

## **Verbs of Perceiving and Verbal Communication in Dutch: Clausal Complementation**

### **Abstract**

The study described in this paper has been undertaken within the framework of the EC DELIS project<sup>1</sup>, combining a corpus-based lexicographical approach and frame-based semantic theory. Syntactic specifications are assigned on two levels: grammatical function and phrase type. In particular, the main verbs of visual, auditive, olfactory and gustatory perception and in addition twelve verbs of verbal communication have been analysed. The focus in this paper is on the clausal complementation of these verbs. In both fields verbs show an interesting variation with respect to their complementation patterns, but they do so for different reasons.

### **1. Introduction**

This section gives an overview of the methods and the kind of descriptive apparatus used in the DELIS encoding of the main perception verbs and some verbs of verbal communication in Dutch. A DELIS dictionary entry combines at least the following elements<sup>2</sup>:

1. A lexical semantic role constellation of the participants involved, based on Fillmore's 'frame semantics' (Fillmore 1993), thus providing a mapping between underlying 'conceptual structure' and the syntactic forms of analysed corpus sentences.

By taking role constellations as a starting point the approach here is onomasiological.

2. A syntactic subcategorization in terms of rather general grammatical functions as well as types of phrase structure. They are of the LFG/HPSG-type. (Pollard/Sag 1994 and Heid 1995.)
3. Optionally additional lexical semantic information may have been added.

A DELIS verb entry, i.e. an identified reading, of a typical two place verb is structured as follows:

There is the lemma and the 'frame element group' (FEG) for each argument consisting of a triple resulting from the inventory of semantic roles (SR), grammatical functions (GF) and phrase types (PT):

LEMMA: "see"

FEG : [ ROLE: [ experiencer ] , GF: [ subject ] , PT: [ np ] ] ,  
       [ ROLE: [ percept ] , GF: [ complement ] , PT: [ np ] ] \_

The descriptive apparatus has been set up for the multilingual description of these verbs. The languages involved were: Danish, Dutch, English, French and Italian and to some extent German (Schwenger 1995). The categories described here have been used in the description of the Dutch verbs. Other or more refined categories may have been used in the description of the other languages. The TFS encoding here ensures the compatibility of the descriptions: by the principle of (multiple) inheritance the transfer of properties from more general to more specific ones is guaranteed.

## 2. Descriptive apparatus

### *Semantic Roles*

Some of the roles are subdivided into subroles, e.g. there is an 'intentional' and a 'non-intentional Experiencer'. In addition to the (sub)roles listed here there are some general roles like Manner, Direction etc. These roles are not specific in both the perception and the communication field:

- Exper(iencer): i(ntentional): *The child watches the duck.*
- Exper(iencer): n(onintentional): *The child sees the duck.*
- Percept:act(ual): *The child hears a duck.*
- Percept:t(ar)g(e)t: *The child looks for the duck.*
- Percept:st(i)m(ulus): *From the sound the child can hear that ...*
- Percept:int(erpretation): *I see that everything is fine.*
- Top(ic): *I heard about the mad cow disease.*
- Mes(sage): *He promised not to be late.*
- Mender: *The prime minister spoke for a while.*
- Receiver: *He spoke to me.*

A more elaborate account of the domain of verbs of communication in terms of roles involved in, as they call it “linguistic action”, but altogether transferrable into the DELIS framework one finds in Dirven et al. (1982:2/3). With respect to sortal information here we refer to metonymic extensions of the types taken up in Copestake & Briscoe (1994). For the Sender one finds f.i. metonymic sense extension: ‘a part of a whole’ i.e. ‘voice’, written/oral texts, organizations/public institutions, locations or scientific branches.

*Sortal restrictions*

In order to cope with sortal information the following raw ontology (Heid 1995) has been used:

sort = entity | proposition, entity = concrete | abstract, concrete = animate | artefact, animate = human | animal, proposition = event | state (the ‘|’ indicates disjunction)

There is of course a rather strong relationship between ‘semantic roles’ and ‘sortal restrictions’, as they both have their origin in verb meaning. ‘Sorts’ or ‘selectional restrictions’ “are just explicit information that the verb supplies about its arguments” (Jackendoff 1990:52). As we have ‘frame elements’ as part of the conceptual structure it is not surprising to have at least some ‘sortal’ information already implicitly represented in the frame elements:

exper-i/n = [human]: a (non-)intentional experiencer is [human]  
 message = [abstract]: realized as np: mostly ‘meta-nouns’ like *story* etc.

I also take the view that ‘proposition’ can be associated with more than one ‘semantic role’ (Jackendoff 1990:49) in the sense that e.g. a ‘percept’ can be realized as a ‘simple’ individuum or as a proposition. In addition it has to be decided whether a given proposition represents an event or not. In order to represent correctly the intricate clausal complementation patterns with perception verbs<sup>3</sup> I will use the boolean features [±F(active)] and [±D(irect)]. ‘Directness’ is expressed by the fact that the same tense in both main and complement clause is obligatory.

*Grammatical functions*

sub(ject) | comp(lement) | xcomp(lement) [C(ontrol): s(ubj) | o(bj)] | adj(unct)

comp: object | i(ndirect) o(bject)

*Phrase types*

np | pp | vp [T(ype):...] | cl(ause) | q(quoted) s(entence)  
 vp[T(ype): b(are)-inf | te-inf | omte-inf]

**3. Dutch verbs and corpora**

The following corpora were used: The Eindhovens Corpus (EC): 750,000 tokens mainly from the sixties: journalism, novels etc.; the NDU/Van Dale (NDU) corpus: journalism; the INL corpus (5MLN): 5,000,000 tokens mainly from journalism in the nineties, including some spoken language. The following verbs have been analysed:

**Verbs of perceiving:** Although olfactive and gustatory verbs have been analysed as well, for reasons of brevity we will restrict ourselves here to the following:

Visual: *zien* (see), *kijken* (look, watch); Auditory: *horen* (hear), *luisteren* (listen).

**Verbs of verbal communication**, including some verbs reflecting the syntactic frame of 'discuss', not having the typical complements of 'speech act verbs' like 'direct speech'. For reasons of brevity we restrict ourselves to those who share the following complementation patterns: te-inf, omte-inf, dat-cl.

*beloven* (promise), *beschuldigen* (accuse), *bevelen* (order), *dreigen* (threaten), *eisen* (demand), *ontkennen* (deny), *toezeggen* (promise), *verwijten* (reproach), *verzoeken* (request), *vragen1* (demand), *vragen2* (invite), *weigeren* (refuse), *zeggen* (say).

**4. Verbs of perceiving**

From table 1 and the examples (1-6) (below) it becomes clear that an alteration of the complement pattern corresponds to a change in meaning. Whereas a 'bare-inf' (s.1a) points to direct perception a 'dat-cl' (s.2a,b) is related to indirect perception as in *see/hear from ... that ...*, even in the sense of 'conclude' (s.4). Also the syntactic difference between a 'bare-inf' (s.1a) and a 'hoe-cl' (s.1b) has its semantic counterpart: the former being not factive may also refer to a merely thought up event or even a hallucination whereas the latter refers to a current event. In cases where

there is not a ‘perception meaning’, f.i. with *see:try* (s.5a,b) it can be seen that there is syntactic variation without semantic counterpart (5). As argued elsewhere (Vliegen 1995) the field of perception verbs is very open to complementation patterns from other domains: compare f.i. the pattern that the verb *sound* takes with that from the field of evaluation in *this sounds great, like a mad cow* (see Heid 1995 for the term *frame blending*). These meaning differences also are nicely reflected in different verb classes in Levin (1993):

- zien: understand: Conjecture verbs* (class 29.5)
- horen: speech: Verbs of Transfer of a Message* (class 37.1)
- zien: try: Psych Verbs* (subclass 31.3: ‘care about/for’ and subclass 32.1 ‘want’)
- luisteren/kijken: Search Verbs* (class 35)

Verb/Reading	Experiencer	Percept, Message, alien
‘( )’: optional element, ‘ ’ disjunction, ‘p:aan’: pp with preposition ‘aan’		
1.horen, zien:direct	n, subj, np	act[-F,+D],xcomp[C:s],vp[T:b-inf] act[+F,+D],comp, hoe-cl
2.horen, zien:indirect	n, subj, np	act, comp, p:aan int[+F,-D], comp, dat-cl  wh-cl
3.horen:speech	n, subj, np	act[hum] comp, p:van/via <sup>4</sup> Message[-F,-D],comp, dat-cl,cl
4.zien:understand	n, subj, np	int[+F,-D],comp,dat-cl wh-cl cl
5.zien:try	i, subj, np	alien, comp, dat-cl <sup>5</sup> alien, xcomp, vp[T:te-inf[sc]]
6.luisteren, kijken: try to hear, look (for)	i, subj, np	tgt[-F,-D],comp, of-cl tgt[+F,-D],comp, wh-cl

Table 1 (‘alien’ indicates that this reading belongs to another not specified verb field.)

**Examples:**

- (1a) Dat hij naar de keuken gegaan is omdat hij het raam  
\*\*hoorde\_mc255 klappen.  
 (“... because he the window rattle hears”)
- (1b) Ik \*\*hoorde\_mc255 duidelijk hoe het schaakbord diepe krassen  
maakte in het koperen tafelblad.  
 (“I heard clearly how the chessboard deep scratches made ...”)

- (1c) ... de twee mannen, die hij zonder kind uit de bosjes \*\*zag\_mc255 terugkomen en in draf naar hun auto zag lopen.  
 (“...the two men, which he without child from the wood saw return and at a trot to their car saw run”)
- (2a) ... aan de geluiden hadden ze \*\*gehoord\_mc216 wat er mis was.  
 (“from the sounds had they heard what was wrong”)
- (2b) ... want aan het gezicht van de jongeman kon ik \*\*zien\_mc210 dat applaus of geen applaus ik toch zou worden afgeranseld.  
 (“..., because on the face of the young man could I see that applause or not I would get beaten up”)
- (3) ... maar toen we eenmaal hadden \*\*gehoord dat u al vier maal de vierdaagse hebt gelopen ...  
 (“... but as we once had heard that you already ... have walked ...”)
- (4) We \*\*zien\_mc254 nu dat Amerika zich met zijn economisch oneindig veel sterkere verkeersluchtvaart ...  
 (“we see now that America ...”)
- (5a) ... tot 'n gewoon verhaal \*\*zien\_mc210 te komen, ...  
 (“to a normal story see to come, ...”)
- (5b) ... we zullen toch \*\*zien\_mc210 dat we 'n voertje hooi op de wagen kennen krijgen, ...  
 (“..., we will nevertheless see that we a load of hay can get, ..”)
- (6a) Zij glimlacht slechts, zij zit zwijgend neer en \*\*luistert\_mc243 naar wat er wordt gezegd.  
 (“she smiles just, she sits wordless down und listens to what is said”)
- (6b) ..., die op instigatie van de burens kwam \*\*kijken\_mc210 of we wel onder de juiste wol lagen ... .  
 (“..., who ... came look whether we ...”)

### 5. Verbs of communication

The following Frame Group Elements were identified:

- I. {Sender, (Receiver), Message}: The Receiver not obligatory.
- II. {Sender, Receiver, (Message)}: Here the Receiver is obligatory, the Message partly is not obligatory.

In the syntactic terminology of generative grammar the object of the main clause c-commands the infinitive clause expressing the Message (van Haften 1991).

For the sake of brevity the levels of Grammatical Function and Phrase Type in table 2 are combined: expressions like 'GF: xcomp[control: subjobj], PT:vp[Type:te-inf]' are combined into 'teinfsc':

Verb/Reading: Sp.act	Se	Receiver	Message
I zeggen 1:ass	np	(iolp:aan)	np DAT qslteinfsc OF
zeggen 2:dir	np	(io)	qslteinfoc
beloven 1:com	np	(iolp:aan)	np DAT qsl(om)teinfsc <sup>6</sup>
bevelen:dir	np	(io)	np DAT qslteinfoc
dreigen 1:com	np	(io)	p:met DAT qslteinfsc
eisen 1:dir	np	(p:van)	np DAT qslteinfoc
ontkennen:ass	np	—	np DAT qslteinfsc
toezeggen:com	np	(iolp:aan)	(np DAT qslteinfsc)
verzoeken:dir	np	(io)	p:om OF qsl(om)teinfoc
vragen 1:dir	np	(iolp:aan)	np p om (om)teinfoc
weigeren 1:ass	np	(iolp:aan)	np (om)teinf(sc)
II beschuldigen:decl	np	np	(p van DAT teinfoc)
verwijten:decl	np	iolp:aan	np DAT qslteinfoc
vragen 2:dir	np	np	(p:voor (om)teinfoc)

Table 2

With respect to this domain I first will make some remarks about the alternation of the 'te-inf' and the 'dat-c', then I will have a closer look at the alternation of 'te-inf' and 'omte-inf': The choice between the first two appears to be motivated syntactically; the 'dat-cl' is found almost exclusively in the following cases:

1. There is no coreference between subject/object of the main clause and the subject of the subclause (no control relation):

(7a) ...beval de Algerijnse regering zeven Iraanse diplomaten het land te verlaten ...  
 (“... ordered the Algerian government seven Iranian diplomats the country to leave”)

2. Verb clustering in the subclause, f.i. with modals:

(7b) ...vader en moeder \*\*zeggen dat ze nooit hadden durven denken dat ...  
 (“... father and mother say that they never had dare think that ...”)

3. If the subclause is fronted (made up example):

(7c) \*De auto te wassen/Dat hij de auto wast heeft Jan beloofd.  
 (“the car to wash/ That he the car washes has Jan promised”)

The alternation between the ‘te-inf’ and the ‘omte-inf’ has been a topic in Dutch grammar for quite some time (Janssen 1993). The Dutch reference grammar ANS (1984:790) points out that *om te* is not accepted with verbs like *mededelen* (inform) or *voelen* (feel) and also that with other verbs *om* in object positions is not obligatory. It is used mainly in spoken, not in written language.

(8) Hij heeft nog geprobeerd (om) de deur open te krijgen.  
 (“He has still tried (for) the door open to get”)

The number of verbs having both these complements is rather small. It is usually said that verbs that can have *om-te* can not be combined with the verb *zullen* (shall). This verb indicates future tense, but it also has modal aspects. In this sample however there are some verbs allowing both possibilities: *beloven*, *toezeggen* and *dreigen*:

(9a) ... en we hebben beloofd om er haar een te brengen.  
 (“.. and we have promised for her one to bring”)

(9b) Een stalknecht nam het paard over en beloofde het flink te zullen afwrijven.  
 (“A stable hand took the horse over and promised it firmly to shall rub off.”)

There are even speakers of Dutch, including the author, who accept the combination of the two (example made up):

- (9c) Hij beloofde (om) de auto te zullen wassen.  
 (“...for the car to shall wash”)

Janssen (1991:338) argues that *om* indicates a ‘hypothetical aim’ and therefore f.i. cannot be combined with past tense in the subclause:

- (10) Zij is bang om zich te verraden/(*\*om*) zich verraden te hebben  
 “She is afraid for herself to betray/for herself betrayed to have”

Now this is very much in line with the verbs analysed here, since they all are commissives or directives: the prospect of action from subject or object in the main clause is being held out but not guaranteed, i.e. hypothetical. Janssen also concludes that *zullen* with these verbs expresses an event presented as real and that therefore in general *zullen* is not possible here. As seen in (9a–c) there are some possibilities. The following explanation at least offers a beginning. The verb *zullen* can be used to express something like a promise (ANS 1984:146). (Compare: “I shall burn this letter immediately”). Whereas *om* expresses the hypothetical aspect, *zullen* here is used to express the commitment of the subject to perform the action. It thus is by no means a coincidence that the verbs that show these possibilities usually have subject control.

## 6. Conclusions

1. If complementation patterns are stored in a lexicographical context one has to be very much aware of the fact that complemental alternation can have very different i.e. semantic, syntactic or pragmatical causes.
2. The realization of these and other complementation patterns can not always be assigned to a lemma as such but rather to the lemma in a specific context.

## Notes

- 1 DELIS stands for "Descriptive Lexical Specifications and tools for corpus-based lexicon building". It was a shared-cost project funded by CEC, Luxembour, (LRE 61.034). All relevant references to the project in Heid (1995).
- 2 The 'typed feature structure (TFS) approach' is used as a lexical representation language in DELIS, but I don't want to go into technicalities here (Emele/Heid 1993).
- 3 Evidence for these features for Dutch (and German) Vliegen (1986).
- 4 p-act[hum] equals p-s(ou)rc(e).
- 5 The possibility of a 'dat-cl' or 'vp[TYPE:te-inf]' doesn't seem to support the the 'Interclausal Relations Hierarchy' as stated in Van Valin/Wilkins (1993:514).
- 6 Some verbs, f.i. *beloven*, by default have subject control, they can however be forced into object control: *Jan belooft mij de auto te mogen wassen* ("Jan promises me the car to can wash").

## References

- Algemene Nederlandse Spraakkunst* 1984. Groningen.
- Copestake, A. and Briscoe, T. 1994. "Semi-Productive Polysemy and Phrasal Tagging", *Esprit BRA 7315 Acquilex-II WP 23*.
- Dirven, R. et. al. 1982. *The scene of linguistic action and its perspectivization by 'speak', 'talk', 'say' and 'tell'*. Amsterdam.
- Emele, M. and Heid, U. 1993. "Formal specification of a typed feature logic based lexical representation language. Deliverable D-V-2 of DELIS.
- Fillmore, Ch. 1993. "Frame Semantics and Perception Verbs", in *Universals in the Lexicon: At the intersection of Lexical Semantic Theories*, Kamp, H. and Pustejovsky, J. (eds.), ms. Dagstuhl.
- van Haaften, T. 1991. *De interpretatie van verzwegen subjecten*. ICG, Dordrecht.
- Heid, U. (Coord.). 1995. A lexicographic and formal decription of the lexical classes of perception and and speech act verbs. Deliverable D-III-1 of DELIS.
- Heid, U. and Krueger, K. 1994: "On the DELIS corpus Evidence Encoding Schema" (CEES). Deliverable D-III- 0 of DELIS.
- Jackendoff, R. 1990. *Semantic Structures*. Cambridge Mass.
- Janssen, T. 1991. "Controle: Een onbeheersbaar onderwerp in de Regeeren-Bindtheorie", *Spektator* 21.4, pp.327-344.
- Levin, B. 1993. *English Verb Classes & Alterations*. Chicago.
- Pollard, C. and Sag, I. 1994. *Head-Driven Phrase Structure Grammar*, Chicago.

- Schwenger, S. 1995. *Korpusbasierte Erstellung von kontrastiven Lexikoneinträgen für Verben der olfaktiven Wahrnehmung*. M.A.th. ms. Stuttgart.
- Vliegen, M. 1988. *Verben der auditiven Wahrnehmung im Deutschen*. Tübingen.
- Vliegen, M. 1995. "Linking and Grinding our Way through". DELIS-internal.
- Vanparijs, J. 1993. *The categorization of illocutinary verbs*. PhD ms. Leuven.
- Verschueren, J. 1985. *What people say they do with words: Prolegomena to an empirical-conceptual approach to linguistic action*. Norwood N.J.
- Van Valin, R. and Wilkins, D. 1993. "Predicting Syntactic Structure from Semantic Representations: 'Remember' in English and its equivalents in Mparntwe Arrernte", in *Advances in Role and Reference Grammar*, van Valin, R. (ed.), Amsterdam, pp.499–534.