Syntactic haplology and the Dutch proform *er*

Gert Webelhuth [iD]

University of Frankfurt

Olivier Bonami [iD]

Universite de Paris, LLF, CNRS

Abstract

Dutch has four pronouns 'er' which show an intriguing pattern of syntactic haplology when a finite verb has more than one 'er' dependent. We present a theory that captures this pattern by relying on two central aspects of HPSG: (i) the distinction between ARG-ST and COMPS and (ii) the distinction between canonical and non-canonical synsem objects. No deletion rules of the kind used in transformational analyses of 'er' are necessary.

1 Introduction

Dutch has four expressions spelled 'er' with different syntactic and semantic functions and syntactic distributions that display unusual and intriguing interdependencies. We give an overview of the major data and show that it can be captured through the interaction of a small number of constraints on argument realization[1]

2 The Data

The sentences in (1) each contain a single example of each type of *er*.[2] The first example features existential *er*, which cooccurs with indefinite subjects and is the only *er* that can fill the first position of a Dutch main clause. In (1b), pronominal *er* expresses the obligatory complement of the preposition *op*. The example shows that pronominal *er* does not need to be adjacent to its selector. (1c) contains *er* in its function as a locational expression, comparable to the English referential locational adverb *there*. Finally, (1d) illustrates quantitative *er*: it serves as the complement of the numeral *drie* in this example and performs a function similar to the partitive elements *en* in French or *ne* in Italian.

(1)   a.   *Er$_X$* loopt een man op straat.
           there walks a   man in the.street

      b.   Jan wacht *er$_P$* al   tijden **op**.
           Jan waits there for ages   for

      c.   Jan staat   *er$_L$* al.
           Jan stands there already

      d.   Jan heeft *er$_Q$* [$_{NP}$ **drie** [*e*]]
           Jan has   there   three

---

[2]We use the following system to label the four *er*s: *er$_X$* = existential, *er$_P$* = pronominal, *er$_Q$* = quantitative, *er$_L$* = locational.

There are many previous analyses of *er* in the literature. Space limitations make it impossible to do anything other than listing the most important ones here: Bech (1952), van Riemsdijk (1978), Bennis (1986), Odijk (1993), and Broekhuis (2013).

As non-native speakers of Dutch we are faced with the problem that on *er* "conflicting judgments can be found in the literature" (Broekhuis (2013, p. 338)). We chose to handle this problem by citing the data and judgments of just a single author. With the exception of (12), which was supplied to us in a personal communication by Hans Broekhuis, all examples are drawn from Broekhuis (2013), which is an extremely comprehensive and detailed treatment of *er*.

We decided to develop a new analysis of *er* in HPSG, as we felt that we can improve on existing HPSG analyses. Bouma (2001), Van Eynde & Augustinus (2014), and Van Eynde (2019) all do not cover quantitative *er*, which behaves differently from the three remaining *er*s, as will be demonstrated below. Campbell-Kibler (2001) was meant to account for different judgments than those considered here. Moreover, since the author draws examples from different works in the literature, it is unclear that this data reflects a consistent set of judgments.

2.1 Linear structure of finite clauses in Dutch

We assume that Dutch sentences can be analyzed as consisting of a number of topological (= linear) fields, as follows:[3]

			C	Middle field	Verb(s)
Subordinate clause:					
Main clause:		Prefield	V_{finite}	Middle field	(Verb(s))

The prefield is limited to a single constituent whereas the middle field can contain zero, one, or several constituents. The next two sections will describe the distribution of *er* in these two fields.

2.2 Clauses without existential *er* in the prefield

As the linear field schema above illustrates, both main and subordinate clauses contain a middle field. The present section deals with *er* in subordinate clauses and in those main clauses whose prefield is *not* filled by er_X, i.e. with main clauses like (1b)-(1d). The middle field in these kinds of sentences satisfies the simple generalization that it can contain at most one overt *er*.[4]

(2) illustrates that an overt existential *er* cannot cooccur with any of the three other other *er*s:

---

[3]On topological fields, see Drach (1937), Reis (1980), Höhle (1983), Höhle (1986). Note that our analysis in this article is restricted to the occurrences of *er* in finite sentences.

[4]Neeleman & van de Koot (2006) accept certain sentences with two *er*s in the middle field as long as the *er*s are not adjacent. Hans Broekhuis and the native speakers we were able to consult consider these examples ungrammatical (personal communication). The theory developed below is only meant to cover Broekhuis' judgments.

(2) a. * dat er_X er_L gedanst wordt.
      that there there danced is
      Intended reading: 'People are dancing there.'

   b. * dat er_X er_P over gesproken wordt.
      that there there about spoken   is
      'Intended reading: 'People are talking about it.''

   c. * dat er_X er_Q [$_{NP}$ twee e]] gestolen zijn.
      that there there   two   stolen have.been
      'Intended reading: 'Two [e.g., computers] have been stolen.''

All three sentences become grammatical if one of the two *er*s is dropped.

(3a)-(3b) show that overt pronominal *er* cannot cooccur with an overt locational or quantitative *er* and (4) provides evidence that the remaining potential combination of overt *er*s is impossible as well:

(3) a. * dat Jan er_P er_L over praatte.
      that Jan there there about talked
      'that Jan talked about it there.'

   b. * dat Jan er_P er_Q drie in stopte.
      that Jan there there three into put
      'that Jan put three [e.g., cigars] in it.'

(4)   * dat Jan er_Q er_L [$_{NP}$ twee [e]] gezien heeft.
      that Jan there there   two   seen has
      'that Jan saw two [e.g., rats] there.'

Again, these sentences become grammatical, if only a single overt *er* appears.

2.3 Sentences with an overt and an implicit *er*

The data presented in the previous subsection jointly illustrate the generalization that in the ideolect studied here the middle field of sentences without an expletive *er* in the prefield can contain only a single overt *er*. Interestingly, however, when one overt *er* appears, one or more additional *er*s can be understood. The sentences in (5) demonstrate this. The existential subordinate clause (5a) contains an overt er_X, an indefinite subject NP, and an object PP.

(5) a. dat er_X gisteren [$_{NP}$ drie potloden] [$_{PP}$ op tafel]   lagen.
      that there yesterday   three pencils   on the.table lay
      'that there were three pencils lying on the table yesterday.'

   b. dat er_{XP} gisteren [$_{NP}$ drie potloden] [$_{PP}$ op] lagen.
      that there yesterday   three pencils   on lay
      'that there were three pencils lying on it yesterday.'

    c.   dat er_{XQ} gisteren [$_{NP}$ drie] [$_{PP}$ op tafel]   lagen.
        that there yesterday   three   on the.table lay

        'that there were three lying on the table yesterday.'

    d.   dat er_{XL} veel   mensen wonen.
        that there many people live

        'that many people live there.'

In (5b), the object of the preposition *op* gets a deictic interpretation 'there', even though the object is unexpressed. If the sentence did not contain an expletive *er*, then the object of the preposition would need to be expressed as pronominal *er*. In (5c) the quantitative *er* of the partitive NP *drie* remains implicit. (5d), finally, illustrates the case where the adverbial complement of the verb *wonen* with the sense of 'reside' can remain unexpressed in the presence of an overt expletive *er* in the middle field.

Intriguingly, but in light of the examples just provided perhaps no longer surprising, it is also possible for a single pronominal *er* to represent the objects of two separate prepositions in a sentence. This is shown in (6). The first sentence contains two PPs with non-pronominal NPs. The second and third examples show that pronomional *er* can serve as the object of each preposition:[5]

(6)   a.   Jan heeft de sleutel [**met een tang**]   [**uit het slot**] gehaald
            Jan has   the key   with a pair.of.tongs out.of the lock taken

            'Jan took the key out of the lock with pliers.'

    b.   Jan heeft er_P de sleutel [**mee**] [uit het slot] gehaald.

    c.   Jan heeft er_P de sleutel [met een tang] [**uit**] gehaald.

    d.   Jan heeft er_{PP} de sleutel [**mee**] [**uit**]] gehaald.

    e. * Jan heeft er_P er_P de sleutel [**mee**] [**uit**] gehaald.

(6d)-(6e) demonstrate what happens when both prepositions are stranded at the same time: the objects of the prepositions must be represented by a single *er*, as two *er*s in the Dutch middle field are forbidden.

The same pattern occurs with quantitative *er*. The second conjunct of the following example contains two partitive NPs but only a single quantitative *er*:[6]

(7)   Iedere student heeft een onvoldoende   gekregen …
        every student has   an unsatisfactory mark gotten

        'Every student got an unsatisfactory mark …'

    a.   … en [$_{NP}$ **drie** e] hebben er_Q zelfs [$_{NP}$ **twee** e].
           and   three   have   there even   two

        '… and three even got two.'

---

[5] *er* occurs in a position for clitics in these examples, thus stranding the prepositions. Also note that when the preposition *met* is stranded, it takes on the allomorphic form *mee*.

[6] Observe that the partitive subject precedes the quantitative clitic *er* in this example.

The final examples of this section show that a single overt *er* can represent four different functions in a single sentence. The initial example contains expletive *er*, as the sentence is existential:

(8) a. dat er$_X$ [**twee studenten**] [**drie boeken**] [**uit de boekkast**] gehaald
         that there two students   three books   out.of the bookcase fetched
         hebben.
         have

   b. dat er$_{XQQ}$ [$_{NP}$ **twee** e] [$_{NP}$ **drie** e] **uit de boekkast** gehaald hebben.

   c. dat er$_{XQQP}$ [$_{NP}$ **twee** e] [$_{NP}$ **drie** e] **uit** gehaald hebben.

In (8b), the single *er* in addition represents the quantitative *er*s of the two partitive noun phrases *twee* and *drie*. Finally, in (8c), the object of the preposition *uit* is interpreted as pronominal *er*, leading in sum to the single overt *er* carrying four different functions within that sentence.

In sum, the examples in this section support the following two descriptive generalizations about sentences without existential *er* in the prefield:

1. Only one overt *er* can occur in the middle field.
2. When one overt *er* is present in the middle field, additional *er*s may be understood.

2.4   Clauses with existential *er* in the prefield

We now turn to sentences like (1a), repeated for convenience below, which contain an existential *er* in the prefield:

(9) *Er$_X$* loopt een man op straat.
     there walks a   man in the.street

These structures need to be discussed separately because unlike the clauses without *er$_X$* in the prefield, they permit more than a single overt *er* in a single clause under some circumstances. All of these clauses are verb-second main clauses and existential *er* is the only *er* permitted to fill the prefield. Moreover, *er$_X$* in the prefield can co-occur with all other *er*s in the middle field, however the latter differ in whether they are allowed to be overt or not.

The behavior of locational and pronominal *er* is simple: both have to remain unexpressed when expletive *er* fills the prefield, as the examples below demonstrate:

(10) a. *Er$_X$* wordt (*er_L) morgen   gedanst.
          there is   there   tomorrow danced

    b. *Er$_X$* wordt (*er_P) morgen   *over* gesproken.
         there is   there   tomorrow about spoken

105

The examples become grammatical if the *er* in the middle field does not appear.

Quantitative *er* shows the opposite behavior: it cannot remain implicit but must be spelled out separately from the initial existential *er* in the middle field:

(11) Er_X zijn er_Q gisteren $[_{NP}$ **twee** $[e]]$ gestolen.
         there have.been there yesterday   two   stolen

In sentences with two quantitative NPs in the middle field, only one quantitative *er* can be spelled out, however (Hans Broekhuis, p.c.):

(12)   [Er   hebben veel   studenten een onvoldoende   gekregen] en
          there(E) have   many students   an   unsatisfactory_mark gotten   and
          er   hebben **er**$_{QQ}$ [**een paar** e] zelfs [**twee** e] gekregen.
          there(E) have   there(QQ)   a   couple   even two   gotten

We sum up the generalizations for sentences with existential *er* in the prefield:

1. Only existential *er* can occur in the prefield, the other ones cannot.
2. When existential *er* occupies the prefield, then
     • an additional single overt quantitative *er* can appear in the middle field
     • implicit locational and pronominal *er*s are possible.

3   The Analysis

As we saw above, both main and subordinate clauses in Dutch show the phenomenon that one or more *er*s can remain implicit when at least one *er* is expressed overtly. In order to capture this in a grammatical theory, a mechanism is needed that makes it possible for an overt *er* to influence whether additional *er*s can or must be expressed. Moreover, this mechanism must be sensitive to the location of the overt *er* in phrase and/or linear structure.

The guiding ideas of our analysis are the following. We assume that existential and locational *er* are arguments of finite verbs, perhaps directly or through argument extension. Moreover, finite verbs attract to their ARG-ST the quantitative and pronominal *er*-complements of their NP and PP arguments that have not been realized within these phrases. Thus, all the *er*s that in principle can be realized at the sentence level appear in one place, namely the ARG-ST of the finite verb heading the sentence. The haplological effect then arises through the interaction of a number of constraints on the ARG-ST and COMPS lists of finite verbs.

In the remainder of the article, we will make these guiding ideas more precise and apply the resulting theory to representative examples.

3.1   Assumptions about Dutch phrase structure

We assume that the Dutch phrase structure system creates a number of linear fields and that in every sentence where they are realized, the fields occur in the left-to-right order that corresponds to their top-to-bottom ordering in Table 1:

Field	Description
pre-fld	the initial position in main clauses
lb	the left sentence bracket, filled by either a finite verb or a complementizer
mid-fld	the middle field contains the elements inbetween the two sentence brackets
rb	the right sentence bracket is made up of one or more verbs
fin-fld	the final field follows the right sentence bracket

Table 1: Description of Dutch topological fields

Except for the pre-field, which is restricted to main clauses, every field can occur in both main and subordinate clauses. We postulate an attribute FLD appropriate for objects of type *synsem*. It is crucial to our account that phrase structure configurations as well as the lexicon may constrain the FLD value of signs.

The trees in Figure 1 sketch the phrase and linear structure of verb-second and subordinate clauses we assume. The units that are connected to their mothers by
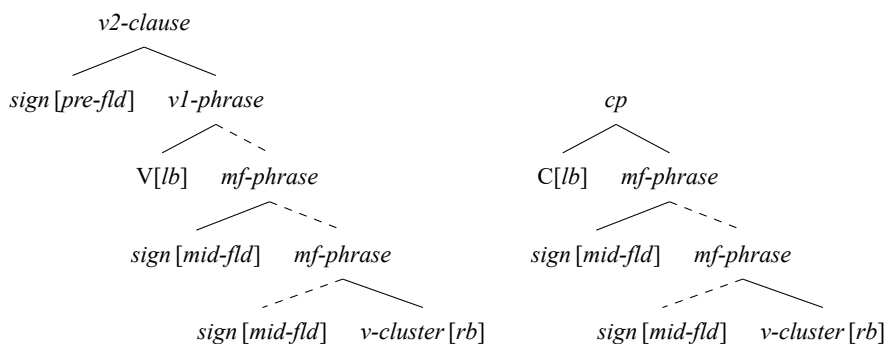

Figure 1: Basic phrase structure of verb-second and subordinate clauses

dashed lines are optional. The phenomenon we are dealing with in this paper reflects the appearance of expletive *er* in the pre-field of main clauses and of one or more *er*s in the middle field of both main and subordinate clauses.

3.2 Assumptions about *er*

In order to capture that the four *er*s on the one hand share properties and yet have different meanings and distributions, we postulate a general lexical identifier *er-lid* that all four *er*s share and a specific subtype for each different *er*: *er-X*, *er-Q*, *er-P*, and *er-L*, as shown in Figure 2.

Using these LID values, we can impose field constraints on the four *er*s lexically. The partial lexical entries in figure 3 permit existential *er* to occur in the pre-field and the middle field (see Broekhuis (2013, p. 337, 338) for this constraint) whereas the three remaining *er*s are restricted to the value *mid-fld* for the FLD attribute.

Second, quantitative *er* must be prevented from being realized within its partitive NP, as it always occurs outside of that NP when it is realized overtly. The
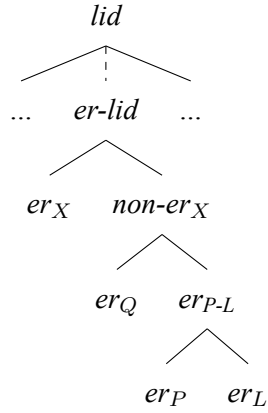
$$lid$$

$$\ldots \quad er\text{-}lid \quad \ldots$$

$$er_X \quad non\text{-}er_X$$

$$er_Q \quad er_{P\text{-}L}$$

$$er_P \quad er_L$$

Figure 2: Partial hierarchy of lexeme identifier types

(13)

$$
\begin{bmatrix} word \\ \text{SYNSEM} \begin{bmatrix} \text{HEAD} & er_X \\ \text{FLD} & \textit{pre-fld} \lor \textit{mid-fld} \end{bmatrix} \end{bmatrix}
\qquad
\begin{bmatrix} word \\ \text{SYNSEM} \begin{bmatrix} \text{HEAD} & non\text{-}er_X \\ \text{FLD} & \textit{mid-fld} \end{bmatrix} \end{bmatrix}
$$

Existential *er* **Non-existential *er*s**

Figure 3: Lexical entries for *er*

following constraint has the desired consequence by ruling out noun phrases with er_Q as a non-head daughter:

(14) $\begin{bmatrix} hd\text{-}comp\text{-}ph \\ \text{HEAD} \ \text{noun} \end{bmatrix} \longrightarrow \begin{bmatrix} \text{NON-HD-DTR} & \begin{bmatrix} \text{LID} \ \neg ER_Q \end{bmatrix} \end{bmatrix}$

Like other units, *er*s can be canonical and noncanonical synsems. It will be important for our analysis that like clitics in languages such as French (Miller & Sag (1997)), *er*s in Dutch have the option of the *synsem* value *pro-synsem*, as shown in Figure 4. This causes them to remain unrealized in phrase structure.

Next, we state argument realization constraints for finite verbs and nouns. Finite verbs map all and only their canonical arguments to their COMPS list. Note that it follows from this constraint that *pro-synsem er* arguments of finite verbs cannot appear on the verbs' COMPS list:

(15) $\begin{bmatrix} \text{HEAD} \ \text{V}[\textit{fin}] \end{bmatrix} \longrightarrow \begin{bmatrix} \text{SUBJ} & \langle \rangle \\ \text{COMPS} & \boxed{1} \ \textit{list(canon-ss)} \\ \text{ARG-ST} & \boxed{1} \bigcirc \textit{list(noncanon-ss)} \end{bmatrix}$
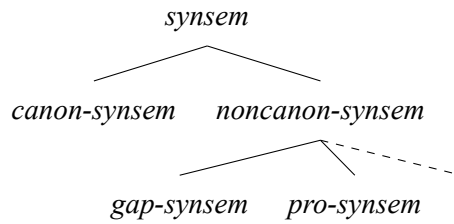
$$synsem$$

Figure 4: Partial hierarchy of *synsem* types

Nouns and prepositions differ from (finite) verbs in one crucial respect. Recall from the introduction to this section that we are assuming that finite verbs attract to their ARG-ST the quantitative and pronominal *er*-complements of their NP and PP arguments that have not been realized within these phrases. Whether or not such a raised *er* is expressed at the sentence level is a function of constraints on the ARG-S and the COMPS lists of the finite verb heading the sentence. As that determination is made only after the argument raising of quantitative and pronominal *er*-complements, such raising must be possible, no matter whether the *er*s' *synsem* type is canonical or non-canonical. Therefore, nouns and prepositions map all their *er*-arguments to their COMPS list, independent of the *er*'s canonicality. Below we present the constraint on nouns. The *er* (which is optional, since not every use of a noun is partitive) carries the tag $\boxed{2}$:

(16) $\begin{bmatrix} \text{HEAD} & \text{N} \end{bmatrix} \longrightarrow$

$$\begin{bmatrix} \text{SUBJ} & \langle\rangle \\ \text{COMPS} & \boxed{1}\ list(canon\text{-}ss)\ \bigcirc\ \boxed{2} \\ \text{ARG-ST} & \boxed{1}list\left(\neg\begin{bmatrix}\text{LID} & er\end{bmatrix}\right)\ \bigcirc\ \boxed{2}\big\langle(er_Q)\big\rangle\ \bigcirc\ list(noncanon\text{-}ss) \end{bmatrix}$$

The constraint on prepositions is analogous.

3.3   Constraints on the Argument Structures and COMPS Lists of Finite Verbs

With these preliminaries out of the way, we are now in a position to state the constraints that will interact to create the haplology effects illustrated in section 2. All constraints regulate the occurrence or co-occurrence of *er*s on the argument structure or COMPS lists of finite verbs.

3.3.1   *er*-Expression Constraints

The first constraint simply states that at least one *er*-argument of a finite verb must appear on the verb's COMPS list and be overtly expressed:

(17) *er*-EXPRESSION CONSTRAINT

$$\begin{bmatrix} \text{HEAD} & \text{V}[\textit{fin}] \\ \text{ARG-ST} & \left\langle \begin{bmatrix} \text{LID} & \textit{er} \end{bmatrix} \right\rangle \bigcirc \textit{list} \end{bmatrix} \longrightarrow \begin{bmatrix} \text{COMPS} & \left\langle \begin{bmatrix} \text{LID} & \textit{er} \end{bmatrix} \right\rangle \bigcirc \textit{list} \end{bmatrix}$$

Given that the system will permit *er* arguments to remain implicit, the constraint above is epistemologically plausible, as it requires at least one of the *er* arguments of a verb to be expressed. The expression of this *er* can thus serve as a signal to the possibility of implict *er*s.

The next two constraints contribute to the opposing behavior of quantitative *er* on the one hand and pronominal and locational *er* on the other in main clauses like (10)-(12), whose prefield is filled by existential *er*. (18) requires that verbs with a quantitative *er* argument must realize an *er* complement in the middle field:

(18) MIDDLE FIELD *er*-EXPRESSION CONSTRAINT

$$\begin{bmatrix} \text{HEAD} & \text{V}[\textit{fin}] \\ \text{ARG-ST} & \left\langle \begin{bmatrix} \text{LID} & \textit{er}_Q \end{bmatrix} \right\rangle \bigcirc \textit{list} \end{bmatrix} \longrightarrow \begin{bmatrix} \text{COMPS} & \left\langle \begin{bmatrix} \text{LID} & \textit{er} \\ \text{FLD} & \textit{mid-fld} \end{bmatrix} \right\rangle \bigcirc \textit{list} \end{bmatrix}$$

The next constraint applies to verbs which have a canonical pronominal or locational *er* argument. The COMPS list of these verbs is well formed only if it does not contain an expletive *er* with field value *pre-fld*.

(19) P-L *er*-EXPRESSION CONSTRAINT

$$\begin{bmatrix} \text{HEAD} & \text{V}[\textit{fin}] \\ \text{ARG-ST} & \left\langle \begin{bmatrix} \textit{canon-synsem} \\ \text{LID} & \textit{er}_{P\text{-}L} \end{bmatrix} \right\rangle \bigcirc \textit{list} \end{bmatrix} \longrightarrow$$

$$\begin{bmatrix} \text{COMPS} & \textit{list}\left(\neg \begin{bmatrix} \text{LID} & \textit{er} \\ \text{FLD} & \textit{pre-fld} \end{bmatrix} \right) \end{bmatrix}$$

Finally, we state the constraint that creates the syntactic haplology effect of *er*. (20) says that a finite verb selects at most one *er*-complement in the middle field:

(20) MIDFIELD SINGLE-*er* CONSTRAINT

$$\begin{bmatrix} \text{HEAD} & \text{V}[\textit{fin}] \\ \text{COMPS} & \left\langle \begin{bmatrix} \text{LID} & \textit{er} \\ \text{FLD} & \textit{mid-fld} \end{bmatrix} \right\rangle \bigcirc \textit{list} \end{bmatrix} \longrightarrow$$

$$\begin{bmatrix} \text{COMPS} & \left\langle \begin{bmatrix} \text{LID} & \textit{er} \\ \text{FLD} & \textit{mid-fld} \end{bmatrix} \right\rangle \bigcirc \textit{list}\left(\neg \begin{bmatrix} \text{LID} & \textit{er} \\ \text{FLD} & \textit{mid-fld} \end{bmatrix} \right) \end{bmatrix}$$

4 Illustration of the major cases

We now illustrate the interplay of the lexical, phrasal, and linear constraints that we have formulated in the previous section. We discuss five cases in detail.

4.1 Case 1: two overt *er*s in the middle field are ruled out

(21)   * dat er_X er_L gedanst wordt.
  that there there danced is
  Intended reading: 'People are dancing there.'

According to our assumptions, the presence of two overt *er*s in the middle field of this sentence would require the verb *wordt* to have two *er* complements (in addition to its verbal complement), as shown in Figure 5. This structure is obviously not licensed by our approach, as *wordt* violates the MIDFIELD SINGLE-*er* CONSTRAINT (20), which permits verbs to have at most one *er* complement with field value *mid-fld*.
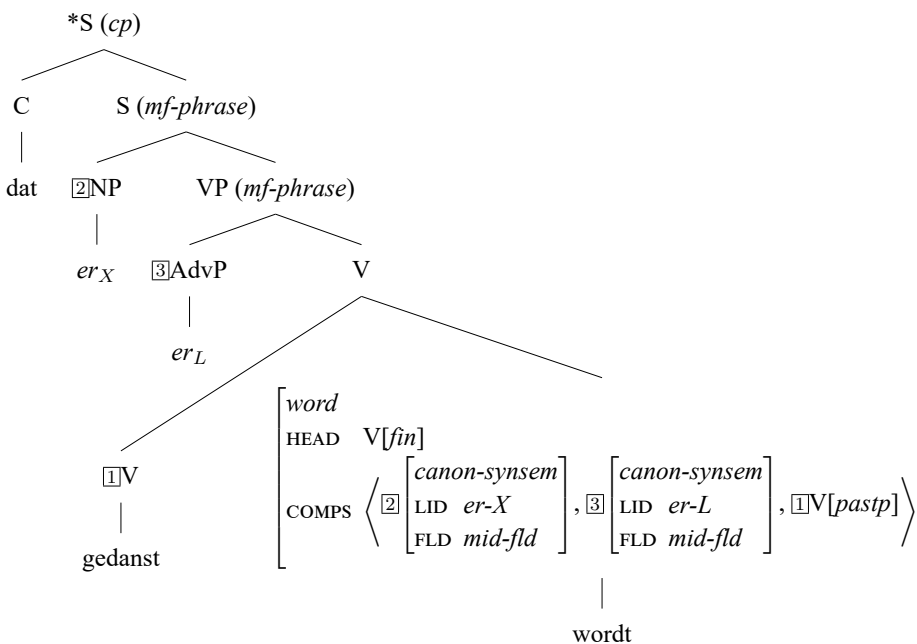

Figure 5: Analysis for example (21)

4.2 Case 2: one overt and one implicit *er* in the middle field

The next case differs from the previous one in that it contains a single overt *er* in the middle field and a second understood *er*, as the verb *wonen* selects a locational complement. The example is grammatical.

(22)   dat *er_{XL}* veel mensen **wonen**.
           that there many people live
           'that many people live there.'

    To license the structure above, in addition to its two overt complements *er* and *veel mensen*, the verb *wonen* must have an implicit *er_L* argument which is not mapped to the verb's COMPS list, as shown in Figure 6. Unlike the finite verb in Case
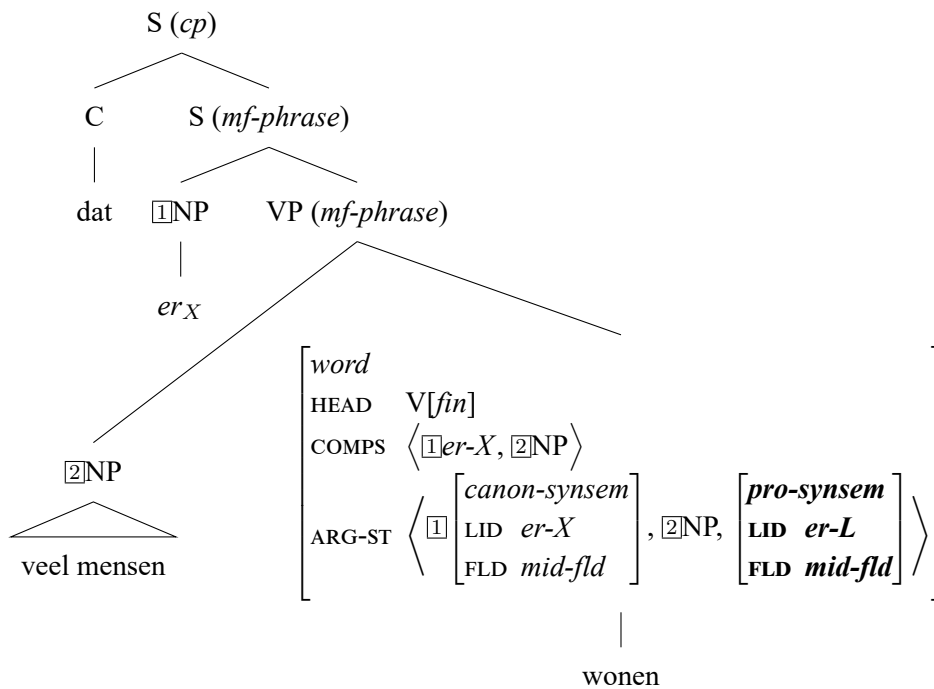
Figure 6: Partial analysis for example (22)

1, the word *wonen* in the tree immediately above satisfies all of our constraints:

1. *er*-EXPRESSION CONSTRAINT: *wonen* selects an *er* complement.
2. MIDDLE FIELD *er*-EXPRESSION CONSTRAINT: is vacuously satisfied, as *wonen* doesn't have an *er_Q* argument.
3. P-L *er*-EXPRESSION CONSTRAINT: is satisfied, as there is no *er* in the prefield.
4. MIDFIELD SINGLE-*er* CONSTRAINT: *wonen* has no more than a single mid-field *er* complement.

4.3   Case 3: an expletive *er* in the prefield and a pronominal *er* in the middle field

The next two cases deal with main clauses whose prefield is filled by expletive *er*. Recall that quantitative *er* parts ways with locational and pronominal *er* in this

sentence type. When present, the latter two have to remain implicit. This is why (23) with an overt pronominal *er* in the middle field is ungrammatical:

(23)　　* *Er*$_X$ wordt **er**$_P$　morgen　　*over*　gesproken.
　　　　　there is　　　there　tomorrow　about　spoken

For the string above to be licensed, it would need to have the structure shown in Figure 7.
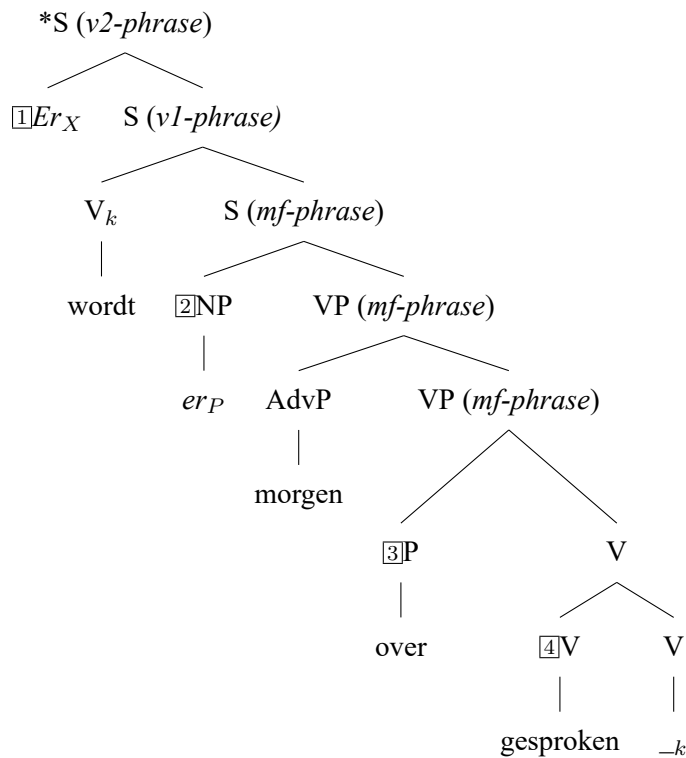

Figure 7: Partial analysis of example (23)

In this structure, the gap of the verb *wordt* would need to have four complements: (i) the expletive *er* occuring in the prefield, (ii) the overt pronominal *er* in the middle field which the verb has inherited from (iii) the preposition *over*, and (iv) the passive participle *gesproken*:

$$
(24) \quad
\begin{bmatrix}
\textit{word} \\
\text{HEAD} \quad \text{V}[\textit{fin}] \\
\text{COMPS} \quad \left\langle \boxed{1}\begin{bmatrix} \textit{canon-synsem} \\ \text{LID} \quad \textit{er-X} \\ \text{FLD} \quad \textit{pre-fld} \end{bmatrix}, \boxed{2}\begin{bmatrix} \textit{canon-synsem} \\ \text{LID} \quad \textit{er-P} \\ \text{FLD} \quad \textit{mid-fld} \end{bmatrix}, \boxed{3}\begin{bmatrix} \textit{canon-synsem} \\ \text{HEAD} \quad \text{P} \\ \text{COMPS} \quad \langle \boxed{2} \rangle \end{bmatrix}, \boxed{4}\text{V}[\textit{pass}] \right\rangle
\end{bmatrix}
$$

According to our approach, the COMPS list above is illicit:

The P-L *er*-EXPRESSION CONSTRAINT is violated, as a pronominal *er* is incompatible with an *er* in the prefield. (10a) with an overt locational *er* in predicted to be ungrammatical for the same reason.

Since (23) violates only one of our constraints, it is correctly predicted that it becomes grammatical when the offending pronominal *er* remains implicit:

(25)   *Er$_{XP}$* wordt morgen   *over* gesproken.
         there   is   tomorrow about spoken

The constraint profile of this structure is as follows:

1. *er*-EXPRESSION CONSTRAINT: the verb selects an *er* complement.
2. MIDDLE FIELD *er*-EXPRESSION CONSTRAINT: is vacuously satisfied, as the verb doesn't have an *er$_Q$* argument.
3. P-L *er*-EXPRESSION CONSTRAINT: is vacuously satisfied, as there is no canonical P-L *er* argument in the middle field.
4. MIDFIELD SINGLE-*er*: CONSTRAINT the verb has no more than a single midfield *er* complement.

4.4   Case 4: an expletive *er* in the prefield and a quantitative *er* in the middle field

Quantitative *er* differs from locational and pronominal *er* in that it must appear overtly in the middle field of sentences introduced by existential *er*. Without the second *er* in the middle field, (26) is ungrammatical.

(26)   *Er$_X$* zijn   *(*er$_Q$*) gisteren   [$_{NP}$ **twee** [*e*]] gestolen.
         there have.been there   yesterday   two   stolen

Under our assumptions, the structure of this sentence is as shown in Figure 8. Expletive *er* appears in the pre-field and quantitative *er* in the middle field. The main verb *gestolen* has inherited the quantitative *er* from the noun *twee* and the auxiliary *zijn* has interited all the arguments of *gestolen*. Altogether, this requires the auxiliary (and its gap) to have two *er* and one N-complement (= *twee*), plus the passive participle of the main verb, as indicated in example (27).

$$(27) \quad \begin{bmatrix} \textit{word} \\ \text{HEAD} \quad \text{V}[\textit{fin}] \\ \text{COMPS} \quad \left\langle \boxed{1}\begin{bmatrix} \textit{canon-synsem} \\ \text{LID} \quad \textit{er-X} \\ \text{FLD} \quad \textit{pre-fld} \end{bmatrix}, \boxed{2}\begin{bmatrix} \textit{canon-synsem} \\ \text{LID} \quad \textit{er-Q} \\ \text{FLD} \quad \textit{mid-fld} \end{bmatrix}, \boxed{3}\begin{bmatrix} \textit{canon-synsem} \\ \text{HEAD} \quad \text{N} \\ \text{COMPS} \quad \langle\boxed{2}\rangle \end{bmatrix}, \boxed{4}\text{V}[\textit{pass}] \right\rangle \end{bmatrix}$$

With quantitative *er* expressed in the middle field, the sentence obeys all constraints:

1. *er*-EXPRESSION CONSTRAINT: the verb selects an *er* complement.

S (*v2-phrase*)
├── [1]*Er*$_X$
└── S (*v1-phrase)*
    ├── V$_k$
    │   └── zijn
    └── S (*mf-phrase*)
        ├── [2]NP
        │   └── *er*$_Q$
        └── VP (*mf-phrase*)
            ├── AdvP
            │   └── gisteren
            └── VP (*mf-phrase*)
                ├── [3]N
                │   └── twee
                └── V
                    ├── [4]V
                    │   └── gestolen
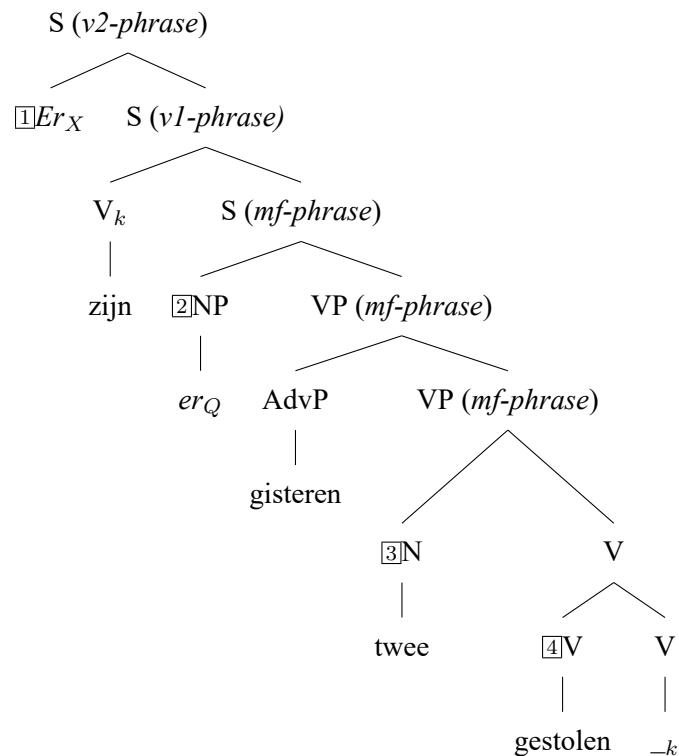   └── V
                        └── _$_k$

Figure 8: Partial analysis of example (26)

2. MIDDLE FIELD *er*-EXPRESSION CONSTRAINT: is satisfied, as there is an *er* complement in the middle field.

3. P-L *er*-EXPRESSION CONSTRAINT: is vacuously satisfied, as there is no P-L *er* argument in the middle field.

4. MIDFIELD SINGLE-*er*: CONSTRAINT the verb has no more than a single midfield *er* complement.

Without the *er* in the middle field, the sentence becomes ungrammatical, as the MIDDLE FIELD ER-EXPRESSION CONSTRAINT is now violated because the constraint requires a verb with an *er*$_Q$ argument to have an overt mid-field *er* complement.

4.5 Case 5: an *er* with four functions

This brings us to the final case. We will demonstrate that our constraints predict the following sentence to be grammatical, in which a single overt *er* expresses four functions at once. As the sentence is existential, the existential function must be present, the two partitive NPs *twee* and *drie* each require a quantitative function, and the stranded preposition *uit* requires the pronominal function.

(28) dat er$_{XQQP}$ [NP **twee** e]   [NP **drie** e]   **uit**   gehaald hebben.
      that there   two   students   three books out.of fetched have

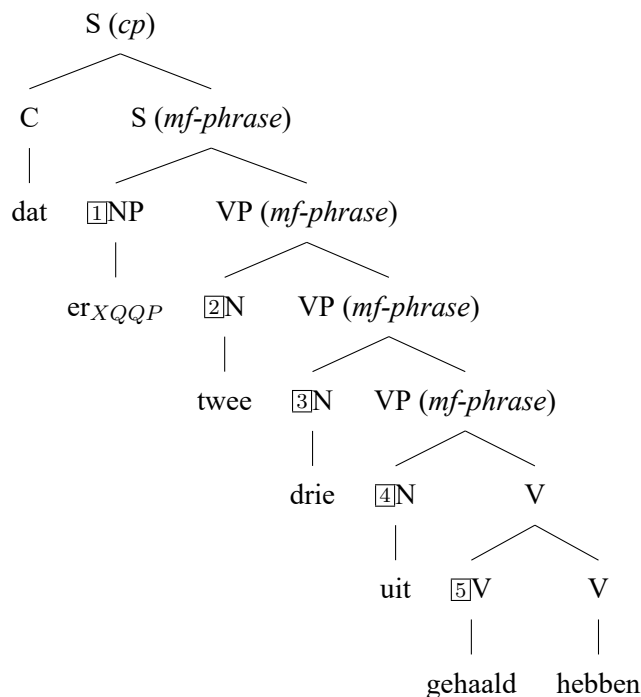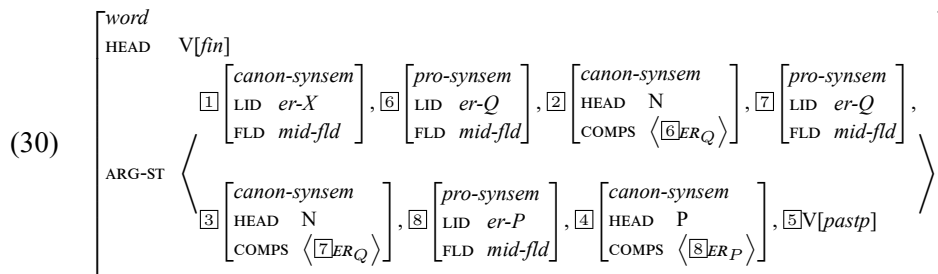(28) has the structure shown in Figure 9. The head of the finite sentence *hebben*

Figure 9: Partial analysis of example (28)

has the following COMPS list in the structure above. From left to right, the verb's complements are as follows: (i) the existential *er*, (ii)-(iii) the partitive nouns *twee* and *drie*, (iv) the preposition *uit*, and (v) the main verb *gehaald*.

$$
(29) \quad
\begin{bmatrix}
\textit{word} \\
\text{HEAD} \quad \text{V[\textit{fin}]} \\
\text{COMPS} \left\langle
\boxed{1}\begin{bmatrix}\textit{canon-synsem}\\ \text{LID}\ \textit{er-X}\\ \text{FLD}\ \textit{mid-fld}\end{bmatrix},\
\boxed{2}\begin{bmatrix}\textit{canon-synsem}\\ \text{HEAD}\ \text{N}\\ \text{COMPS}\ \langle ER_Q\rangle\end{bmatrix},\
\boxed{3}\begin{bmatrix}\textit{canon-synsem}\\ \text{HEAD}\ \text{N}\\ \text{COMPS}\ \langle ER_Q\rangle\end{bmatrix}, \\
\boxed{4}\begin{bmatrix}\textit{canon-synsem}\\ \text{HEAD}\ \text{P}\\ \text{COMPS}\ \langle ER_P\rangle\end{bmatrix},\ \boxed{5}\text{V[\textit{pastp}]}
\right\rangle
\end{bmatrix}
$$

Note that the two stranded partitive nouns and the preposition each have an *er*-complement on their COMPS lists which is inherited by *gehaald* and ultimately by the head *hebben* of the whole structure. These *er*s are not visible in the COMPS list of *hebben*, since then the verb would have more than a single *er* on its COMPS list in

116

violation of the MIDFIELD SINGLE-*er* CONSTRAINT. But they are present on the verb's ARG-ST, where they immediately precede their source.

$$
(30) \quad
\begin{bmatrix}
\textit{word} \\
\text{HEAD} \quad \text{V}[\textit{fin}] \\
\text{ARG-ST} \quad
\left\langle
\boxed{1}
\begin{bmatrix}
\textit{canon-synsem} \\
\text{LID} \quad \textit{er-X} \\
\text{FLD} \quad \textit{mid-fld}
\end{bmatrix},
\boxed{6}
\begin{bmatrix}
\textit{pro-synsem} \\
\text{LID} \quad \textit{er-Q} \\
\text{FLD} \quad \textit{mid-fld}
\end{bmatrix},
\boxed{2}
\begin{bmatrix}
\textit{canon-synsem} \\
\text{HEAD} \quad \text{N} \\
\text{COMPS} \quad \langle \boxed{6}_{ER_Q} \rangle
\end{bmatrix},
\boxed{7}
\begin{bmatrix}
\textit{pro-synsem} \\
\text{LID} \quad \textit{er-Q} \\
\text{FLD} \quad \textit{mid-fld}
\end{bmatrix}, \\
\boxed{3}
\begin{bmatrix}
\textit{canon-synsem} \\
\text{HEAD} \quad \text{N} \\
\text{COMPS} \quad \langle \boxed{7}_{ER_Q} \rangle
\end{bmatrix},
\boxed{8}
\begin{bmatrix}
\textit{pro-synsem} \\
\text{LID} \quad \textit{er-P} \\
\text{FLD} \quad \textit{mid-fld}
\end{bmatrix},
\boxed{4}
\begin{bmatrix}
\textit{canon-synsem} \\
\text{HEAD} \quad \text{P} \\
\text{COMPS} \quad \langle \boxed{8}_{ER_P} \rangle
\end{bmatrix},
\boxed{5}\text{V}[\textit{pastp}]
\right\rangle
\end{bmatrix}
$$

Despite the relative complexity of this argument structure and its relation to the verb's COMPS list, *hebben* satisfies all of the constraints we formulated in section 3.

1. *er*-EXPRESSION CONSTRAINT: the verb selects an *er* complement.
2. MIDDLE FIELD *er*-EXPRESSION CONSTRAINT: is satisfied, as there is an *er* complement in the middle field.
3. P-L *er*-EXPRESSION CONSTRAINT: is satisfied, as there is no *er* in the pre-field.
4. MIDFIELD SINGLE-*er* CONSTRAINT: the verb has no more than a single mid-field *er* complement.

(28) is thus correctly predicted to be grammatical with one overt *er* that carries four different functions.

5   Conclusion

Dutch has four pronouns *er* which show an intriguing pattern of syntactic haplology when a finite verb has more than one *er* argument. We presented a theory that captures this pattern by relying on two central aspects of HPSG:

1. the distinction between ARG-ST and COMPS
2. the distinction between canonical and non-canonical *synsem*.

No deletion rules of the kind used in transformational analyses of *er* are necessary. We are not aware of any other formal theory that captures all the data presented in this paper. It remains to be seen whether other cases of syntactic haplology are susceptable to the kind of analysis used here.

References

Bech, Gunnar. 1952. Über das niederländische Adverbialpronomen *er*. *Traveaux du Cercle Linguistique de Copenhague* 8. 5–32.

Bennis, Hans. 1986. *Gaps and Dummies*. Dordrecht: Foris.

Bouma, Gosse. 2001. Argument Realization and Dutch R-Pronouns: Solving Bech's Problem without Movement or Deletion. In Ronnie Cann, Claire Grover & Philip Miller (eds.), *Grammatical Interfaces in HPSG*, Stanford: CSLI Publications.

Broekhuis, Hans. 2013. *Syntax of Dutch: Adpositions and Adpositional Phrases*. Amsterdam: Amsterdam University Press.

Campbell-Kibler, Kathryn. 2001. Bech's Problem, Again: Using Linearization on Dutch R-Pronouns. In Frank Van Eynde, Lars Hellan & Dorothee Beermann (eds.), *The Proceedings of the 8th International Conference on Head-Driven Phrase Structure Grammar*, 87—102. Stanford: CSLI Publications.

Drach, Erich. 1937. *Grundgedanken der deutschen Satzlehre*. Frankfurt: Diesterweg.

Höhle, Tilman N. 1983. Topologische Felder. Ms.

Höhle, Tilman N. 1986. Der Begriff '''Mittelfeld''', Anmerkungen über die Theorie der topologischen Felder. In Walter Weiss, Herbert Ernst Wiegand & Marga Reis (eds.), *Akten des VII. Kongresses der Internationalen Vereinigung für germanische Sprach- und Literaturwissenschaft. Göttingen 1985. Band 3. Textlinguistik contra Stilistik? – Wortschatz und Wörterbuch – Grammatische oder pragmatische Organisation von Rede?* (Kontroversen, alte und neue 4), 329–340. Tübingen: Max Niemeyer Verlag.

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

Neeleman, A. & H. van de Koot. 2006. Syntactic Haplology. In Martin Everaert & Henk van Riemsdijk (eds.), *The Blackwell Companion to Syntax*, vol. 4, 685—710. London: Blackwell.

Odijk, Jan. 1993. *Compositionality and Syntactic Generalizations*: Tilburg University dissertation.

Reis, Marga. 1980. On Justifying Topological Frames: 'Positional Field' and the Order of Nonverbal Constituents in German. *Documentation et Recherche en Linguistique Allemande Contemporaine* 22/23. 59–85.

van Riemsdijk, Henk C. 1978. *A Case Study in Syntactic Markedness: The Binding Nature of Prepositional Phrases*. Dordrecht: Foris.

Van Eynde, Frank. 2019. Clustering and Stranding in Dutch. *Linguistics* 57(5). 1025–1071.

Van Eynde, Frank & Liesbeth Augustinus. 2014. Complement Raising, Extraction and Adposition Stranding in Dutch. In Stefan Müller (ed.), *Proceedings of the 21st International Conference on Head-Driven Phrase Structure Grammar, University at Buffalo*, 156–175. Stanford, CA: CSLI Publications. http://cslipublications.stanford.edu/HPSG/2014/vaneynde-augustinus.pdf.