Abstract
We present an analysis of Dutch cross serial dependencies in Head-driven Phrase Structure Grammar ([P&S(1994)]). We start out from the assumption that causative and perceptual verbs, like auxiliaries, can lexically 'raise' the arguments of the non-finite verbs they govern to their own list of arguments through "argument composition" ([H&N(1989)]).

1 Introduction
Dutch cross serial dependencies (DCSDs), well-known from (1) and (2), still challenge computational linguistics for an efficient treatment.

(1) dat ik haar de nijlpaarden zag voeren
"that I saw her feed the hippos"

(2) dat ik Henk haar de nijlpaarden helpen voeren
"that I saw Henk help her feed the hippos"

The problematic aspects of DCSDs are of course the bounded discontinuous relation between the NPs and the verbs of which they are arguments, indicated in (1) and (2) by the subscripted integers, and the recursiveness of the phenomenon. The construction is only licensed by members of two closed classes of verbs, the class of perceptual verbs like zien ("see"), horen ("hear") and voelen ("feel"), and the class of causative verbs like laten ("let/make") and helpen ("help"). In the analysis put forward here we emphasize this lexical aspect of the phenomenon; in our analysis DCSDs are strictly tied to the subcategorization and semantics of perceptual and causative verbs. We analyze them as verbs which select, apart from their subject, a nonfinite V-projection which denotes an event. More particularly, as is proposed for German auxiliaries in [H & N(1989)], they subcatego-

2 Event Semantics in HPSG
The choice of semantics in terms of a theory of events, known from [Davidson(1967)], offers interesting advantages and explanations of logical and linguistic phenomena, motivating the development of a constraint-based version of it. So, in the spirit of event semantics we propose that main verbs like voeren ("feed") in (3) should denote a discourse referent, which is in fact a very natural assumption. In (3) and throughout the paper, recurring [i]'s indicate structure sharing, that is token-identity of information, as is common usage in HPSG. Note also that we follow [Borsley(1987)] in representing subjects as values of S(subject) and follow [P&S(1994)] (chapter 9) in representing non-subject
arguments as values of COMPS.

\[
\text{LEX} + \approx \begin{array}{c}
\text{PHON} \quad \{\text{voeren}\} \\
\text{HEAD} \quad \{\text{major}\} \\
\text{SUBJ} \quad \{\text{NP}\{\text{case}\}\{2\}\} \\
\text{COMPS} \quad \{\text{NP}\{\text{acc}\}\{3\}\} \\
\text{GOV} \quad \{\} \\
\text{CONT} \quad \{\text{det} \quad \{\text{para}\}\{1\}\} \\
\text{RESTRI} \quad \{\text{reln} \quad \{\text{phon}\}\{1\}\} \\
\end{array}
\]

The constraint-based event semantics of the base form verb *voeren* as it is depicted in (3), with the quasi-determiner *kwant*, should be interpreted as an existentially quantified event with a parameter [1] which is restricted to involve a relation of *feeling*, an argument with the role of agent which is associated with a semantic content [2] and an argument associated with a semantic content [3] which is the theme.9

Here the value of DET is a 'shallow' representation of a quantifier,6 and the value of PARA, which is an abbreviation for 'parameter', is structure shared with the value of a feature INST which is short for 'instance'. We will suppose that the value of PARA corresponds with a discourse referent in the discourse representation associated with a natural language expression, without formally defining this relation here. The value of RESTRI, which abbreviates 'restrictions', is a set of constraints on the value of this parameter.

3 An Argument Composition Analysis

We assume that the clause structure of DCSDs is one where we have a binary left-branching verbal complex. This verbal complex then locally selects the sum of the arguments of the verbs which constitute it. We feel that a binary branching analysis is empirically motivated by auxiliary flip in the same way as auxiliary flip motivates a binary right-branching structure for the German verbal complex, following [H&N(1989)].

A governing auxiliary will apply argument composition and raise all the complements from the governed verb(s) to become arguments of the auxiliary, as proposed in [H&N(1989)]. We assume that causative and perceptual verbs syntactically behave just like auxiliaries in this respect.

The difference between auxiliaries on the one hand and causative and perceptual verbs on the other we view as basically semantic. We take it that auxiliaries semantically more or less operate on events, affecting features for tense and aspect or modality. Causative and perceptual verbs on the other hand will be analyzed as events themselves, events which take other events as their argument, in general as a theme (viz., a value of ARG2, cf. the entry in (7) below).

In chapter 9 of [P&S(1994)] the approach to local selection from [Borsley(1987)] is developed further and leads to the Valence Principle, which refers to the valence features SUBJ and COMPS through \(P\):

\[
\text{Valence Principle, Chapter 9, [P&S(1994)]}
\]

In a headed phrase, for each valence feature \(P\), the \(P\) value of the head-daughter is the concatenation of the phrase's \(P\) value with the list of SYNSYM values of the \(P\)-daughters' value.

The general effect of the principle on a phrase which is headed by some sign is that this headed sign can only become 'complete' (or 'saturated') if it is combined with the appropriate arguments. For example, in the case of a transitive verb, such a verb must find a subject NP (selected through SUBJ) and some object (selected through COMPS). If we assume a flat clause structure analysis of Dutch and we furthermore assume lexical signs like (3) and (7), then the immediate dominance statements (5) and (6) will suffice to describe the construction of Dutch we are concerned with here. 7 Here the \(H\), \(S\) and \(C\) indicate that the daughters of the phrase include a head, a subject and complements, not necessarily in that order (cf. chapter 9 of [P&S(1994)] for details). Note that in addition to the valence features SUBJ and COMPS, we also assume the presence of the GOV-feature, ranging over 1 complement: 8

\[
\begin{align*}
(5) & \quad \text{XP[lex]} \rightarrow \text{S}, \text{C}_1, \ldots, \text{C}_n, \text{I[go\(\langle C\rangle \), lex\(\langle C\rangle \)]} \\
(6) & \quad \text{X[lex]} \rightarrow \text{H[go\(\langle C\rangle \), lex\(\langle C\rangle \)]}, \text{C}_i
\end{align*}
\]

The second schema is in a sense not a "phrase" structure schema but is instead a "cluster-formation" schema. This is because normally the combination of two or more words leads to a sign which is lex+, a phrasal sign, but here it leads to a 'complex word' which is lex+. Also (6) is strictly binary: it takes one argument, namely the argument which is the value of...

---

819
GOV. We arrange the lexicon so that any value of GOV will always be an unsaturated base form verb which is defined as LEX+ as well. By the Valency Principle, this selection requirement of the governing verb will be appropriately 'cancelled' after string concatenation during parsing.

Central to our analysis of the case-markings of NPs in the Dutch Mittelfeld is the assumption from [Pollard(1986)] that base forms of verbs do not assign any case to their subject. The value for the subject-NP's CASE-feature in (3), "c^s~", is the supertype in the type hierarchy for those atomic types that are appropriate values of the feature CASE. So, the value CASH is the supertype of NOM and ACC in Dutch and English, and in German also of DAT and GEN. The result of assigning the subject-NP this supertype for case in practice boils down to giving this NP some kind of "any"-value for case; the case-value CASE of such an NP will unify with any other possible case value.

In our analysis, the discontinuous relation between arguments and verbs in DCSDs is brought about firstly by lexically defining finite perceptuals like zag (and finite causatives) as argument composition verbs, along the following lines:9

(7)  

\[
\begin{align*}
\text{PHON} & \quad \{ \text{ zag } \} \\
\text{HEAD} & \quad \text{VFORM} \quad \text{FIN} \\
\text{SUBJ} & \quad \{ \text{NP[NOM]} \} \\
\text{COMPS} & \quad \{ \text{NP[ACC]} \} \otimes \text{L} \\
\text{GOV} & \quad \{ \text{V[BASH]} \} \\
\text{SUBJ} & \quad \{ \text{LNP} \} \\
\text{COMPS} & \quad \text{L} \\
\text{CONT} & \quad \text{4} \\
\text{LEX+} & \\
\text{CONTENT} & \quad \text{DET EVENT} \\
\text{PARA} & \quad \text{6} \\
\text{RESTR} & \quad \{ \text{HELPHELM} \} \\
& \quad \{ \text{INST [6]} \} \\
& \quad \{ \text{ARG1 [1]} \} \\
& \quad \{ \text{ARG2 [4]} \}
\end{align*}
\]

The finite argument composition verb zag selects a singular nominative NP through its SUBJ-feature. As non-subject arguments it selects through its COMPS-feature first the NP tagged as [5] which is unified with the SUBJ-value of the governed verb(s), and secondly the list [L] of zero or more non-subject arguments of the governed verb(s). And crucially, being a governing verb, zag selects through GOV a governed base form verb,10 with as SUBJ-value [E], as COMPS-value \([\text{L}]\) and as semantics \([\text{A}]\). Note that, since the governed V[form] is selected as missing a subject and a list of complements, it must not 'find' this subject or these complements, which it indeed doesn't (cf. the tree in Figure 1).

As it were in passing, the governing perceptual verb (or causative verb alike) imposes accusative case on the NP which denotes the subject-argument of the governed verb. The unification of \([\text{CASE CASE}]\) and \([\text{CASE ACC}]\) will be forced through the structure-sharing indicated in (7) as \([\text{E}]\), and will result in the more specific restriction \([\text{CASE ACC}]\). This accounts for the accusative case-marking on haar ("her") in examples (1) and (2), and in general on all non-subject arguments in such constructions.

The second and crucial step in our account of the discontinuity is accounting for the linear order in the verb cluster with DCSDs. The linear order of the verb cluster in Dutch we account for through (8):

(8) Linear Precedence Rule Dutch Verb Clusters

\[
\text{GOV} \{ \text{X} \} < \text{X}
\]

(9) Linear Precedence Rule German Verb Clusters

\[
\text{X} < \text{GOV} \{ \text{X} \}
\]

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.11 It is straightforward to show that the above approach has the desired effect also for the sentence (2) mentioned in the introduction if we define a lexical entry for the causative helpen with a syntax and semantics along the same lines as the perceptual zag. The only difference must be that such nonfinite entries do not assign NOM to their subject, but "CASH". Other than that, there will just be additional embeddings in the semantics as well as in the verb cluster. Thus, by the ID-rule in (6) and the lexical entries for causatives and perceptuals, we account for the recursiveness of the phenomenon, cf. the tree in Figure 2.

4 Conclusion

We extended the [H&N(1989)]-analysis of German to Dutch, accounting for the difference, resp. nested vs. cross serial dependencies, through one single LP-parameter. Also, we argued that such an argument composition approach is to be preferred over several alternative approaches, since argument composition isn't an 'additional' mechanism. Further linguistic advantages of this approach, i.e. accounts of irregular case assignments and constraints on double infinitives, are discussed in [Rentier(1994)]. We are able to derive verb second constructions by standard application of the ID-schema in (6) either will be LEX+, so that we are able to recursively build up bigger and bigger LEX+-complexes.

9In this entry and throughout the paper, \(\otimes\) stands for concatenation of arbitrary-length lists of arguments.

10One base form verb, or a base form verb-headed verbal cluster...
Figure 1: The discontinuous relation: Valence Principle, schema's (5) & (6), entries (3) & (7), LP-rule (8).
the Dutch versions of the extraction lexical rules (see Rentier(1993)) to the verbs at the lexical level.

References

[Borsley(1987)] R. Borsley, “Subjects and Complements in HPSG”, CSLI Report 107, Stanford University, USA

[Davidson(1967)] D. Davidson, “The Logical Form of Action Sentences”, reprinted in “Essays on Actions and Events”, Clarendon Press, Oxford, 1980

[H&N(1989)] E. Hinrichs, T. Nakazawa, “Flipped out: AUX in German”, in “Proceedings of the 23rd Regional Meeting of the Chicago Linguistic Society”, CLS, USA

[Pollard(1994)] C. Pollard, “On Head Non-Movement”, in: Hunt, H. & van Horck, A. (eds.), “Proceedings of the Symposium on Discontinuous Constituency”, Mouton-de Gruyter, Germany

[P&S(1994)] C. Pollard, I.A. Sag, “Head-driven Phrase Structure Grammar”, University of Chicago Press and CSLI Publications, USA

[Reape(1994)] M. Reape, “Getting Things in Order”, in: Hunt, H. & van Horck, A. (eds.), “Proceedings of the Symposium on Discontinuous Constituency”, Mouton-de Gruyter, Germany

[Rentier(1994)] G. Rentier, “Dutch Object Clitics, Preposition Stranding and Across-the-Board Extraction”, in Sijtsma, W. & Zweeckhorst, O. (eds.), “Papers from Computational Linguistics in the Netherlands (CLIN) III, 1992”, the Netherlands