eigenaardige woordje er. http://www.let.uu.nl/~nltaal/er/, Universiteit Utrecht, Faculteit der Letteren.

MENN, LISA, and BRIAN MACWHINNEY. 1984. The repeated morph constraint: Toward an explanation. *Language* 60.519–41.

PULLUM, GEOFFREY. 1982. Free word order and phrase structure rules. In *Proceedings of the 12 annual meeting of the Northeast Linguistic Society*, 209–220, Amherst. Graduate Linguistics Student Association.

REAPE, MIKE, 1993. *A Formal Theory of Word Order: A Case Study in West Germanic*. University of Edinburgh dissertation.

——. 1996. Getting things in order. In *Discontinuous Constituency*, ed. by Harry Bunt and Arthur van Horck, number 6 in Natural language processing, 209–253. Berlin, New York: Mouton de Gruyter.

RENTIER, GERRIT M. 1994. A lexicalist approach to dutch preposition stranding. In *KONVENS ’94*, ed. by Harald Trost, 280–289, Berlin. Springer-Verlag. corrected version: Extraction in Dutch with Lexical Rules.

SAG, IVAN, 2001. Six dimensions of natural language locality. paper presented at HPSG 2001. Trondheim, Norway.

VAN RIEMSDIJK, H. C. 1978. *A case study in syntactic markedness: The binding nature of prepositional phrases*. Lisse: The Peter de Ridder Press.

ZWICKY, ARNOLD. 1986. Concatenation and liberation. In *Papers from the 22nd regional meeting, Chicago Linguistic Society*, 65–74, Chicago. Chicago Linguistic Society.