

# Argument Realization in Dutch

Frank Van Eynde

Centrum voor Computerlinguïstiek - Universiteit Leuven

## Abstract

The treatment of argument realization is rather straightforward for a language like English, but for a language with relatively free word order, such as Dutch, it is a complex matter. It is not surprising then that the devices which are commonly used to deal with it show a high degree of computational complexity. They typically include movement, as in transformational grammar, or the dissociation of order-in-the-representation from the surface order, as in certain types of monostratal grammar. For the purpose of natural language *description* these devices are certainly convenient, but for the purpose of natural language *processing* they are less attractive. For this reason, I propose an alternative treatment of argument realization, which is consistently monostratal and surface-oriented. Its cornerstone is the GENERALIZED ARGUMENT REALIZATION PRINCIPLE. It is a generalization of the Argument Realization Principle which is proposed in (Ginzburg and Sag 2000) to deal with English.

## 1 The argument realization principle

The lexical entries of argument taking words commonly include information about the number and the kinds of arguments which they select. Intransitive verbs, for instance, select one NP and transitive verbs two. This information is useful for syntactic and semantic processing, on condition that it is complemented with information on how the arguments are realized in sentences. For a language with a relatively rigid word order, such as English, the constraints on argument realization are relatively straightforward. In an active English clause, for instance, the first NP argument is realized as the subject and precedes the verb, whereas the other arguments follow the verb in a fixed order.

- (1) a. They gave her a bike.
- b. \* Gave they her a bike.
- c. \* They her a bike gave.
- d. \* They gave a bike her.

To spell this out in formal terms HEAD-DRIVEN PHRASE STRUCTURE GRAMMAR employs two kinds of features. There is the ARG(UMENT)-ST(RUCTURE) feature, which specifies the syntactic and semantic properties of the arguments which a word selects, and there are the valence features SUBJ(ECT), SP(ECIFIC)R and COMP(LEMENT)S, which spell out how these arguments are realized. The link between them is defined by the ARGUMENT REALIZATION PRINCIPLE (Ginzburg and Sag 2000, 23).<sup>1</sup>

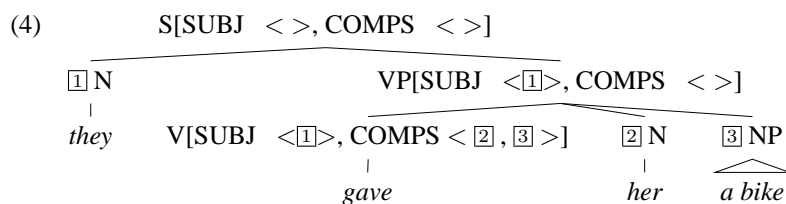
<sup>1</sup>Throughout the text, the boxed alphabetic characters stand for lists of objects, whereas the boxed integers stand for individual objects, usually of type *synsem*.  $\oplus$  is a concatenation operation on lists.

$$(2) \text{ word} \Rightarrow \left[ \begin{array}{l} \text{ARG-ST} \quad \boxed{A} \oplus \boxed{B} \oplus \boxed{C} \\ \text{SYNSEM} \mid \text{LOC} \mid \text{CAT} \quad \left[ \begin{array}{l} \text{SUBJ} \quad \boxed{A} \\ \text{SPR} \quad \boxed{B} \\ \text{COMPS} \quad \boxed{C} \end{array} \right] \end{array} \right]$$

In words, the list of arguments is divided in three sublists; the members of the first sublist are realized as subjects ( $\boxed{A}$ ), those of the second one as specifiers ( $\boxed{B}$ ) and those of the third one as complements ( $\boxed{C}$ ). Subsets of words are associated with more specific constraints. Verbs, for instance, have a singleton SUBJ list and an empty SPR list, which means that they take one subject and no specifier (o.c., 22). When applied to the ditransitive *give*, this yields the following result.

$$(3) \left[ \begin{array}{l} \text{ARG-ST} \quad \langle \boxed{1} \text{ NP}, \boxed{2} \text{ NP}, \boxed{3} \text{ NP} \rangle \\ \text{SYNSEM} \mid \text{LOC} \mid \text{CAT} \quad \left[ \begin{array}{l} \text{SUBJ} \quad \langle \boxed{1} \rangle \\ \text{SPR} \quad \langle \rangle \\ \text{COMPS} \quad \langle \boxed{2}, \boxed{3} \rangle \end{array} \right] \end{array} \right]$$

The first argument is realized as a subject and the remaining arguments as complements. These requirements interact with the constraints on headed phrases: the combination of the verb with its complements is subsumed by the type *head-comps-phrase*, yielding a VP, and the combination of the VP with the subject is subsumed by the type *head-subj-phrase*, yielding a fully saturated clause, as in (4).



The distinction between the subject and the complements is not only motivated by their position (preverbal vs. postverbal), but also by other factors. The subject, for instance, can be separated from the verb by a modifier, but the complements cannot.

- (5) a. They never gave her a bike.  
 b. \* They gave never her a bike.  
 c. \* They gave her never a bike.

Similarly, in the case of raising, it is only the subject of the embedded verb which is affected, not any of its other arguments.

- (6) a. She seems to like her bike.

- b. \* She seems her bike to like.

In sum, the distinction between subject and complements is easy to draw in a language like English. Dutch, by contrast, is a language with relatively free word order, and this seriously complicates the treatment of argument realization. First, there is the variation between SVO and SOV order, which implies that the complements can either follow or precede the verb.

- (7) a. Ze gaven haar een fiets.  
       they gave her a bike  
       b. ... dat ze haar een fiets gaven.  
       ... that they her a bike gave

Second, the mutual order of the arguments is not fixed. In English the indirect object NP invariably precedes the direct object NP, and this is also a possibility in Dutch, as illustrated by (7), but the alternative order is also possible and sometimes obligatory, for instance, when the direct object is a weak pronoun, such as *het* ‘it’.

- (8) We geven het haar morgen.  
       we give it her tomorrow

Third, VP-adjuncts are routinely interleaved with the complements and often separate them from the verb, both in SVO and SOV clauses.

- (9) a. Ze gaven haar elk jaar een nieuwe fiets.  
       they gave her every year a new bike  
       b. ... dat ze haar elk jaar een nieuwe fiets gaven.  
       ... that they her every year a new bike gave

Fourth, raising is not restricted to subjects, as will be illustrated in section 5. The modeling of argument realization is, hence, considerably more complex for Dutch than it is for English. To deal with this complexity there are basically two approaches.

One approach is to separate the issues of argument realization and linear order. More specifically, one can use the same general constraints on argument realization as for English, and add a proviso that they only hold for a level of syntactic structure in which the constituents do not necessarily occur in the surface order. This approach is typical of multistratal models, such as those of transformational grammar, which postulate a level of structure, at which the relation between the selected arguments and their realization is straightforward (deep structure or d-structure), but at which the order of the constituents does not correspond to their linear order in the clause. The resulting gap is bridged by movement operations, such as raising and scrambling. A similar approach is taken in certain variants of monostratal grammar. Dowty’s distinction between tectogrammatical and phenogrammatical structure, for instance, has—in this respect—roughly the same function as the distinction between deep and surface structure, and has inspired a strand of HPSG in which the order of the words in the syntactic representation does not necessarily correspond to their order in the PHONOLOGY value. This approach was pioneered in (Reape 1994) and further developed in (Kathol 2000).

The alternative approach is to modify the constraints on argument realization. For instance, in order to account for the interleaving of arguments and adjuncts in Dutch, (Van Noord and Bouma 1994) adds the adjuncts to the COMPS lists of the verbs, so that there is no longer any need for scrambling operations. The proposal to be presented in this paper is another instance of this approach, but develops it in another direction. Instead of obliterating the distinction between arguments and adjuncts, I will preserve it and argue in favor of making some finer-grained distinction between different kinds of arguments, depending on how they are positioned with respect to the adjuncts. The approach is similar in spirit to the one proposed for Norwegian in (Hellan and Haugereid 2004).

## 2 The generalized argument realization principle

The cornerstone of the proposal is the GENERALIZED ARGUMENT REALIZATION PRINCIPLE (GARP).<sup>2</sup>

$$(10) \text{ word} \Rightarrow \left[ \begin{array}{l} \text{ARG-ST } \boxed{A} \oplus \boxed{B} \\ \text{SYNSEM} \mid \text{LOC} \mid \text{CAT} \left[ \begin{array}{l} \text{L-ARGS } \boxed{A} \\ \text{COMPS } \boxed{B} \end{array} \right] \end{array} \right]$$

In words, the arguments which some given word selects are realized as either complements or l(ef)t-arguments.<sup>3</sup> The complements are those arguments which are adjacent to their head in SOV clauses and which cannot be separated from their head by means of adjuncts. The l(ef)t-arguments, by contrast, are the arguments which can be separated from their head by adjuncts. This can be captured in the formula [L-Arg\* - Adj\* - Comp\* - Head], in which \* is the Kleene star. When applied to (11), it turns out that the subject and the indirect object are l-arguments, since they both precede the adjunct *elk jaar*.

- (11) ... dat ze haar elk jaar een nieuwe fiets gaven.  
 ... that they her every year a new bike gave

If the head takes the leftmost position, as in SVO clauses, the formula reads as follows: [Head - L-Arg\* - Adj\* - Comp\*]. Notice that it is only the position of the head which changes; the mutual order of the arguments and the adjuncts is the same as in the head-final sequence. In (12), for instance, the pronoun *haar* ‘her’ is an l-argument, just as in (11).

- (12) Ze gaven haar elk jaar een nieuwe fiets.  
 they gave her every year a new bike

The term *left-argument* is not only chosen because it applies to the arguments which remain after the complements have been subtracted but also because the l-arguments

<sup>2</sup>This is a preliminary version. The final version will be given in section 5.

<sup>3</sup>For the reasons given in (Van Eynde 2003) I do not employ a separate valence feature for the selection of specifiers. I, hence, drop SPR from the inventory of valence features.

invariably occur to the left of the complements, both in head final and head first constructions.

As compared to the familiar distinction between complements and subjects the one between complements and I-arguments is more general. More specifically, the former can be derived from the later if one adds the constraint that the L-ARGS list contain at most one member. This extra constraint may well be appropriate for some languages, such as English, but for a language like Dutch it is too strict. At the same time, the differentiation between complements and I-arguments in Dutch is not entirely without constraints either. In fact, there is a number of factors which steer the partitioning. One such factor concerns the properties of the selecting words; another one concerns the properties of the arguments themselves, and a third factor concerns the constraints on information packaging. The first two factors are discussed and illustrated in sections 3 and 4 of this paper; the third one is explored in (Van de Cruys this volume).

### 3 Predicate selectors

To illustrate the role of the selecting words in the partitioning of the arguments I take the verbs which select a predicate. Such verbs take at least two arguments. One is the predicate, which denotes a property, and the other one is an NP which denotes the entities to which the property is attributed. The adjectival predicate of the copula in (13), for instance, denotes a property which is attributed to the individual denoted by *Patrick*.

- (13) ... dat Patrick elke morgen voor zes uur wakker is.  
 \* ... dat Patrick wakker elke morgen voor zes uur is  
 ‘... that Patrick is awake before six every morning.’

As illustrated by the starred clause, the predicate cannot be separated from the verb by the temporal adjuncts. This implies that it must be realized as a complement. To express this I introduce a specific lexical type for the words which select a predicate, to be called *pred(icate)-sel(ector)*, and assign it the following properties:

- (14) 
$$\left[ \begin{array}{l} \textit{pred-sel} \\ \text{ARG-ST } \boxed{X} \oplus \langle \boxed{1} \rangle \oplus \langle \boxed{2} \rangle \oplus \boxed{Y} \\ \text{SS | LOC | CAT } \left[ \begin{array}{l} \text{L-ARGS } \boxed{X} \oplus \langle \boxed{1} \text{ NP}_i \rangle \\ \text{COMPS } \langle \boxed{2} [\text{LOC | CONT | NUCL | THEME } i] \rangle \oplus \boxed{Y} \end{array} \right] \end{array} \right]$$

In words, the property denoting argument of a predicate selector must be realized as a complement (2) and the argument to which the property is attributed must be realized as a left argument (1).

Besides the fact that it must be realized as a complement, there are no syntactic constraints on the predicate. Its part of speech, for instance, need not be adjectival, but can also be nominal, prepositional or adverbial, as in (15).

- (15) a. ... dat hij later dokter wordt.  
 \* ... dat hij dokter later wordt  
 ‘... that he’ll become a doctor later.’  
 b. ... dat zijn geduld nu wellicht op geraakt.  
 \* ... dat zijn geduld op nu wellicht geraakt  
 ‘... that his patience is probably running out by now.’  
 c. ... dat hij nog niet weg is.  
 \* ... dat hij weg nog niet is  
 ‘... that he is not away yet.’

The predicate may also be phrasal, as illustrated in (16).

- (16) a. ... dat zijn rapport blijkbaar [vol fouten] staat.  
 \* ... dat zijn rapport [vol fouten] blijkbaar staat  
 ‘... that his report is apparently full of mistakes.’  
 b. ... dat die boeken volgens hem niet [voor kinderen] zijn.  
 \* ... dat die boeken [voor kinderen] volgens hem niet zijn  
 ‘... that those books are not for children according to him.’

In these clauses the predicate denotes a relation between the referent of its external argument and the referent of its own complement. The adjective *vol* in (16a), for instance, denotes a relation between the referents of *zijn rapport* and *fouten*, and the preposition *voor* in (16b) denotes a relation between the referents of *die boeken* and *ons*.

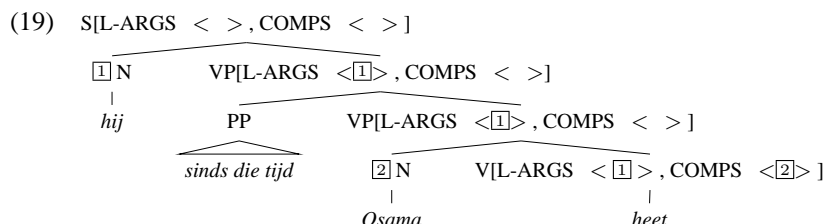
The only restriction on the predicate is a semantic one: it must denote a relation (with arity equal or greater than one) in which the role of theme is realized by the external argument. The relevance of this constraint is illustrated by (17).

- (17) a. ... dat ze zichzelf niet meer zijn.  
 ... that they themselves not more are  
 ‘... that they are not themselves anymore.’  
 b. ... dat dat Osama niet is.  
 ... that that Osama not is  
 ‘... that that is not Osama.’

In contrast to the predicate nominal in (15a), the reflexive pronoun in (17a) and the proper noun in (17b) precede the negation marker, which implies that they are 1-arguments, rather than complements. This, however, is not a violation of the *pred-sel* constraint, for since the NPs in (17) denote individuals, rather than properties or relations, they are not subsumed by the constraint. Further evidence for the semantic nature of the constraint is provided by (18).

- (18) ... dat hij sinds die tijd Osama heet.  
 \* ... dat hij Osama sinds die tijd heet.  
 ‘... that he is called Osama since then.’

In this clause, *Osama* does not denote an individual, but a name, and since names can be attributed to individuals, they are subsumed by the *pred-sel* constraint. This accounts for the fact that *Osama* must precede the temporal adjunct. Employing the distinction between COMPS and L-ARGS, (18) is analyzed as follows:



The verb takes the name as its complement (2) and the subject as its left argument (1). The lower VP has an empty COMPS list and combines with the PP adjunct yielding the higher VP which then combines with the l-argument *hij* ‘he’. In this example, the distinction between complement and left argument coincides with the distinction between complement and subject. This, however, is not always the case. In (11), for instance, the indirect object sides with the l-arguments, and the same holds for the reflexive pronoun and the proper noun in (17).

As for the external argument of the predicate, the only restriction is that it be an NP with a referential index. This NP can be the subject, as in the examples above, but it can also be the direct object, as in (20).

- (20) ... dat ze hem elke morgen voor zes uur wakker maakt.  
 \* ... dat ze hem wakker elke morgen voor zes uur maakt.  
 ‘... that she wakes him before six every morning.’

In this clause, the adjectival complement of the causative verb *maakt* ‘makes’ denotes a property which is attributed to the individual denoted by *hem* ‘him’. To model this instance of object oriented predication we do not need to make any changes to the constraint on predicate selectors, since (14) foresees that the external argument of the predicate may be preceded by other arguments (X). In the case of *maken*, X is a list which contains one NP (the subject); in the case of the copula, it is the empty list.

As the reader can verify, the constraint also foresees the possibility that the predicate complement is followed by other arguments (Y). The relevance of this addition will become clear in the next section.

#### 4 Minor arguments

It is not only the properties of the argument selecting words which influence the partitioning between complements and l-arguments. Equally important are the properties of the selected arguments themselves. To demonstrate this I will now discuss two classes of arguments which share the property that they invariably consist of a single word, i.e. the particles and the weak pronouns.

#### 4.1 Particles

Particles are closely tied to the verb, both semantically and syntactically. In the orthography this is sanctioned by the convention to treat them as verbal prefixes, as in (21a).

- (21) a. ... dat ik ze onmiddellijk opbel.  
 ... that I her immediately upcall  
 '... that I call her immediately.'
- b. Ik bel ze onmiddellijk op.  
 I call her immediately up  
 'I call her immediately.'
- c. ... dat ik ze onmiddellijk op zal bellen.  
 ... that I her immediately up will call  
 '... that I will call her immediately.'

In spite of its intuitive appeal, though, the affixal treatment of the particles is not satisfactory, since it is only applicable in those cases in which the particle happens to immediately precede the verb. If this condition is not fulfilled, as in (21b) and (21c), the particle must be treated as a separate word.<sup>4</sup> As a consequence, since the contribution of the particle is the same in (21a) as it is in (21b) and (21c), it makes more sense to treat it as a separate syntactic unit, also when it is adjacent to the verb.

To model the distribution of the particles, I start from the assumption that they are on the ARG-ST list of the verbs to which they are related. This is motivated a.o. by the fact that their presence may influence the rest of the verb's argument structure. *Lachen* 'laugh', for instance, is intransitive when it is not accompanied by a particle, but when it is accompanied by the particle *uit*, it is transitive. Conversely, *zien* 'see' is transitive when used without a particle, but when it is accompanied by the particle *af*, it has the meaning of suffering and in that sense it is intransitive.

As for the realization of the particles, they invariably occur after the adjuncts, both in SOV and SVO clauses.

- (22) a. ... dat ik ze onmiddellijk op                      zal bellen.  
 \* ... dat ik ze op                      onmiddellijk zal bellen
- b. Ik bel ze onmiddellijk op.  
 \* Ik bel ze op                      onmiddellijk

This demonstrates that they are complements. To express this in formal terms I employ the following constraint:

<sup>4</sup>Some of the particles are homophonous to prefixes, but it is easy to distinguish them, since the prefixes cannot be separated from the verb, whereas the particles can. Compare, for instance, the prefix in *Hij overdrijft altijd* 'he always exaggerates' with the homophonous particle in *Hij steekt de straat over* 'he crosses the street'.



$$(23) \left[ \begin{array}{l} \text{ARG-ST } nelist \oplus \left\langle \boxed{1} \left[ \text{LOC} \mid \text{CAT} \left[ \begin{array}{l} \text{minor} \\ \text{HEAD } \neg p\text{-noun} \end{array} \right] \right] \right\rangle \\ \text{SYNSEM} \mid \text{LOC} \mid \text{CAT} \mid \text{COMPS } \boxed{X} \oplus \left\langle \boxed{1} \right\rangle \end{array} \right]$$

In words, the particles are the most oblique arguments. They are preceded by at least one other argument and must be realized as complements. Besides their rightmost position on the ARG-ST list, there is one other characteristic which distinguishes the particles from the other arguments, i.e. the fact that they are minor. The distinction between major and minor words, introduced in (Van Eynde 1999), differentiates the words which can take local dependents from those which cannot. The former are major and can head a branching XP; the latter are minor and lack this possibility. This differentiates a.o. the prepositions which are used as predicate complements from the homophonous prepositional particles.

- (24) a. ... dat de voorraad nu [helemaal op] is.  
 ... that the stock now [entirely up] is  
 '... that the stock is entirely finished now.'
- b. \*... dat hij je morgen [helemaal op] zal bellen.  
 \*... that he you tomorrow [entirely up] will call

While the prepositional predicate in (24a) can take a modifying adverb and project a PP, the particle in (24b) cannot. Another difference between major and minor words is that the former can be stressed and topicalized, whereas the latter cannot. Compare, for instance, the predicate complement in (25a) with the particle in (25b).

- (25) a. Op is het nog niet, maar we moeten nu wel zuinig zijn.  
 up is it not yet, but we must now economical be  
 'It is not finished yet, but we have to be economical now.'
- b. \*Op moet je hem niet bellen.  
 \*up must you him not call

Since the major/minor distinction is applied to the values of CAT(EGORY), it is orthogonal to the part of speech distinction, and, hence, applicable to all parts of speech. This is the way it should be, since the particles can also take the form of adjectives, common nouns and adverbs.

- (26) a. ... dat ze dit niet (\*bijzonder) goed zal keuren.  
 ... that she this not (\*particularly) good will deem
- b. ... dat het concert vorige week (\*een) plaats had moeten vinden.  
 ... that the concert last week (\*a) place had must find  
 '... that the concert should have taken place last week.'
- c. ... dat hij die bloemkool niet (\*ver) weg had moeten geven.  
 ... that he that cauliflower not (\*far) away had must give  
 '... that he should not have given away that cauliflower.'

The impossibility of adding a dependent to *goed*, *plaats* and *weg* demonstrates that they are minor, and the fact that they cannot be topicalized confirms this.

- (27) a. \* *Goed zal hij het niet keuren.*  
       \* good will he it not deem  
       b. \* *Plaats zal het morgen vinden.*  
       \* place will it tomorrow find  
       c. \* *Weg zal hij het niet geven.*  
       \* away will he it not give

The only part of speech to which the particles cannot belong is the one of the pronouns.<sup>5</sup> A justification for this exclusion will be given in the next section.

If a verb takes both a predicate complement and a particle, the constraint requires that the former precede the latter. This follows from the fact that the particle is the rightmost member of COMPS in (23) and from the fact that the predicate complement can be followed by another complement in the *pred-sel* constraint, see  $\bar{Y}$  in (14). The correctness of this prediction is illustrated by (28).

- (28) ... *dat we hem vanmorgen dood aan hebben getroffen.*  
       \* ... *dat we hem vanmorgen aan dood hebben getroffen*  
       ‘... that we found him dead this morning.’

The adjectival predicate (*dood*) and the particle (*aan*) are both complements of *getroffen* and must occur in that order.<sup>6</sup>

## 4.2 Weak pronouns

Most of the Dutch pronouns come in two kinds. Besides the strong forms, such as *wij* ‘we’ and *dat* ‘that’, which have a clear vowel, there are the weak forms, such as *we* ‘we’ and *het* ‘it’, which have a mute vowel. Syntactically, the strong forms behave like major words, whereas the weak forms behave like minor words. The strong forms, for instance, can take local dependents, whereas the weak forms cannot, as illustrated by the following examples, quoted from (Model 1991, 287).

- (29) *Wij/\*we allen hebben daar aan meegewerkt.*  
       we all have that on collaborated  
       ‘We all have contributed to that.’  
       (30) *Wij/we hebben allen daar aan meegewerkt.*  
       we have all that on collaborated  
       ‘We have all contributed to that.’

<sup>5</sup>The part of speech *p-noun* contrasts with *c-noun*. The former subsumes all of the pronouns and most of the proper nouns; the latter subsumes the common nouns and a subset of the proper nouns. For a motivation of this distinction, see (Van Eynde 2003).

<sup>6</sup>The particles and the predicates are not the only arguments which must be realized as complements. Other such arguments are the NP objects of measure verbs, as in *vijf euro kosten* ‘to cost five euro’ and *drie kilo wegen* ‘to weigh three kilos’, and the arguments which are part of idiomatic expressions, as in *de benen nemen* ‘to take the legs’, i.e. ‘to escape’.

Since verb-second clauses can only have one constituent before the verb, the NP in (29) must be branching and its head must, hence, be a nominal which can take local dependents. This accounts for the fact that the strong form *wij* can be used in this position, whereas the weak form *we* cannot. In (30), however, both forms can be used, for since the quantifier is in postverbal position, the subject NP is nonbranching and may, hence, also consist of a pronoun which cannot take any local dependents. The minor status of the weak pronouns is confirmed by the fact that they cannot be stressed or topicalized.

- (31) Dat/\*het zou ik niet doen.  
 that/\*it would I not do  
 ‘That I wouldn’t do.’

Returning now to the issue of argument realization, there is a conspicuous difference between the major and the minor pronouns. Both tend to precede the VP adjuncts, but while the major ones can also be realized after the adjuncts if they receive stress, the minor ones cannot.

- (32) Ze willen nu eerst en vooral JOU/\*je ondervragen.  
 they want now first and foremost YOU interrogate  
 ‘They first want to interrogate YOU now.’

Arguments which take the form of a minor pronoun must, hence, be realized as left arguments. In other words, they must be assigned to the L-ARGS list of the selecting word.

- $$(33) \left[ \begin{array}{l} \text{ARG-ST } \boxed{X} \oplus \left\langle \boxed{1} \left[ \text{LOC} \mid \text{CAT} \left[ \begin{array}{l} \text{minor} \\ \text{HEAD } p\text{-noun} \end{array} \right] \right] \right\rangle \oplus \boxed{Y} \\ \text{SYNSEM} \mid \text{LOC} \mid \text{CAT} \mid \text{L-ARGS } \boxed{X} \oplus \left\langle \boxed{1} \right\rangle \end{array} \right]$$

There are no constraints on the relative position of the pronoun in the ARG-ST list: it may but need not be preceded by other arguments ( $\boxed{X}$ ) and it may but need not be followed by other arguments ( $\boxed{Y}$ ). There are also no constraints on the part of speech of the selector.

### 4.3 Summing up

Arguments which take the form of minor words are subject to some specific constraints: when they are pronouns, they must be realized as l-arguments, and when they belong to another part of speech they must be realized as complements.

## 5 Argument raising

In all of the examples discussed thus far the argument selecting word was a verb. Verbs, however, are not the only words which select syntactic arguments. Adjectives

and prepositions are argument selectors as well, and are, hence, also subsumed by the GENERALIZED ARGUMENT REALIZATION PRINCIPLE. This implies that their arguments can also be differentiated in complements, on the one hand, and left arguments, on the other hand. The PP argument of the adjective *bewust* ‘aware’ in (34), for instance, is separated from its head by the adverb *zeer* ‘very’, and the pronominal argument of the preposition *onder* ‘under’ in (35) is separated from its head by *vlak* ‘right’.

- (34) ... dat we ons van dat probleem [zeer bewust] waren.  
 ... that we us of that problem [very aware] were  
 ‘... that we were very aware of that problem.’
- (35) ... dat hij daar [vlak onder] stond.  
 ... that he that [right under] stood  
 ‘... that he stood right under that.’

This demonstrates that they are both left-arguments. Moreover, the left arguments of the adjective and the preposition can also be separated from their head by adjuncts which belong to the verb, such as the negation *nog niet* ‘not yet’ in (36) and the temporal *toen* ‘then’ in (37).

- (36) ... dat we ons van dat probleem nog niet bewust waren.  
 ... that we us of that problem still not aware were  
 ‘... that we were not yet aware of that problem then.’
- (37) ... dat hij daar toen vlak onder stond.  
 ... that he that then right under stood  
 ‘... that he then stood right under that.’

To model this, I assume that the L-ARGS requirements of the adjective, c.q. preposition, are inherited by the verb which selects the adjective, c.q. preposition. More specifically, I introduce a lexical type *l-arg-raiser* with the following properties:

- (38) 
$$\left[ \begin{array}{l} l\text{-arg-raiser} \\ \text{ARG-ST } \boxed{X} \oplus \left\langle \left[ \text{LOC} \mid \text{CAT} \mid \text{L-ARGS } \boxed{A} \right] \right\rangle \oplus \boxed{Y} \\ \text{SYNSEM} \mid \text{LOC} \mid \text{CAT} \mid \text{L-ARGS } \boxed{Z} \oplus \boxed{A} \end{array} \right]$$

The words which belong to this type add the L-ARGS requirements of their arguments to their own L-ARGS list. When applied to the VP in (37), this yields the following structure:

- (39) 
$$\begin{array}{c} \text{VP[L-ARGS } \langle \boxed{1}, \boxed{2} \rangle, \text{COMPS } \langle \rangle] \\ \text{Adv} \quad \text{VP[L-ARGS } \langle \boxed{1}, \boxed{2} \rangle, \text{COMPS } \langle \rangle] \\ | \\ \text{toen} \quad \boxed{3} \text{PP[L-ARGS } \langle \boxed{2} \rangle] \quad \text{V[L-ARGS } \langle \boxed{1} \rangle \oplus \langle \boxed{2} \rangle, \text{COMPS } \langle \boxed{3} \rangle] \\ \quad \quad \quad \underbrace{\hspace{2cm}} \quad \quad \quad | \\ \quad \quad \quad \text{vlak onder} \quad \quad \quad \text{stond} \end{array}$$

The verb takes the PP as its complement (③) and inherits its L-ARGS requirement (②), which is appended to its own L-ARGS list (①).<sup>7</sup> This analysis of l-arg raising also accounts for the contrasts in (40).

- (40) a. ... dat ze het hier nog steeds niet gewoon is.  
 \* ... dat ze hier nog steeds niet [het gewoon] is.  
 ‘... that she is still not used to it here.’
- b. ... dat ze er toen nog niet tegen was.  
 \* ... dat ze toen nog niet [er tegen] was.  
 ‘... that she was not yet against it.’

The pronominal arguments of the adjective *gewoon* ‘used to’ and the preposition *tegen* ‘against’ are realized by minor pronouns (*het* and *er*), and must, hence, be realized as l-arguments. Moreover, since the copula is an l-arg-raiser, the L-ARGS list of the predicate is appended to the L-ARGS list of the copula, which implies that the minor pronouns must precede the VP adjuncts.

A consequence of this treatment of argument raising is that a selector may have elements on its L-ARGS list which do not correspond to any of the elements on its ARG-ST list. The PP in (34) and the accusative pronouns in (35) and (40), for instance, are l-arguments of the verb, but do not figure on its ARG-ST list. This implies that we have to revise the formulation of the GENERALIZED ARGUMENT REALIZATION PRINCIPLE as follows:

$$(41) \textit{word} \Rightarrow \left[ \begin{array}{l} \text{ARG-ST} \quad \boxed{A} \oplus \boxed{B} \\ \text{SYNSEM} \mid \text{LOC} \mid \text{CAT} \left[ \begin{array}{l} \text{L-ARGS} \quad \boxed{A} \oplus \boxed{X} \\ \text{COMPS} \quad \boxed{B} \end{array} \right] \end{array} \right]$$

This version allows the L-ARGS list of a selector to include arguments which do not figure on its own ARG-ST list ( $\boxed{X}$ ). This is not only useful to capture the intuition that the raised arguments have another status than the verb’s own arguments, it also provides an account of the contrast between (42) and (43).

- (42) ... dat hij elke morgen (door haar) wakker wordt geschud.  
 ... that he each morning (by her) awake is shaken  
 ‘... that he is shaken awake (by her) every morning.’
- (43) \*... dat deze omgeving niet (door hem) gewoon geraakt wordt.  
 \*... that this environment not (by him) used-to gotten is

As illustrated by (42), the external argument of an object-oriented predicate can become the subject of a passive clause. This can be made more specific in terms of the lexical rule which standard HPSG employs to model passivization. This rule maps the

<sup>7</sup>It may be worth stressing that this analysis does not involve any movement; the pronominal argument of *onder* ‘under’ is not moved out of the PP. Instead, the link between its surface position and its canonical (or semantically ‘natural’) position is defined in terms of the sharing of selection requirements. Constraints on movement, hence, become constraints on the sharing of selection requirements.

stem of a transitive verb onto its passive participle form and reorders its ARG-ST list in such a way that the NP in the first position ( $\langle NP_i, NP_j, \dots \rangle$ ) is demoted to a more oblique position and realized as an optional PP complement ( $\langle NP_j, \dots, (PP[P + NP_i]), \dots \rangle$ ). An automatic consequence of this reshuffling is that the originally second argument becomes the first and gets realized as the subject. Turning now to (43), we can account for its ungrammaticality, if we assume that the NP *deze omgeving* is not on the ARG-ST list of the verb, for in that case it cannot be promoted to its first position. Interestingly, this assumption need not be stipulated anywhere, since it automatically follows from the fact that it is a raised argument, so that it only figures on the L-ARGS list of the verb and not on its ARG-ST list.

Summing up, the distinction between complements and left arguments is not only useful to model linear order in the VP, it also provides a way to identify those arguments which can undergo raising. For a language like English, these arguments only include the subject, since the other arguments are complements, but for a language like Dutch in which most of the arguments are left arguments, raising can be applied to a much wider range of arguments.

## 6 Conclusion

For a language with relatively free word order, such as Dutch, the relation between argument selection and argument realization is considerably more complex than for a language like English (section 1). The devices which are commonly employed to deal with this problem involve movement, as in transformational grammar, or the dissociation of order-in-the-representation from order-in-the-clause, as in certain types of monostratal grammar. While these devices are convenient for the purpose of language description, they are less appealing for the purpose of language processing. From a computational point of view it is more attractive to stick to the monostratal surface-oriented approach of early HPSG. This, I claim, is possible, if one draws the distinction between the different kinds of arguments in another way than is commonly done in current HPSG. Taking the surface order as a criterion, rather than as a distorting nuisance, I replace the distinction between complements and subjects with a more general distinction between complements and left arguments. The resulting treatment of argument realization is modeled in terms of the GENERALIZED ARGUMENT REALIZATION PRINCIPLE (section 2), supplemented with a number of constraints which apply to specific types of argument selecting words (section 3) and specific types of arguments (section 4). Further evidence for the generalized treatment of argument realization is provided by the phenomenon of raising (section 5).

## References

- Ginzburg, J. and Sag, I.(2000), *Interrogative Investigations*, CSLI, Stanford.  
 Hellan, L. and Haugereid, P.(2004), *Norsource - an exercise in the matrix grammar-building design*, NTU Trondheim.  
 Kathol, A.(2000), *Linear Syntax*, Oxford University Press.

- Model, J.(1991), *Grammatische analyse. Syntactische verschijnselen van het Nederlands en Engels*, ICG Publications, Dordrecht.
- Reape, M.(1994), Domain union and word order variation in German, in J. Nerbonne, K. Netter and C. Pollard (eds), *German in HPSG*, CSLI Publications, Stanford, pp. 151–197.
- Van de Cruys, T.(this volume), Between VP adjuncts and second pole in Dutch, University of Leuven.
- Van Eynde, F.(1999), Major and minor pronouns in Dutch, in G. Bouma, E. Hinrichs, G.-J. Kruijff and R. Oehrle (eds), *Constraints and Resources in Natural Language Syntax and Semantics*, CSLI Publications, Stanford, pp. 137–151.
- Van Eynde, F.(2003), Morpho-syntactic agreement and index agreement in Dutch NPs, in T. Gaustad (ed.), *Computational Linguistics in the Netherlands 2002*, Rodopi, Amsterdam - New York, pp. 111–127.
- Van Noord, G. and Bouma, G.(1994), Adjuncts and the processing of lexical rules, *Proceedings of the 15th International Conference on Computational Linguistics (Coling)*, University of Kyoto, Kyoto, pp. 250–256.