A lexicalist approach to Dutch cross serial dependencies

Rentier, G.M.

Publication date:
1994

Document Version
Publisher's PDF, also known as Version of record

[redacted]

RESEARCH REPORT

INSTITUTE FOR LANGUAGE TECHNOLOGY AND ARTIFICIAL INTELLIGENCE INSTITUUT VOOR TAAL- EN KENNISTECHNOLOGIE

ITK Research Report No. 50

A Lexicalist Approach to Dutch Cross Serial Dependencies

Gerrit Rentier

ITK
Warandelaan 2
P.O. Box 90152
5000 LE TILBURG
rentier@kub.nl

May 1994

ISSN 0924-7807

A Lexicalist Approach to Dutch Cross Serial Dependencies

Gerrit Rentier*
Institute for Language Technology and Artificial Intelligence
Tilburg University, PO Box 90153, 5000 LE Tilburg, The Netherlands
rentier@kub.nl

1 Approaches to Dutch Cross Serial Dependencies

Cross serial dependencies in Dutch (DCSDs) confront linguistic theory with a recursive and intriguingly systematic bounded discontinuity, indicated here by subscripts;

(1) a. omdat ik_1 $haar_2$ de nijlpaarden$_2$ zag_1 voeren$_2$
because I[NOM] her[ACC] the hippos saw[FIN] feed[BSE]
"because I saw her feed the hippos"

b. dat ik_1 $haar_2$ hem_3 de nijlpaarden$_3$ zag_1 helpen$_2$ voeren$_3$
that I[NOM] her[ACC] him[ACC] the hippos saw[FIN] help[BSE] feed[BSE]
"that I saw her help him feed the hippos"

Transformational analyses, starting with (Evers 1975), have generally involved repeated rightward *head movement* to get the shuffled word order effect, where this structural operation is triggered by appropiate configurations. In categorial grammar DCSDs have been analyzed through a derivational step which is known as *function composition* (Steedman 1985). In CG, this combinatory rule of function composition is allowed to apply when the proper categories are adjacent during a derivation.

We present what we would call an in comparison more strictly lexicalist approach which builds on a mechanism typically available within Head-driven Phrase Structure Grammar (HPSG, cf. (Pollard & Sag 1994)), namely the mechanism of *structure-sharing* (token-identity of information; formally, re-entrancy in graphs). In the abovementioned approaches DCSDs are analyzed through a 'generally' available structural or derivational operation; an appealing aspect of our approach appears to be

*This report is a preprint of a paper which will appear in the *Proceedings of the 30th Regional Meeting of the Chicago Linguistic Society*, April 14-16 1994, Chicago, USA. Author was partly sponsored by EC projects ESPRIT P5254 (PLUS), ESPRIT P6665 (DANDELION) and by two travel grants from the Netherlands Organization for Scientific Research (NWO). Many thanks to Bob Borsley, Jo Calder, Bart Geurts, Josée Heemskerk, John Nerbonne, Paola Monachesi, Ivan Sag, Wietske Sijtsma and Craig Thiersch for comments and discussion; also to several attendants of the *First International Workshop on HPSG* in Columbus, Ohio, August '93 as well as attendants of the *30th Regional Meeting of the CLS*.

that, as suggested by the data, the explanatory mechanism is triggered firstly obligatorily by lexical stipulation, and thus, secondly, applies only 'locally' when it is triggered by the presence of a member of a lexically determined, relatively small closed class subset of verbs in Dutch.

We start out then from the straightforward assumption that Dutch causative and perceptual verbs can lexically force 'inheritance' of arguments from the nonfinite verb forms which they govern. This lexically enforced effect, *"argument composition"*, has been introduced in HPSG for German auxiliary and modal verbs in (Hinrichs & Nakazawa 1989)(henceforth H&N). Since then it has been furthermore independently motivated to account for restructuring verbs and clitics in Italian by (Monachesi 1993) and for auxiliaries and causatives in French, cf. (Abeillé & Godard 1994).

German shows nested instead of cross serial dependencies with causative and perceptual verbs, since German word order in verbal clusters is in general (modulo so-called "auxiliary flip", cf. (H&N 1989) and (Baker 1994)) the mirror image of that in Dutch, cf. (2);

(2) weil ich$_1$ Cecilia$_2$ Hans$_3$ die Nilpferde$_3$ füttern$_3$ helfen$_2$ sah$_1$
because I Cecilia Hans the hippos feed help saw
"because I saw Cecilia help Hans feed the hippos"

We will extend the H&N analysis of German towards an account of Dutch CSDs, and account for the different word orders through one single linear precedence parameter.

However, our analysis of the verbal cluster is different from H&N's in important technical details. Furthermore we go into details of case assignment to subjects of bare infinitives (section 3) and work out the general idea that in such clusters it is not the governed but the governing verb which assigns case to the subject of the governed verb. The governed verb does, however, assign it's subject it's semantic role. Section 4 presents the analysis of (1a) and (1b), whereas in section 5 we show how our account of the local selection properties of base-form verbs allows for an interesting analysis of so-called 'irregular' case assignment to subjects of nonfinite verbs in German. Also we show how we can define infinitive markers in such a manner that we can account for 'switching' case with subjects of Dutch infinitival constructions as well. Finally, in section 6 we generalize the analysis such that we arrive at an integrated account of Dutch clausal word order and an important constraint on double infinitive constructions.

2 Assumptions on Valency and Dutch Clause Structure

Instead of the feature SUBCAT, which lists all the arguments locally selected by a head in (Pollard & Sag 1987), we will adopt the division between subject and non-subject arguments which is motivated and put forward for English in (Borsley 1987). This means we assume that subjects are locally selected through a feature SUBJ, and that non-subject arguments are locally selected through another valency feature, COMPS. In chapter 9 of (Pollard & Sag 1994) this approach is developed further and leads to the Valence Principle, which refers to the valence features SUBJ and COMPS through 'F':

(3) **Valence Principle** Chapter 9, (Pollard & Sag 1994)
In a headed phrase, for each valence feature F, the F value of the head-daughter is the concatenation of the phrase's F value with the list of SYNSEM values of the F-daughters' value.

The effect of this combinatory principle on any headed sign is that such a sign can only become "complete" or *saturated* if the sign is combined with the appropriate arguments, where appropriate arguments for each selecting sign are defined strictly lexically. As for clause structure, we agree with considerations mentioned in (Nerbonne forthc.) for the German Mittelfeld, and assume that they apply for Dutch as well. Therefore, we aim at a flat Mittelfeld analysis of Dutch instead of H&N's more contoured analysis of the Mittelfeld. Furthermore, we assume the presence of one additional valence-feature, GOV, ranging over a verbal argument, following (Chung forthc.) and discussions of Webelhuth, Ackerman, Sag and Pollard at WCCFL XIII suggesting this for German. The assumption of a flat Mittelfeld and the GOV-feature together suggest the following immediate dominance schemata (which will be illustrated in Figures 1 to 5):

(4) a. $\mathrm{XP}_{[\mathrm{LEX}-]} \rightarrow \mathrm{S},\ \mathrm{C}_1, ..., \mathrm{C}_n,\ \mathrm{H}_{[\mathrm{GOV}\langle\,\rangle,\mathrm{LEX}+]}$

b. $\mathrm{X}_{[\mathrm{LEX}+]} \rightarrow \mathrm{H}_{[\mathrm{GOV}\,\langle\,[\,\ldots\,]\rangle,\,\mathrm{LEX}+]},\quad \mathrm{C}_1$

Here H, S and C indicate that the daughters of the phrase include a head, a subject and complements, not necessarily in that order. Schema (4a) typically allows for flat root and complement clauses with a finite verbal head H which is LEX+. The Head Feature Principle will make sure that the HEAD-info on any head H will become instantiated on the mother, XP or X.

(5) **Head Feature Principle** Chapter 1, (Pollard & Sag 1994)
The HEAD value of any headed phrase is structure-shared with the HEAD value of the head-daughter.

That a verbal cluster too can be treated as 'word-like', that is, can be LEX+, we guarantee by our definition of schema (4b). Schema (4b) is in a sense not a phrase structure schema but is instead a "cluster-formation"-schema. Normally the combination of two or more words leads to a sign which is LEX−, a phrasal sign, as is the case with structures licensed by the ID-schema in (4a). However, in the case of (4b) the combination of signs which are LEX+ leads to a complex predicate which is LEX+ as well. Note also that (4b) is strictly binary: it takes one argument, namely the argument which is the value of GOV. We arrange the lexicon so that any value of GOV (the "[. . .]" in (4b)) will always be an unsaturated base form verb(-cluster) which is defined as LEX+ as well. This approach eliminates the possibility of *spurious ambiguities* simply because the verb-cluster rule is binary and arguments of the complex predicate can only be 'checked off' in a structure licensed by (4a) where the head is GOV⟨⟩. Also, by extension, this approach suggests an account for the fact that no modifiers can appear in the verb cluster (whereas modifiers can appear almost anywhere in the Mittelfeld), provided we treat adverbial modifiers as arguments of verbs as well (as proposed in (Miller 1992)).

We assume that the clause structure of Dutch CSDs is one where we generally have a binary right-branching verbal complex. This verbal complex then locally selects the sum total of the arguments which are 'originally' selected by the individual verbs which constitute the verbal cluster. We feel that such a structure is motivated by *auxiliary flip* in the same way as auxiliary flip motivates a binary left-branching structure for the German verbal complex, cf. (H&N 1989). In (6b) we give an auxiliary-flipped version of (6a):

(6) a. omdat Frits het nijlpaard niet [heeft [kunnen voeren]]
because Frits the hippo not has[FIN] can[BSE,AUX+] feed[BSE]
"Because Frits hasn't been able to feed the hippo"

b. omdat Frits het nijlpaard niet [[kunnen voeren] heeft]
because Frits the hippo not can[BSE] feed[BSE] has[FIN]

The reordering possibility in (6b) is most easily (without recourse to crossing branches or discontinuities) explained if we say that "*heeft*" has as a complement the verbal cluster "*kunnen voeren*". We assume that causative and perceptual verbs syntactically behave just like auxiliaries and modals. So they too will apply argument composition and raise all the complements of the governed verb(s) to become arguments of the complex predicate. The verbal cluster will be licensed by ID-schema (4b), and the result will give rise to a complex predicate which will be marked LEX+. This word-like complex can act as the head of any flat clause that is licensed by ID-schema (4a).

3 Case Assignment to Subjects of Nonfinite Verbs

In our approach to valency, the local syntactic properties of the base form of the transitive verb *voeren* ("feed") will look as in (7), displaying some technical detail:

(7)
$$\left[\begin{array}{ll} \text{PHON} & \langle\, \textit{voeren} \,\rangle \\ \text{SYNSEM} & \left[\begin{array}{ll} \text{LOC} & \left[\begin{array}{ll} \text{CAT} & \left[\begin{array}{ll} \text{HEAD} & \left[\begin{array}{l} \text{MAJOR V} \\ \text{VFORM BSE} \end{array}\right] \\ \text{SUBJ} & \langle\, \left[\text{LOC}\left[\text{CAT}\left[\text{HEAD}\left[\begin{array}{l}\text{MAJOR N}\\ \text{CASE CASE}\end{array}\right]\right]\right]\right]_{\textit{synsem}} \,\rangle \\ \text{COMPS} & \langle\, \left[\text{LOC}\left[\text{CAT}\left[\text{HEAD}\left[\begin{array}{l}\text{MAJOR N}\\ \text{CASE ACC}\end{array}\right]\right]\right]\right]_{\textit{synsem}} \,\rangle \\ \text{GOV} & \langle\,\rangle \end{array}\right] \end{array}\right]_{\textit{synsem}} \\ \text{LEX} & + \end{array}\right]_{\textit{word}}$$

We work out a proposal made for German in(Pollard forthc.) which holds that nonfinite verbs do not assign any case to their subject. In (7), the value for the subject-NP's

CASE-feature, "CASE", is the supertype in the type hierarchy for those atomic types that are appropriate values of the feature CASE. So, the value CASE is the supertype of NOM and ACC in Dutch and English, and in German also of DAT and GEN. The result of assigning the NP this supertype for case in practice boils down to giving this NP some kind of "any"-value for case that will unify with any other case value.

This assumption allows us to account for the fact that whereas subjects of governed nonfinite verb forms like "*voeren*" usually get nominative case, they may in certain constructions get accusative (or even dative) case. That is, whereas modals (and auxiliaries) seem to assign nominative case to subjects of nonfinite governed verbs (cf. (8a)), Dutch causative and perceptual verbs seem to assign it accusative case, cf. (8b):

(8) a. dat hij het konijn kan voeren.
that he[NOM] the rabbit can[FIN] feed[BSE]
"that he can feed the rabbit"

b. dat Karina hem het konijn zag voeren
that Karina him[ACC] the rabbit saw[FIN] feed[BSE]
"that Karina saw him feed the rabbit"

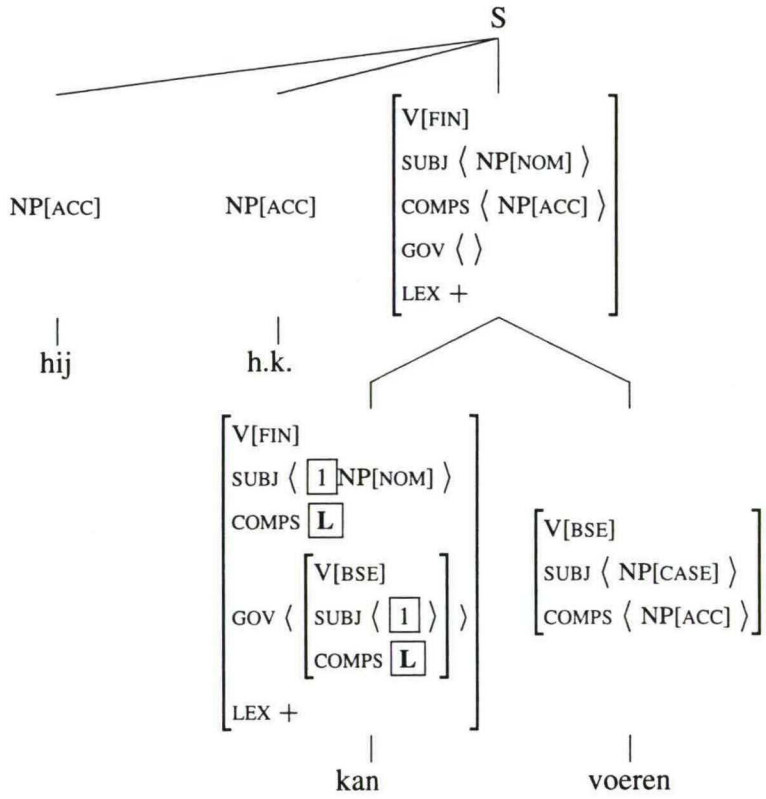
It is straightforward to assume that the case assignment to the subject argument of "*voeren*", more in general any NP which is the logical subject of a nonfinite verb, does not arise from the nonfinite verb but originates with the immediately governing verbs. We define modal verbs like *kan*("can") as argument composition verbs, cf. (9):[1]

(9)
$$\left[\begin{array}{ll}\text{PHON} & \langle\, \textit{kan} \,\rangle \\ \text{SYNSEM} & \left[\text{LOC}\left[\text{CAT}\left[\begin{array}{ll}\text{HEAD} & \left[\begin{array}{l}\text{MAJOR V} \\ \text{VFORM FIN}\end{array}\right] \\ \text{SUBJ} & \langle\, \boxed{1}\,\text{NP[NOM]} \,\rangle \\ \text{COMPS} & \boxed{\text{L}} \\ \text{GOV} & \left\langle \left[\begin{array}{l}\text{V[BSE]} \\ \text{SUBJ} \langle\, \boxed{1} \,\rangle \\ \text{COMPS} \boxed{\text{L}} \\ \text{GOV} \langle\,\rangle \\ \text{LEX} + \end{array}\right] \right\rangle\end{array}\right]\right] \\ \text{LEX} + & \end{array}\right]$$

In Figure 1, due to the Valence Principle, the sign which dominates the headed phrase "*kan voeren*" no longer selects anything through GOV. This local selection requirement of "*kan*" has been 'fulfilled' by "*voeren*". Note also that by ID-schema (4b), "*kan*

[1] In the figures and examples throughout this paper, recurring $\boxed{\text{i}}$'s will indicate structure-sharing, that is token-identity of information, as is common usage in HPSG.

Figure 1: Argument Composed Cluster with a Modal Verb



voeren" will be a sign which is marked as LEX+, and therefore can in turn be the head of any flat clause which is licensed through ID-Schema (4a). Since the finite "*kan*" (cf. (9)) is the head of "*kan voeren*", by the Head Feature Principle the resulting sign will be finite, [VFORM FIN], as well.

The unification of [CASE CASE] and [CASE NOM] will be forced through the structure-sharing indicated in (9) and Figure 1 as "$\boxed{1}$", and will result in the more specific restriction [CASE NOM]. In this way, the governing verb determines the case-marking of the subject of the governed verb. Similarly, the list $\boxed{\mathbf{L}}$ of non-subject arguments of the governed nonfinite verb is 'inherited' by the governing verb. It is this inheritance of local selection properties through structure-sharing which is called "argument composition". Since the governed V[BSE] is selected by "*kan*" through GOV *as* missing a subject and some list **L** of non-subject arguments, "*voeren*" must specifically *not* 'find' a subject or any arguments. And indeed it doesn't, cf. "*voeren*" as it appears in the tree in Figure 1.

The way in which we assigned nominative case-marking naturally suggests a lexicalist account of the accusative marking on the subject of *voeren* in (8b). We define finite perceptuals like *zag* and finite causatives as *argument composition* verbs too, but along the following lines:

(10)
$$\begin{bmatrix} \text{PHON} \langle \textit{zag} \rangle \\ \text{SYNSEM} \begin{bmatrix} \text{LOC} \begin{bmatrix} \text{CAT} \begin{bmatrix} \text{HEAD} \begin{bmatrix} \text{MAJOR} & \text{V} \\ \text{VFORM} & \text{FIN} \end{bmatrix} \\ \text{SUBJ} \langle \text{NP[NOM]} \rangle \\ \text{COMPS} \langle \boxed{1}\text{NP[ACC]} \rangle \oplus \boxed{\mathbf{L}} \\ \text{GOV} \langle \begin{bmatrix} \text{V[BASE]} \\ \text{SUBJ} \langle \boxed{1} \rangle \\ \text{COMPS} \boxed{\mathbf{L}} \\ \text{GOV} \langle \rangle \\ \text{LEX} + \end{bmatrix} \rangle \end{bmatrix} \end{bmatrix} \\ \text{LEX} + \end{bmatrix} \end{bmatrix}$$

Through GOV, the governing verb selects the governed base form verb (or verbal cluster). Furthermore the finite argument composition verb *zag* selects a nominative NP through its SUBJ-feature. As non-subject arguments it selects through its COMPS-feature first an NP which is to be structure-shared with the subject-argument of the governed verb(s) and secondly the list $\boxed{\mathbf{L}}$ of zero or more non-subject arguments of the governed verb(s). The NP, which is lexically forced to be accusative, and the list $\boxed{\mathbf{L}}$ are to be concatenated into one list, as is indicated by the $\oplus$-operator in (10).

Thus we get a syntactic difference between modals and auxiliaries on the one hand and causatives and perceptuals on the other. We assume that with the former, structure-sharing is between the SUBJ-values of the governed verb and the governing verb. If the governing verb is finite, then it will force it's SUBJ-value to be nominative (the effect resembles that of the raising-to-subject transformation). So, intuitively, finite inflection brings about nominative case. Causatives and perceptuals always structure-share the governed verb's SUBJ-value with an accusative NP on their COMPS-list (which could be compared to the effect of raising-to-object), whether they are finite or nonfinite.

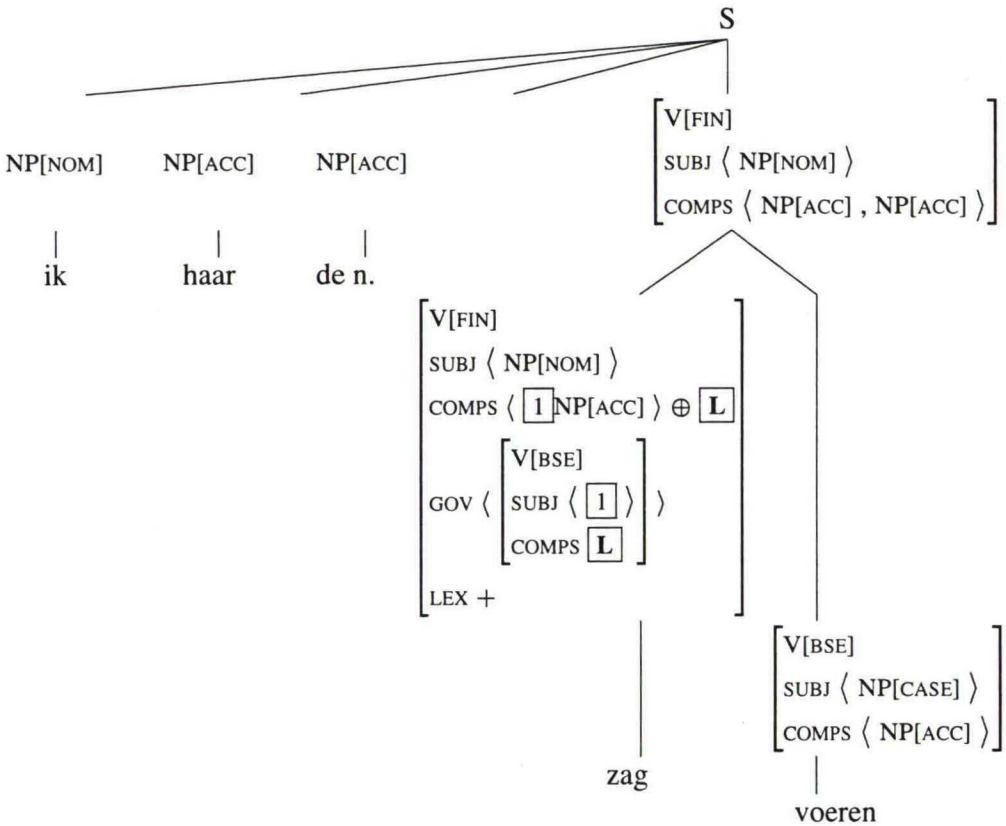
4 An Analysis of Dutch Cross Serial Dependencies

Now the stage is set for our analysis of cross serial dependencies in Dutch, which we illustrate first using example (1a), repeated here as (11):

(11) omdat ik_1 haar_2 de nijlpaarden$_3$ zag_1 voeren_3
because I[NOM] her[ACC] the hippos saw[FIN] feed[BSE]
"because I saw her feed the hippos"

Through ID-schema (4b) and the Valence Principle we can license a sign which is as displayed in (12) where (10) is the head which has selected (7) through its GOV-feature:

Figure 2: Argument Composed Cluster with a Perceptual Verb



(12)
$$\left[\begin{array}{l}\text{PHON} \quad \langle\, \textit{zag voeren} \,\rangle \\ \text{SYNSEM} \left[\text{LOC} \left[\text{CAT} \left[\begin{array}{l}\text{HEAD} \left[\begin{array}{l}\text{MAJOR V} \\ \text{VFORM FIN}\end{array}\right] \\ \text{SUBJ} \langle\, \text{NP[NOM]} \,\rangle \\ \text{COMPS} \langle\, \text{NP[ACC]} , \text{NP[ACC]} \,\rangle \\ \text{GOV} \langle\rangle\end{array}\right]\right]\right] \\ \text{LEX} +\end{array}\right]$$

As it were in passing, the perceptual verb imposes accusative case on the NP which corresponds with the subject-argument of the governed verb. This is because, as discussed in section 3, the governing verb will assign case to the subject argument of the governed verb through the structure-sharing indicated by "1"in (10) and Figure 2. Also, the list **L** in (10) (the entry for the governing verb "*zag*") will become instantiated by the list of length one which constitutes the COMPS-list of the governed verb "*voeren*".

The argument composition only accounts for the possibility of the discontinuity: we still have to account for the fact that the dependencies are cross serial in Dutch, whereas they are nested in German. We do so through two language specific linear precedence rules:

(13) Linear Precedence Rule Dutch Verb Clusters
[GOV ⟨ X ⟩] < X

(14) Linear Precedence Rule German Verb Clusters
X < [GOV ⟨ X ⟩]

By these LP-rules, in each part of the binary branching verb cluster the governing verb will appear head-initial in Dutch, and head-final in German. Our analysis also accounts for the recursive case, for example also for sentence (1b), repeated here as (15):

(15) dat ik_1 $haar_2$ hem_3 de nijlpaarden$_3$ zag_1 $helpen_2$ $voeren_3$
that I[NOM] her[ACC] him[ACC] the hippos saw[FIN] help[BSE] feed[BSE]
"that I saw her help him feed the hippos"

We account for this sentence and all similar constructions with more than one base form verb by defining base form entries for causatives and perceptuals along the same lines as the entry in (10). Note that, e.g., "*helpen*" has the valence that under our proposal all base form verbs have, and assigns case "CASE" to its subject argument. Other than that, "*helpen*" and all other nonfinite causatives and perceptuals are identical to finite causatives and perceptuals. Cf. Figure 3 for "*helpen*" as it appears in (15).

The syntactic derivation of (15) will proceed as the derivation of (11), and in fact we can account for any number of embeddings and any number of arguments in the Mittelfeld of DCSDs by repeated argument composition, since our entries for causatives and perceptuals allow us to simply argument compose a governing verb with governed base form verbs over and over. As is standardly assumed, this process is merely bounded by processing constraints.

5 Irregular Case Assignments in German and Dutch

Our analysis also gives us a handle on the contrast between (16b) and (16a);

(16) a. * ...daß er ihn das Lied singen half
...that he[NOM] him[ACC] the song sing[BSE] helped[FIN]
"that he helped him sing the song"

b. ...daß er ihm das Lied singen half
...that he[NOM] him[DAT] the song sing[BSE] helped[FIN]

We only have to postulate an entry for German *helfen*("help") which is like the entries for Dutch causatives and perceptuals which we motivated above, but make the verb assign DAT to the NP which structure-shares with the SUBJ-value of the governed V[BSE]. In the picture, Figure 4, this is realized through the structure-sharing, notated by "$\boxed{1}$",

Figure 3: Recursion in the Verb Cluster (Sentence (1b)).

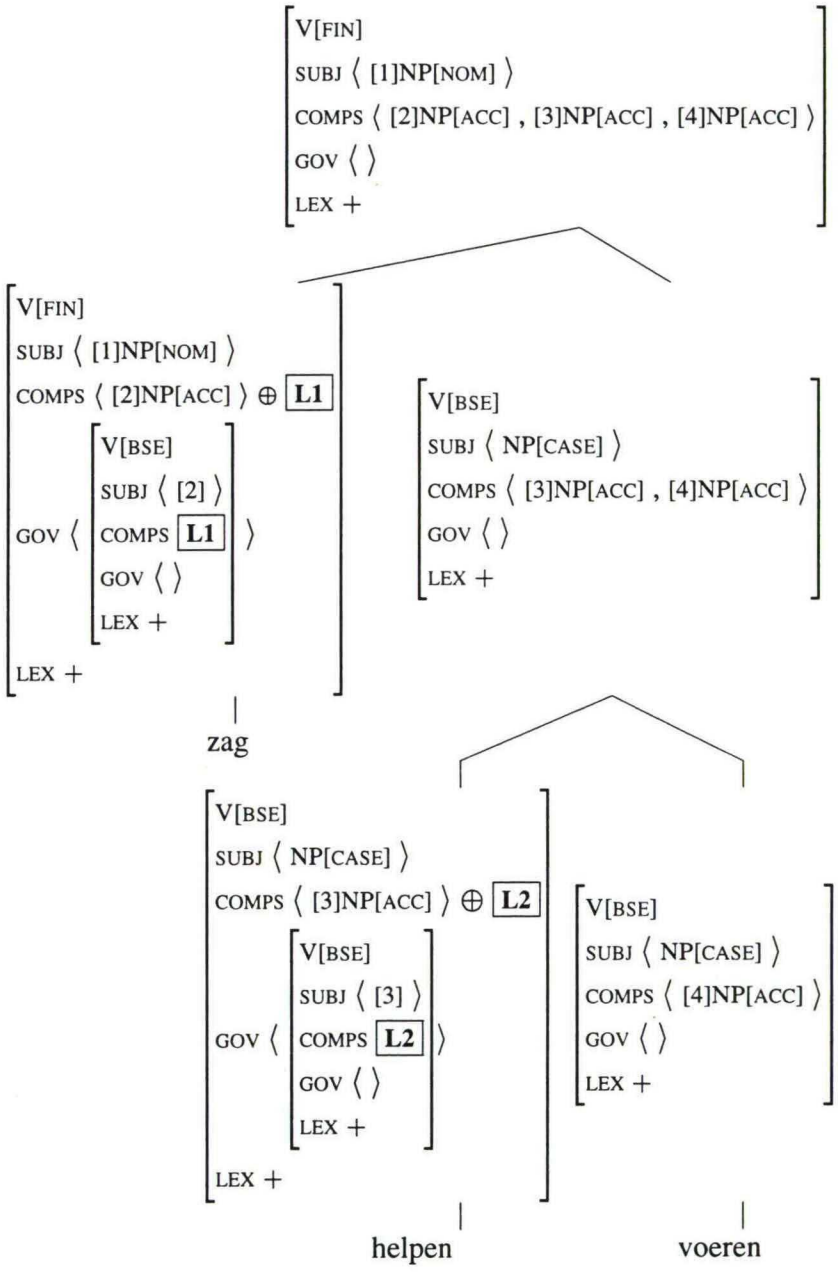
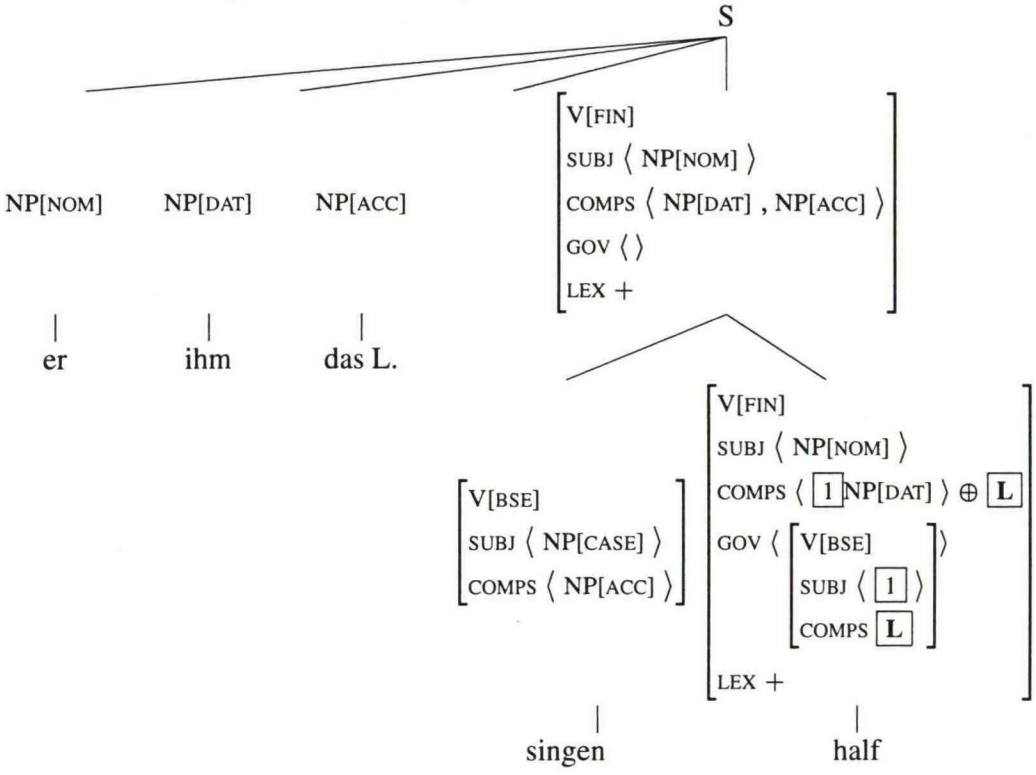


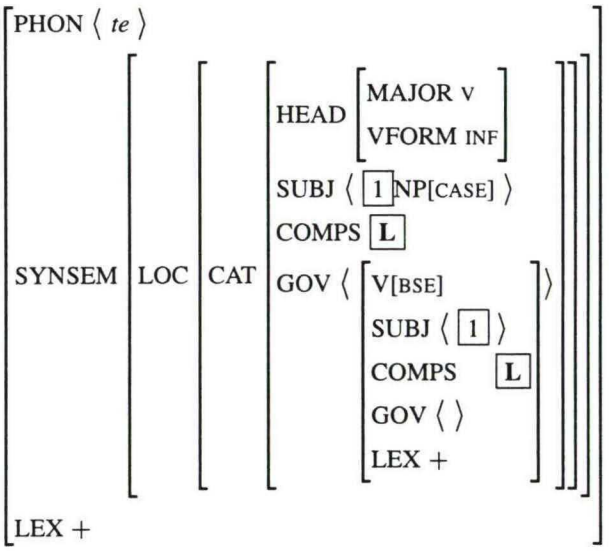
Figure 4: Dative Subjects and the German Verb Cluster



between the first element on the COMPS-list of *half* which gets dative case and the unrealized caseless subject of *singen*. Other than imposing dative instead of accusative on this NP, we only have to refer to the LP-rule in (14) to account for the head-final word order in the verbal cluster in (16b). This gives us the nested dependencies of German instead of the cross serial dependencies of Dutch for the verbal cluster of (16b), cf. Figure 4.

Also, our analysis is fully compatible with existing HPSG-analyses of raising and control. We just have to assume entries for infinitive markers like Dutch *te* with the value INF for the feature VFORM and make them argument compose with base form verbs just as we proposed for modals, but without assigning case to the subject. As any infinitive marker will be the head of a binary branching 'verb' cluster, by the Head Feature Principle such a cluster, like *te vertrekken*("to leave"), will be [VFORM INF] instead of [VFORM BSE]. This distinguishes 'true' infinitives from 'bare' infinitives.

So, in a sense, the infinitive marker combines with some base form verb to form a verbal cluster which is of a different category, but of the same valence as the base form verb.

(17)


Moreover, this allows us to explain why the subject of infinitival complexes can sometimes be accusative in Dutch, as in (18a), whereas it is usually nominative, cf. (18b):

(18) a. dat Jan hem dwingt te vertrekken
that Jan him[ACC] forces [to leave][INF]
"that Jan will force him to leave"

b. dat hij schijnt te vertrekken
that he[NOM] seems [to leave][INF]
"that he seems to leave"

All we have to is define raising verbs like "*schijnt*" along the lines of modals, as discussed for "*kan*" (cf. (9)), and define finite verb-forms like "*dwingt*" along the lines discussed for causative and perceptual verbs, cf. (10). The only difference should be that raising verbs like "*schijnt*" and object control verbs like "*dwingt*" demand that their GOV-value be VFORM INF instead of VFORM BSE.

6 Argument Composing Verbs as Heads of Root Clauses

With respect to word order in Dutch main clauses, we follow the (Pollard forthc.) analysis of German and assume that the verb appears head initial because it is marked as INV+ while linear precedence rule LPR1, a.o., applies:

(19) a. LPR1: [INV+] < [. . .]

b. LPR2: [. . .] < [INV−]

c. LPR3: C' << C"

d. LPR4: GOV ⟨ X ⟩ < X

So, by LPR1 a [INV+]-verb precedes all its sisters, "[. . .]" meaning "whatever sign". In contrast, a [INV−]-verb follows all its sisters by LPR2. Thus LPR2 accounts for the

SOV-order in all complement clauses, where presumably the complementizer always forces the finite verbal head of such a complement clause to be [INV−]. In addition, LPR3 orders arguments in a linear order determined by the obliqueness-relation "<<" and LPR4 is the rule of which we argued above that it orders governing verbs to the left of governed verbs in Dutch verb clusters.

Now we will integrate with this tentative account of Dutch word order an analysis of constructions like the following yes/no-interrogatives:

(20) a. Kan hij het konijn voeren ?
Can[FIN] he[NOM] the rabbit feed[BSE] ?
"Is he able to feed the rabbit ?"

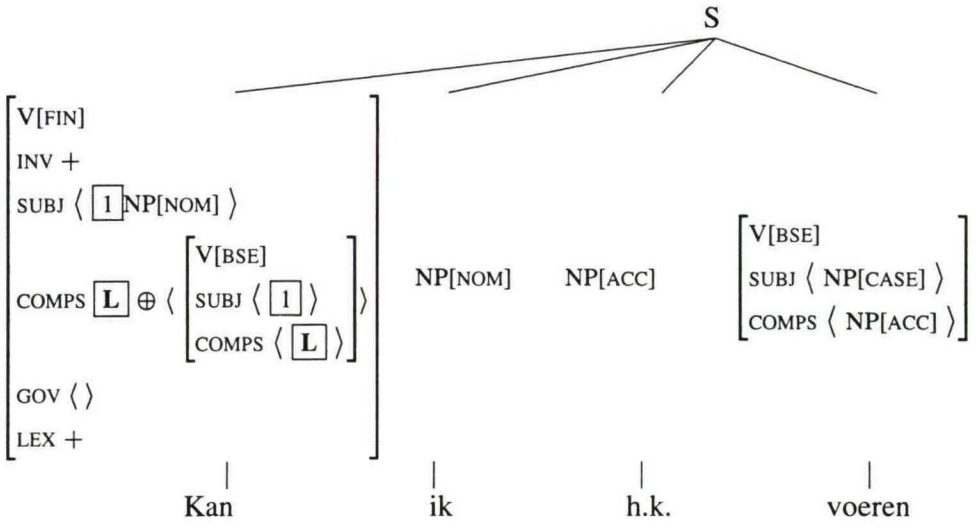
b. Zag Karina hem het konijn voeren ?
Saw[FIN] Karina him[ACC] the rabbit feed[BSE] ?
"Did Karina see him feed the rabbit ?"

As the analysis stands, these constructions cannot be licensed as instances of either ID-schema (4a) or (4b): this caveat we will take care of presently. First we must assume that since syntactically they behave like auxiliaries, causatives and perceptuals should actually be marked as auxiliaries, that is, they must be [AUX+]. If this is the case across the lexicon, then the following lexical rule might apply to all auxiliaries, modals, causatives and perceptuals alike:

$$\begin{bmatrix} \text{HEAD} \begin{bmatrix} \text{MAJOR} & \text{V} \\ \text{VFORM} & \text{FIN} \\ \text{AUX} & + \\ \text{INV} & - \end{bmatrix} \\ \text{SUBJ} \langle \text{NP[NOM]} \rangle \\ \text{COMPS} \ldots \boxed{\text{L}} \\ \text{GOV} \langle \begin{bmatrix} \text{V[BASE]} \\ \text{SUBJ} \langle \text{NP} \rangle \\ \text{COMPS} \boxed{\text{L}} \\ \text{LEX} + \end{bmatrix} \rangle \\ \text{LEX} + \end{bmatrix} \Rightarrow \begin{bmatrix} \text{HEAD} \begin{bmatrix} \text{MAJOR} & \text{V} \\ \text{VFORM} & \text{FIN} \\ \text{AUX} & + \\ \text{INV} & + \end{bmatrix} \\ \text{SUBJ} \langle \text{NP[NOM]} \rangle \\ \text{COMPS} \ldots \boxed{\text{L}} \oplus \langle \begin{bmatrix} \text{V[BASE]} \\ \text{SUBJ} \langle \text{NP} \rangle \\ \text{COMPS} \boxed{\text{L}} \\ \text{LEX} + \end{bmatrix} \rangle \\ \text{GOV} \langle \rangle \\ \text{LEX} + \end{bmatrix}$$

The input for this lexical rule are all [AUX+]-entries which are [INV−], that is verbs which will (by LPR2) appear clause final as in the SOV-order examples in all previous sections. Note that these input-entries are defined as "governing" verbs: they have a non-empty value for GOV. In contrast, the output of this lexical rule is formed by entries which are INV+ and which have an *empty*-list value for GOV. Instead, these INV+-verbs, which will appear clause initial by LPR1, take the verb they argument

Figure 5: An Argument Composition Verb Heading a Root Clause



compose with as their most oblique non-subject argument, at the 'end' of COMPS. Such output-verbs will then function as heads in constructions like those in (20), cf. Figure 5, which are licensed by ID-schema (4a).

Except for Dutch modal verbs, also the perfective auxiliary *hebben*("have") syntactically should be defined more or less like a modal (cf. (9)). Additionally, this entry for "*hebben*" must demand that its V[BSE]-argument is marked [AUX+]. By the assumptions made in this section, the set of [AUX+]-verbs comprises modals, causatives and perceptuals. In this way we account for the fact that "*hebben*" always selects two or more base form verb complements, as in example (6), the double infinitive construction. It must be 'double' in the sense that it should never be just one governed bare infinitive;

(21) * omdat Frits het nijlpaard niet [heeft voeren]
because Frits the hippo not has[FIN] feed[BSE,AUX−]
"because Frits not has feed the hippo"

This constraint we can now impose by making "*hebben*" select a verb cluster which is specified as [AUX+]. Then a main verb, like "*voeren*", which is [AUX−], must always be governed by modals, causatives and/or perceptuals to participate in a [AUX+]-verbal cluster that can be selected as an argument by perfective "*hebben*".

Again following (Pollard forthc.), we derive verb second clauses through extraction from flat VSO root clauses and through appealing to an LP-rule which orders fillers to the left of headed phrases. A traceless theory of extraction in Dutch, motivated along the lines of (Fodor & Sag 1994) and formalized along the lines of (Pollard & Sag 1994), chapter 9, can be integrated with this approach so as to account for Dutch verb second and to account for Dutch preposition stranding (see (van Riemsdijk 1978)), as we have discussed in (Rentier 1993).

References

Abeillé, A. & D. Godard (1994): "The Complementation of French Auxiliaries." *Proceedings of 13th West Coast Conference on Formal Linguistics*, USA

Baker, K. (1994): "An Integrated Account of Modal Flip and PVP-Fronting in German." *Proceedings of the 30th Regional Meeting of the Chicago Linguistic Society*, Chicago, USA

Borsley, R. (1987): "Subjects and Complements in HPSG." Stanford University, USA. CSLI Report 107

Chung, C. (forthc.): "Korean Auxiliary Verb Constructions Without VP Nodes." In: S. Kuno (ed.) *Harvard Studies in Korean Linguistics V*. Hanshin, Seoul

Evers, A. (1975): "The Transformational Cycle in Dutch and German." Indiana University Linguistics Club, Indiana, USA. PhD Thesis

Fodor, J. & I. A. Sag (1994): "Extraction without Traces." *Proceedings of 13th West Coast Conference on Formal Linguistics*, San Diego, USA

Hinrichs, E. & T. Nakazawa (1989): "Flipped out: AUX in German." *Proceedings of the 25th Regional Meeting of the Chicago Linguistic Society*, Chicago, USA

Miller, P. (1992): *Clitics and Constituents in Phrase Structure Grammar*. New York: Garland, USA

Monachesi, P. (1993): "Restructuring Verbs in Italian HPSG Grammar." *Proceedings of the 29th Regional Meeting of the Chicago Linguistic Society*, Chicago, USA

Nerbonne, J. (forthc.): "Partial Verb Phrases and Spurious Ambiguities." In: J. Nerbonne, K. Netter & C. Pollard (eds.) *German Grammar in HPSG*. (working title: forthc.) CSLI Publications and Chicago University Press, USA

Pollard, C. (forthc.): "On Head Non-Movement." In: H. Bunt & A. van Horck (eds.) *Proceedings of the Symposium on Discontinuous Constituency*. Mouton-de Gruyter, Germany

Pollard, C. & I. A. Sag (1987): *Information-Based Syntax and Semantics: Fundamentals*. CSLI Publications, Stanford, USA

_____ (1994): *Head-driven Phrase Structure Grammar*. University of Chicago Press and CSLI Publications, USA

Rentier, G. (1993): "Dutch Object Clitics, Preposition Stranding and ATB-Extraction." *Papers from Computational Linguistics in the Netherlands (CLIN) III, 1992*, Tilburg, the Netherlands

Riemsdijk, H. van (1978): "A Case Study in Syntactic Markedness: The Binding Nature of Prepositional Phrases." Foris , The Netherlands. PhD Thesis

Steedman, M. (1985): "Dependency and Coordination in Dutch and English." *Language: 61* :523–568.

OVERVIEW OF ITK RESEARCH REPORTS

No	Author	Title
1	H.C. Bunt	On-line Interpretation in Speech Understanding and Dialogue Sytems
2	P.A. Flach	Concept Learning from Examples Theoretical Foundations
3	O. De Troyer	RIDL*: A Tool for the Computer-Assisted Engineering of Large Databases in the Presence of Integrity Constraints
4	M. Kammler and E. Thijsse	Something you might want to know about "wanting to know"
5	H.C. Bunt	A Model-theoretic Approach to Multi-Database Knowledge Representation
6	E.J. v.d. Linden	Lambek theorem proving and feature unification
7	H.C. Bunt	DPSG and its use in sentence generation from meaning representations
8	R. Berndsen and H. Daniels	Qualitative Economics in Prolog
9	P.A. Flach	A simple concept learner and its implementation
10	P.A. Flach	Second-order inductive learning
11	E. Thijsse	Partical logic and modal logic: a systematic survey
12	F. Dols	The Representation of Definite Description
13	R.J. Beun	The recognition of Declarative Questions in Information Dialogues
14	H.C. Bunt	Language Understanding by Computer: Developments on the Theoretical Side
15	H.C. Bunt	DIT Dynamic Interpretation in Text and dialogue
16	R. Ahn and H.P. Kolb	Discourse Representation meets Constructive Mathematics

No	Author	Title
17	G. Minnen and E.J. v.d. Linden	Algorithmen for generation in lambek theorem proving
18	H.C. Bunt	DPSG and its use in parsing
19	H.P. Kolb	Levels and Empty? Categories in a Principles and Parameters Approach to Parsing
20	H.C. Bunt	Modular Incremental Modelling Belief and Intention
21	F. Dols	Compositional Dialogue Referents in Prase Structure Grammar
22	F. Dols	Pragmatics of Postdeterminers, Non-restrictive Modifiers and WH-phrases
23	P.A. Flach	Inductive characterisation of database relations
24	E. Thijsse	Definability in partial logic: the propositional part
25	H. Weigand	Modelling Documents
26	O. De Troyer	Object Oriented methods in data engineering
27	O. De Troyer	The O-O Binary Relationship Model
28	E. Thijsse	On total awareness logics
29	E. Aarts	Recognition for Acyclic Context Sensitive Grammars is NP-complete
30	P.A. Flach	The role of explanations in inductive learning
31	W. Daelemans, K. De Smedt and J. de Graaf	Default inheritance in an object-oriented representation of linguistic categories
32	E. Bertino and H. Weigand	An Approach to Authorization Modeling in Object-Oriented Database Systems
33	D.M.W. Powers	Multi-Modal Modelling with Multi-Module Mechanisms: Autonomy in a Computational Model of Language

No	Author	Title
34	R. Muskens	Anaphora and the Logic of Change*
35	R. Muskens	Tense and the Logic of Change
36	E.J. v.d. Linden	Incremental Processing and the Hierarchical Lexicon
37	E.J. v.d. Linden	Idioms, non-literal language and knowledge representation 1
38	W. Daelemans and A. v.d. Bosch	Generalization Performance of Backpropagation Learning on a Syllabification Task
39	H. Paijmans	Comparing IR-Systems: CLARIT and TOPIC
40	R. Muskens	Logical Omniscience and Classical Logic
41	P. Flach	A model of induction
42	A. v.d. Bosch and W. Daelemans	Data-oriented Methods for Grapheme-to-Phoneme Conversion
43	W. Daelemans, S. Gillis, G. Durieux and A. van den Bosch	Learnability and Markedness in Data-Driven Acquisition of Stress
44	J. Heemskerk	A Probabilistic Context-free Grammar for Disambiguation in Morphological Parsing
45	J. Heemskerk and A. Nunn	Dutch letter-to-sound conversion, using a morpheme lexicon and linguistic rules
46	A. HH. Ngu, R. Meersman and H. Weigand	Specification and verification of communication constraints for interoperable transactions
47	J. Jaspars and E. Thijsse	Fundamentals of Partial Modal Logic
48	E. Krahmer	Partial Dynamic Predicate Logic
49	W. Daelemans	Memory-Based Lexical Acquisition and Processing
50	G. Rentier	A Lexicalist Approach to Dutch Cross Serial Dependencies

ITK: P.O. BOX 90153 5000 LE TILBURG THE NETHERLANDS