

Between VP Adjuncts and Second Pole in Dutch

A corpus based survey regarding the complements between VP adjuncts and second pole in Dutch

Tim Van de Cruys

K.U.Leuven

Abstract

In Dutch, verbs are situated at fixed places in the sentence. Those places are called the first and second pole. VP adjuncts seem to function as some kind of pivot place in between these poles. This article investigates, by means of corpus research in the Spoken Dutch Corpus (CGN), which elements are intervening between these VP adjuncts and the second pole. Attention is particularly paid to the reasons and principles that make elements end up between VP adjuncts and second pole. First of all, these elements will often be syntactically and semantically linked to the main verb. Secondly, the functional sentence perspective will be important for the placement of elements before or behind the VP adjuncts. The results will show that the functional sentence perspective is one of the main information dividing principles in Dutch sentences. The functional sentence perspective is then implemented in Head-Driven Phrase Structure Grammar, extending Van Eynde's theory about Argument Realization in Dutch. Being able to handle focus information in an adequate way is important for contemporary issues such as coreference resolution. A better understanding of the principles that order the complements of the Dutch verb will also be helpful in correctly analyzing and parsing Dutch sentences.

1 Concepts

1.1 The Structure of Dutch Clauses

The Dutch grammar *ANS* (Haeseryn, Romijn et al. 1997) describes a Dutch main clause on the basis of a first and second pole, occupied by the verbs. The *Mittelfeld*, in between these two poles, contains three parts. The central part of the *Mittelfeld* is occupied by different kinds of VP adjuncts.

The structure of a Dutch subclause is quite different. The first pole is occupied by a conjunction, that connects the subclause to the main clause. The actual subclause starts with the *Mittelfeld*. All the verbs are put on the second pole.

	1st sentence position	1st pole	MITTELFELD			2nd pole	last sentence position
			1	2 VP Adjuncts	3		
a	Ik	heb	Jan	gisteren	een boek	gegeven	
	I	have	Jan	yesterday	a book	given	
<i>I've given a book to Jan yesterday</i>							
b	–	(...dat)	ik Jan	gisteren	een boek	gegeven heb	
	–	(...that)	I Jan	yesterday	a book	given have	
<i>...that I've given a book to Jan yesterday</i>							

Table 1: The structure of a Dutch main clause (a) and subclause (b)

The actual position in which constituents end up, depends on all kinds of different ordering principles. We will only have a look at the ordering principles that are important for the position between VP adjuncts and second pole.

1.2 Between VP Adjuncts and Second Pole

There are two main reasons why elements are ending up between VP adjuncts and second pole. First of all, some constituents are **inherently connected to the main verb**. Sentence (1) gives an example. The predicative complement *groen* needs to be put between the VP adjuncts and the second pole. A sentence with the predicative complement between first pole and VP adjuncts, as in (2), is ill-formed.

- (1) Ze hebben dat hekje gisteren groen geverfd.
 they have that fence yesterday green painted
 ‘They have painted that fence green yesterday.’
- (2) *Ze hebben dat hekje groen gisteren geverfd.
 they have the fence green yesterday painted

Secondly, the **functional sentence perspective** makes the most informative elements end up between VP Adjuncts and second pole. The ANS states that the VP adjuncts in the Mittelfeld function as some kind of pivot place: elements that are less informative appear before the VP adjuncts, more informative elements appear behind them. Sentence (3) gives an example. Sentence (4), where the new information is put between first pole and VP adjuncts, is highly questionable.

- (3) Ik zal je morgen een boek geven.
 I will you tomorrow a book give
 ‘I will give you a book tomorrow.’
- (4) ?Ik zal je een boek morgen geven.
 I will you a book tomorrow give

It should be noted that the position between VP adjuncts and second pole is not the only position that is signaling focus information. This can also be the case with first sentence position. The first sentence position, which is the normal position for the subject in the main clause, is signaling focus information if it is taken by a constituent other than the subject, as in (5).

Also, the last sentence position often attracts extra attention, as in (6), although there are other reasons why a constituent might end up in this position.

- (5) Een boek zal ik je morgen geven.
 A book will I you tomorrow give
 ‘I will give you a book tomorrow.’
- (6) Ik heb gisteren een boek gegeven aan Jan.
 I have yesterday a book given to Jan
 ‘I’ve given a book to Jan yesterday.’

But in this paper, we will mainly focus on the position between VP adjuncts and second pole.

2 Methodology

The corpus research was carried out in the syntactically annotated part of CGN (Spoken Dutch Corpus). As has been indicated by van der Wouden et al. (2003), the CGN is a new resource for research into contemporary spoken Dutch that is well suited for carrying out statistical research in order to shed some light on certain linguistic issues. Only the Flemish Dutch part was used, as the Northern Dutch part was not available at the time of research. This corpus contained 42479 sentences, of which only the clauses with VP adjuncts and second pole were retained. The remainder consists of 3879 main clauses and 3309 subclauses. This corpus was searched with the syntactic search program TIGERSearch.

A statistical approach was taken in investigating the constituents between VP adjuncts and second pole. A number of queries were developed, that determined, for all complements of the verb (subject, direct object, indirect object, ...), in how many cases they end up between VP adjuncts and second pole. This number is then compared to the other possible places in which the complements can appear. This gives the following possibilities:

- Main clause:
 1. First sentence position
Aan Jan heb ik gisteren een boek gegeven
 2. Between first pole and VP adjuncts
Ik heb Jan gisteren een boek gegeven
 3. Between VP adjuncts and second pole
Ik heb dat boek gisteren Aan Jan gegeven
 4. Final sentence position (extraposition)
Ik heb dat boek gisteren gegeven aan degene die het graag wilde hebben
- Subclause
 1. Before VP adjuncts
... dat ik Jan gisteren een boek gegeven heb
 2. Between VP adjuncts and second pole
... dat ik dat boek gisteren aan Jan gegeven heb
 3. Final sentence position (extraposition)
... dat ik dat boek gisteren gegeven heb aan degene die het graag wilde hebben

3 Results

3.1 Subject

Table 2 gives the percentages of the different subject positions in a Dutch clause. In the main sentence, the subject can appear in three places: in first sentence position, between first pole and VP adjuncts, and between VP adjuncts and second pole. The results show that the subject appears only in few cases between VP adjuncts and second pole. In the majority of cases, the subject comes before the VP adjuncts (96.40%).

The results of the subclause are similar. The subclause lacks a first sentence position, but about 90% of the subjects appear before the VP adjuncts.

position	main clause		subclause	
	n	%	n	%
1st sentence position	2565	68.99%	–	–
1st pole - VP adjuncts	1019	27.41%	2486	89.88%
VP adjuncts - 2nd pole	120	3.23%	247	8.93%
extraposition	14	0.37%	33	1.19%
total	3718	100.00%	2766	100.00%

Table 2: The position of the subject

The interpretation of these results is quite straightforward: the functional sentence perspective is responsible for the distribution of the subject. The subject is usually a known entity, to which an unknown attribute is assigned (7). In passive sentences, however, there are some cases in which the subject can appear between VP adjuncts and second pole. This is the case if there is a pronoun which anticipates the subject (8), or if an adjunct acquires first sentence position (9). The subject then gets the focus of the sentence.

- (7) Ik heb gisteren een koffie gedronken.
I have yesterday a coffee drunk
'I have drunk a coffee yesterday.'
- (8) Er wordt ook wijn gedronken.
there is also wine drunk
'Wine is also drunk.'
- (9) In de krant zijn toen veel spellingsbijlagen verschenen.
in the newspaper are then many spelling supplements published
'Many spelling supplements have been published in the newspaper at that moment.'

Note that, in order to put the subject in focus position, it needs to be placed between VP adjuncts and second pole. This is the only possibility to give focus to the subject,

because the first sentence position is the normal, unmarked position of the subject. For all the other complements, first sentence position (topicalisation) is a marked position, and hence attains focus.

3.2 Indirect Object

Table 3 gives the results of the indirect object. Clearly, the indirect object occurs mostly between first pole and VP adjuncts: in about 3 out of 4 cases in both main clause and subclause.

position	main clause		subclause	
	n	%	n	%
1st sentence position	7	5.22%	–	–
1st pole - VP adjuncts	100	74.63%	27	72.97%
VP adjuncts - 2nd pole	13	9.70%	7	18.92%
extraposition	14	10.45%	3	8.11%
total	134	100.00%	37	100.00%

Table 3: The position of the indirect object

The functional sentence perspective is again responsible for the distribution of the indirect object over the different positions in the clause. But this does not explain why there are more indirect objects that appear before the VP adjuncts. Upon examining the data a bit closer, an explanation comes up: the majority of the clauses with an indirect object is built according to the structure *first pole + personal pronoun + VP adjuncts (+ direct object) + second pole*, as in sentence (10). So in the majority of cases, the indirect object consists of a personal pronoun (a known entity) that does not bear the focus of the sentence. Hence, it is not put in focus position. If the indirect object is put into first sentence position or between VP adjuncts and second pole, it clearly bears the focus of the sentence, as in (11) and (12).

- (10) Ik zal je meteen een voorbeeld geven.
I will you immediately an example give
'I will give you an example immediately.'
- (11) Aan Jan zal ik dat boek morgen geven.
to Jan will I that book tomorrow give
'I will give that book to Jan tomorrow.'
- (12) ... dat ik dat boek morgen aan Jan zal geven.
that I that book tomorrow to Jan will give
'... that I will give that book to Jan tomorrow.'

3.3 Direct Object

Table 4 gives the results of the direct object's position. The four positions are possible, but most of the direct objects end up in the Mittelfeld. The number of direct objects that is put before the VP adjuncts and behind the VP adjuncts is about the same.

position	main clause		subclause	
	n	%	n	%
1st sentence position	244	12.10%	–	–
1st pole - VP adjuncts	797	39.51%	492	40.39%
VP adjuncts - 2nd pole	737	36.54%	590	48.44%
extraposition	239	11.85%	136	11.17%
total	2017	100.00%	1218	100.00%

Table 4: The position of the direct object

Again, the functional sentence perspective is responsible for the position of the direct object. Compare sentences (13) and (14).

(13) Ik heb dat boek gisteren aan Jan gegeven.
I have that book yesterday to Jan given
'I've given that book to Jan yesterday.'

(14) Ik heb Jan gisteren een boek gegeven.
I have Jan yesterday a book given
'I've given that book to Jan yesterday.'

(15) *Ik heb een boek gisteren aan Jan gegeven.
I have a book yesterday to Jan given

In sentence (13), the direct object *dat boek* is presented as a known entity, while the indirect object *aan Jan* gets the focus. Sentence (14) gives the opposite situation: the indirect object *Jan* is known, but it is not known that *een boek* has been given to him.

The fact that the functional sentence perspective really does play an important role, is again proven by sentence (15): introducing an unknown object before the VP adjunct sounds awkward to the native speaker of Dutch.

3.4 Prepositional Complement

Sentence (16) is an example of a normal prepositional complement. But when discussing the prepositional complement (as well as the locative/directional complement in the next section), we need to take into account an extra particularity. In Dutch it is possible to split up a prepositional complement, if the head of the prepositional complement is a pronoun.¹ The pronoun is then put before the VP adjuncts, while

¹This particular construction will be coined *discontinuous prepositional complement*, as opposed to *full prepositional complements*.

the preposition comes after the VP adjuncts (17). Moreover, it is not possible to put the pronoun between VP adjuncts and second pole when referring to inanimate objects (18). This is again a clear indication that the functional sentence perspective plays an important role.

- (16) Ze hebben gisteren weer over voetbal gepraat.
they have yesterday again about soccer talked
'They have been talking about soccer again yesterday.'
- (17) Ze hebben er gisteren weer over gepraat.
They have there yesterday again about talked
'They have been talking about it again yesterday.'
- (18) *Ze hebben gisteren weer over het gepraat.
They have yesterday again about it talked

position	main clause		subclause	
	n	%	n	%
1st sentence position				
full	17	3.91%	–	–
discontinuous	35	8.05%	–	–
1st pole - VP adjuncts				
full	15	3.45%	12	3.93%
discontinuous	110	25.29%	44	14.43%
VP adjuncts - 2nd pole	126	28.97%	123	40.33%
extraposition	132	30.34%	126	41.31%
total	435	100.00%	305	100.00%

Table 5: The position of the prepositional complement

Table 5 gives the results of the prepositional complement. The results show that full prepositional complements mainly end up after the VP adjuncts. It seems that prepositional complements are either full and end up between VP adjuncts and second pole, or they are discontinuous, with the preposition between VP adjuncts and second pole, and the pronoun between first pole and VP adjuncts. There seems to be a strong link between the verb and its preposition, so that it needs to be realized near the verb. But when the prepositional contains known information, this conflicts with the functional sentence perspective. This is why the prepositional complement is split up, with the preposition realized near the verb (between VP adjuncts and second pole), and the pronoun moved between first pole and VP adjuncts.

With regard to these conflicting principles, sentence (19) is particularly interesting. It contains the referring pronoun *hem*, which is normally ending up between first pole and VP adjuncts due to the functional sentence perspective. This is shown by sentence (20), which expresses a similar meaning. Nevertheless, a sentence like (21) is questionable, because the preposition *over* is inherently connected to the main verb.

- (19) Ze hebben gisteren weer over hem gepraat.
they have yesterday again about him talked
'They have talked about him again yesterday.'
- (20) Ze hebben hem gisteren weer uitvoerig bediscussieerd.
they have him yesterday again ample discussed about
'They have talked a lot about him again yesterday.'
- (21) ?Ze hebben over hem gisteren weer gepraat.
they have about him yesterday again talked

3.5 Locative/Directional Complement

The locative/directional complement subsumes all complements that are designating a place or a direction. They are either prepositional (22) or adverbial (23). Again, discontinuous complements are possible (24).

- (22) Hij is dan naar dat eiland gezwommen.
he is then to that isle swum
'He has swum to that isle then.'
- (23) We zijn toen huiswaarts gekeerd.
we are then towards home turned
'We went home then.'
- (24) We gaan er bomen op planten.
we go there trees on plant
'We're going to plant trees on it.'

position	main clause		subclause	
	n	%	n	%
1st sentence position				
full	26	5.96%	–	–
discontinuous	7	1.61%	–	–
1st pole - VP adjuncts				
full	59	13.53%	71	16.95%
discontinuous	37	8.49%	33	7.88%
VP adjuncts - 2nd pole	261	59.86%	288	68.74%
extraposition	46	10.55%	27	6.44%
total	436	100.00%	419	100.00%

Table 6: The position of the locative/directional complement

Table 6 gives the results of the locative/directional complement. They are quite different to the results of the prepositional complement: full locative/directional complements appear more between first pole and VP adjuncts compared to the prepositional complement. Discontinuous complements are possible, but appear to a lesser extent compared to the prepositional complement. Most locative/directional complements end up between VP adjuncts and second pole.

The results might be explained as follows: the preposition is not inherently connected to the main verb, so that full complements can appear before the VP adjuncts. But the locative/directional complement most of the times contains focus information, so that it needs to be realized between VP adjuncts and second pole. Hence, there are also less discontinuous complements. So in this category, the functional sentence perspective plays an important role again.

3.6 Predicative Complement

The results of the predicative complement (table 7) are very straightforward. The majority of predicative complements ends up between VP adjuncts and second pole.

position	main clause		subclause	
	n	%	n	%
1st sentence position	13	4.13%	–	–
1st pole - VP adjuncts	9	2.86%	26	4.66%
VP adjuncts - 2nd pole	275	87.30%	510	91.40%
extraposition	18	5.71%	22	3.94%
total	315	100.00%	558	100.00%

Table 7: The position of the predicative complement

The explanation is simple: predicative complements are inherently connected to the main verb. They need to come obligatorily between VP adjuncts and second pole (except for some special cases like topicalisation). Compare sentence (25) and (26).

(25) Dat zal wel genoeg zijn.
that will well enough be
'That should be enough.'

(26) ?Dat zal genoeg wel zijn.
that will enough well be

The predicative complements occurring between first pole and VP adjuncts are due to spoken language characteristics, as in (27): normally, the modifier *uiteraard* comes before the predicative complement, but in spoken language one might modify the utterance 'on the fly', after having already uttered the predicative complement. Such a syntactic construction is not used in written language.

- (27) Dat mag de Nederlandse tekst uiteraard zijn.
 that may the Dutch text of course be
 ‘That may of course be the Dutch text.’

For a more detailed discussion of the predicative complements, see Van Eynde (this volume).

4 Implementation in HPSG

Van Eynde (this volume) presents an HPSG theory to capture the different sentence positions in Dutch. Instead of using the popular classification of arguments (SUBJ, SPR and COMPS), a difference is made between arguments that need to be realized near the verb (COMPS) and arguments that can be separated from the verb by VP adjuncts (L-ARGS). This theory provides an adequate basis to capture the conclusions that have been deduced from the corpus research. Van Eynde offers a description of the inherently connected arguments, that are dependent on semantic and syntactic factors. I will focus on the description of pragmatics, i.e. the functional sentence perspective, in HPSG.

4.1 Capturing the Functional Sentence Perspective

The way of coding focus information into HPSG is based on the approach of Engdahl and Vallduví (1996). They provide a theory to capture the functional sentence perspective in English and Catalan.

$$\left[\text{SYNSEM|LOCAL|CONTEXT|INFO-STRUCT} \left[\begin{array}{l} \text{FOCUS} \langle [\dots] \rangle \\ \text{LINK} \langle [\dots] \rangle \end{array} \right] \right]$$

Figure 1: Functional sentence perspective structure in HPSG

Figure 1 shows where focus information is coded in the HPSG structure. An extra attribute INFO-STRUCT, which contains the attributes LINK and FOCUS, is included in the CONTEXT attribute. The attribute LINK contains the complements that ‘link’ the sentence to former sentences or to known information. They are the known entities, about which something is said. The FOCUS attribute contains the new, informative information. Both FOCUS and LINK can take lists of values.

The corpus research has proven that the arguments between first pole and VP adjuncts contain the LINK information, while the arguments between VP adjuncts and second pole contain the FOCUS information². Now that we have a way to code the different positions in HPSG (the division between COMPS arguments and L-ARGS arguments), the next step is to design a *Focus Realization Principle* which assigns the correct INFO-STRUCT values according to the position in the sentence. Figure 2 shows what this principle should look like.

²Next to first sentence position and last sentence position, as explained in 1.2.

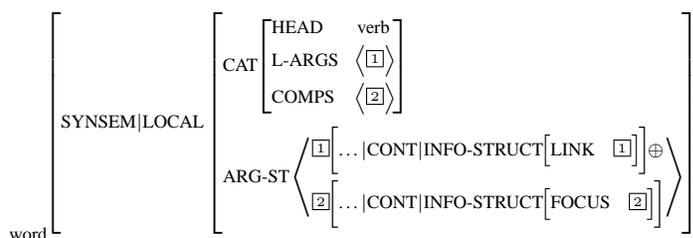


Figure 2: Focus Realization Principle

The information about word order (which determines the focus information) is available in the feature structure of the verb, in the L-ARGS and COMPS list. The Focus Realization Principle makes sure that the arguments which are link or focus, also get this characteristic coded into their feature structures. This is brought about by adding the information to the ARG-ST-list: arguments on the L-ARGS-list become link, arguments on the COMPS-list become focus. Note the fact that the value of LINK and FOCUS is equal to the sign itself.³

The Focus Realization Principle makes sure that the various arguments of the verb signal the right focus information. Now we only have to make sure that this information is passed on to the mother nodes, so that the final root node will also contain the correct focus information. This is done by the *Focus Inheritance Principle* in (28).

(28) **Focus Inheritance Principle**

The INFO-STRUCT value of the mother node is equal to the different INFO-STRUCT values of the child nodes.

This way, the focus information of the sentence is put together correctly.

4.2 An Example

Figure 3 shows the analysis of sentence (29).

- (29) ... dat [ik hem]_L gisteren [een boek]_F gaf.
 that I him yesterday a book gave
 ‘... that I gave him a book yesterday’

The verb *gaf* has in this example two arguments *ik* and *hem* on the L-ARGS-list, and one argument *een boek* on the COMPS-list. When the VP is built up (according to the Argument Realization Principle), the Focus Realization Principle makes sure that the correct focus information is distributed over the various arguments. This way, *een*

³Engdahl and Vallduví note correctly that in this way, semantic information as well as phonological and syntactic information is marked as focus. It would be more correct to make the LINK and FOCUS values equal only to the semantic information of the sign. To keep a clear view, this approach is not elaborated, but we’re assuming that it is done this way.

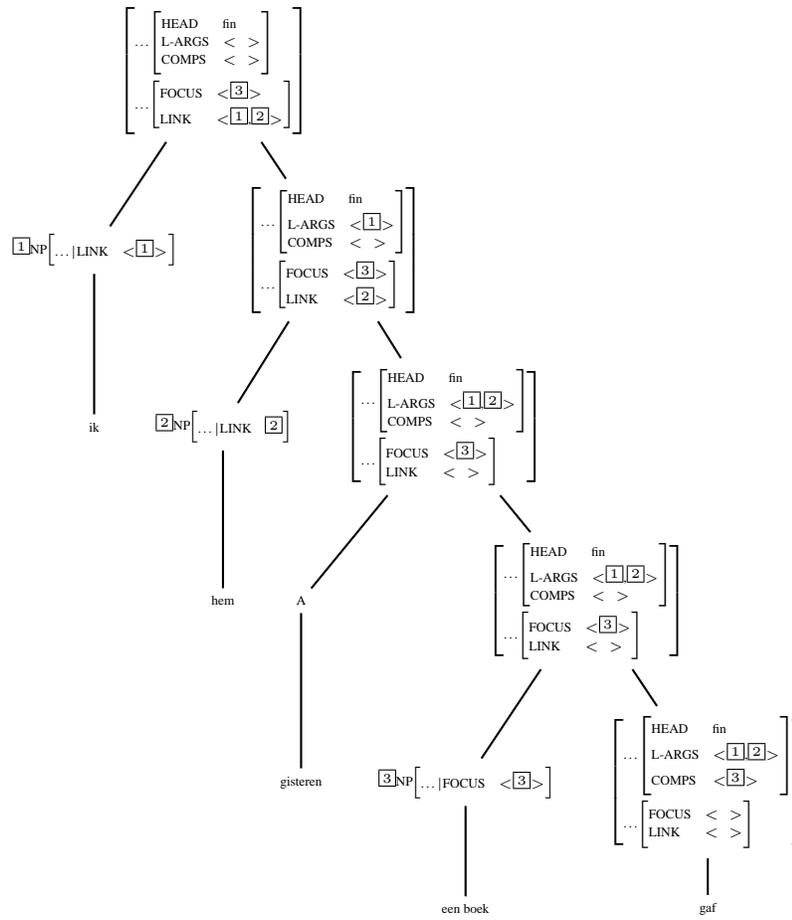


Figure 3: The focus structure of a subclause

boek signals that its meaning must be focus, and *ik* and *hem* signal that their semantic information must be linked to other known information. At the same time, the Focus Inheritance Principle makes sure that the focus information is passed on to the mother nodes, so that the root node of the clause contains the correct focus information.

5 Conclusion and Further Research

This paper has investigated the distribution of the various complements to the verb in a Dutch clause, and the reasons and principles that are responsible for the different distributions of these complements. It has become clear that, for certain complements, semantic (and syntactic) principles play an important role. Predicative complements need to appear close to the main verb because they are semantically linked to it. Also, the preposition of prepositional complements is closely linked to the verb. But these principles alone are not sufficient to explain the distribution of the various complements. The corpus research has clearly indicated that, for other complements, pragmatic principles play an equally important role. Complements that do not have a fixed position in the sentence are distributed according to the functional sentence perspective: unknown, informative information is put behind the VP adjuncts, while known information, that links the unknown information to the speaker's world, appears before the VP adjuncts.

Van Eynde (this volume) provides an HPSG implementation of the semantic and syntactic principles that are important in a Dutch clause. My paper has focused on an implementation of the pragmatic functional sentence perspective. It has been shown by the corpus research that this principle needs to be implemented in the grammar, to be able to describe the formation of Dutch clauses in an adequate way.

At the same time, the description of the pragmatic principles in a Dutch clause is not yet complete. This paper has mainly investigated the position between VP adjuncts and second pole, leaving aside the other focus positions such as first and last sentence position. Also, it needs to be investigated in which way the focus information stemming from word order combines with the focus information that is conveyed by prosodic cues. These topics are to be investigated to get a complete view of how pragmatic principles influence the design of Dutch clauses.

It is only when these pragmatic principles have been investigated and added to the grammar, that a Dutch clause can be analyzed to its full extent.

References

- Engdahl, E.(1999), Integrating pragmatics into the grammar, in L. Mereu (ed.), *Boundaries of Morphology and Syntax*, John Benjamins, Amsterdam.
- Engdahl, E. and Vallduví, E.(1996), Information Packaging in HPSG, *Working Papers in Cognitive Science*.
- Eynde, F. V.(this volume), Argument realization in an SOV language.
- Haeseryn, W., Romijn, K. et al.(1997), *Algemene Nederlandse Spraakkunst*, Martinus Nijhoff uitgevers/Wolters Plantyn, Groningen/Deurne.

- König, E., Lezius, W. and Voormann, H.(2003), *TIGERSearch 2.1 User's Manual*, University of Stuttgart, Stuttgart.
- Sag, I. and Wasow, T.(1999), *Syntactic Theory. A Formal Introduction*, CSLI Publications, Stanford.
- van der Wouden, T., Schuurman, I., Schoupe, M. and Hoekstra, H.(2003), Harvesting Dutch trees: Syntactic properties of spoken Dutch, in T. Gaustad (ed.), *Computational Linguistics in the Netherlands 2002. Selected Papers from the Thirteenth CLIN Meeting*, Rodopi, Amsterdam/New York, pp. 129–141.